Methods in
Molecular Biology 824

Springer Protocols

Argelia Lorence *Editor*

Recombinant Gene Expression

Reviews and Protocols

Third Edition

Humana Press

Methods in Molecular Biology™

Series Editor
John M. Walker
School of Life Sciences
University of Hertfordshire
Hatfield, Hertfordshire, AL10 9AB, UK

For further volumes:
http://www.springer.com/series/7651

Recombinant Gene Expression

Reviews and Protocols, Third Edition

Edited by

Argelia Lorence

Arkansas Biosciences Institute, Arkansas State University, Jonesboro, AR, USA

Editor
Argelia Lorence, Ph.D.
Arkansas Biosciences Institute and
Department of Chemistry & Physics
Arkansas State University, Jonesboro
University Loop E. 504
State University, AR 72401, USA
alorence@astate.edu

ISSN 1064-3745 e-ISSN 1940-6029
ISBN 978-1-61779-432-2 e-ISBN 978-1-61779-433-9
DOI 10.1007/978-1-61779-433-9
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2011943340

© Springer Science+Business Media, LLC 2012
All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Humana Press, c/o Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.
The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed on acid-free paper

Humana Press is part of Springer Science+Business Media (www.springer.com)

Dedication

To David and Noah, my two big loves

Preface

Formidable progress in the field of recombinant gene expression has taken place since 2004, the date of publication of the previous edition of this book. In particular, the emergence of the "omics" technologies has revolutionized all areas of biology. In the industrial arena, *Escherichia coli*, *Saccharomyces cerevisiae*, and insect cells continue to be the dominant production platforms of recombinant proteins. However, in the last few years plants and animals have grown in importance as viable sources of more complex proteins.

This third edition of *Recombinant Gene Expression* is a lot more than just an update of the previous edition. Although some of the authors that contributed in 2004 were also invited to participate in this new project, this volume contains brand new protocols and topics not covered before.

I am indebted to the experts in the field and their students and post-doctoral associates whose talent and experience is reflected in the outstanding quality of the chapters here included. I would like to thank Dr. Paulina Balbás, from whom I have learned a lot about managing a project like this, and two members of my laboratory, Austin Slaven and Gwendolyn Wilson who helped me edit the reference section of each chapter.

While organizing a book of such an extensive topic as gene expression, it was indispensable to pick and choose from the multitude of strategies, vectors, promoters, and so on, so the coverage of topics is far from exhaustive. Some expression systems were omitted because of size limitations and even within areas presented unavoidably; some research approaches were unevenly treated.

The information provided in *Recombinant Gene Expression*, is organized in sections by biological host: Bacteria, lower eukaryotes, fungi, plants and plant cells, and animals and animal cells, presenting one or two authoritative reviews and several protocol chapters in each section. Each chapter concludes with a section containing excellent notes where authors offer their valuable expertise of scientists and their personal views of strategy planning, as well as a variety of approaches, and alternatives that will surely be useful and inspiring to you, the reader.

Jonesboro, AR, USA ***Argelia Lorence***

Contents

Preface . . . . *vii*
Contributors . . . . *xiii*

PART I GENERAL ASPECTS

1 Using Folding Promoting Agents in Recombinant Protein Production: A Review. 3
Beatrix Fahnert

2 Routine Identity Confirmation of Recombinant Proteins by MALDI-TOF Mass Spectrometry 37
Brett J. Savary and Prasanna Vasu

3 A Matter of Packaging: Influence of Nucleosome Positioning on Heterologous Gene Expression 51
María de la Cruz Muñoz-Centeno, Gonzalo Millán-Zambrano, and Sebastián Chávez

4 Tools of the Trade: Developing Antibody-Based Detection Capabilities for Recombinant Proteins. 65
Maureen C. Dolan, Giuliana Medrano, Anthony McMickle, and Carole L. Cramer

PART II PROKARYOTES

5 Heat-Shock Protein Fusion Vectors for Improved Expression of Soluble Recombinant Proteins in *Escherichia coli*. 109
Christos A. Kyratsous and Christos A. Panagiotidis

6 The Use of a Flagellar Export Signal for the Secretion of Recombinant Proteins in *Salmonella*. 131
Ferenc Vonderviszt, Ráchel Sajó, József Dobó, and Péter Závodszky

7 Optimization of Purification Protocols Based on the Step-by-Step Monitoring of the Protein Aggregates in Soluble Fractions 145
Ario de Marco

8 Heterologous Protein Expression by *Lactococcus lactis*. 155
Julio Villatoro-Hernández, Oscar P. Kuipers, Odila Saucedo-Cárdenas, and Roberto Montes-de-Oca-Luna

9 An Extended Suite of Genetic Tools for Use in Bacteria of the *Halomonadaceae*: An Overview. 167
Montserrat Argandoña, Carmen Vargas, Mercedes Reina-Bueno, Javier Rodríguez-Moya, Manuel Salvador, and Joaquín J. Nieto

10 Regulated Recombinant Protein Production in the Antarctic Bacterium *Pseudoalteromonas haloplanktis* TAC125 203
Valentina Rippa, Rosanna Papa, Maria Giuliani, Cinzia Pezzella, Ermenegilda Parrilli, Maria Luisa Tutino, Gennaro Marino, and Angela Duilio

11 A Novel Strategy for the Construction of Genomic Mutants of the Antarctic Bacterium *Pseudoalteromonas haloplanktis* TAC125 219
Maria Giuliani, Ermenegilda Parrilli, Cinzia Pezzella, Valentina Rippa, Angela Duilio, Gennaro Marino, and Maria Luisa Tutino

12 A New Bacterial Co-expression System for Over-expressing Soluble Protein and Validating Protein–Protein Interaction 235
Jumei Zeng and Zheng-Guo He

13 Heterologous High-Level Gene Expression in the Photosynthetic Bacterium *Rhodobacter capsulatus* 251
Nadine Katzke, René Bergmann, Karl-Erich Jaeger, and Thomas Drepper

14 Plasmid DNA Production for Therapeutic Applications 271
Alvaro R. Lara and Octavio T. Ramírez

Part III Lower Eukaryotes

15 Recombinant Protein Production in the Eukaryotic Protozoan Parasite *Leishmania tarentolae*: A Review 307
Tomoaki Niimi

16 Expression of Multisubunit Proteins in *Leishmania tarentolae* 317
Marisa Sugino and Tomoaki Niimi

Part IV Fungi

17 Recombinant Protein Production in Yeasts 329
Diethard Mattanovich, Paola Branduardi, Laura Dato, Brigitte Gasser, Michael Sauer, and Danilo Porro

18 Yeasts as a Tool for Heterologous Gene Expression 359
Raja Mokdad-Gargouri, Salma Abdelmoula-Soussi, Nadia Hadiji-Abbès, Ines Yacoubi-Hadj Amor, Istabrak Borchani-Chabchoub, and Ali Gargouri

19 The Cre/Lox System: A Practical Tool to Efficiently Eliminate Selectable Markers in Fungal Endophytes 371
Simona Florea, Caroline Machado, Kalina Andreeva, and Christopher L. Schardl

20 Aptamer-Regulated Expression of Essential Genes in Yeast 381
Beatrix Suess, Karl-Dieter Entian, Peter Kötter, and Julia E. Weigand

21 Cloning and Expression of Hemicellulases from *Aspergillus nidulans* in *Pichia pastoris* 393
Prasanna Vasu, Stefan Bauer, and Brett J. Savary

22 A Thiamine-Regulatable Epitope-Tagged Protein Expression System in Fission Yeast 417
Tiina Tamm

23 Heterologous Gene Expression by Chromosomal Integration in Fission Yeast 433
Akihisa Matsuyama and Minoru Yoshida

24 Genetic Engineering of Industrial Strains of *Saccharomyces cerevisiae* 451
Sylvie Le Borgne

PART V PLANTS AND PLANT CELLS

25 Recombinant Protein Production in Plants: Challenges and Solutions 469
Elizabeth E. Hood and Deborah V. Vicuna Requesens

26 A Novel Plant Cell Bioproduction Platform for High-Yield Secretion of Recombinant Proteins 483
Jianfeng Xu and Marcia J. Kieliszewski

27 Super-promoter:TEV, a Powerful Gene Expression System for Tobacco Hairy Roots 501
Luis Ñopo, Bonnie J. Woffenden, Deborah G. Reed, Scott Buswell, Chenming Zhang, and Fabricio Medina-Bolivar

28 Bioseparation of Recombinant Proteins from Plant Extract with Hydrophobin Fusion Technology 527
Jussi J. Joensuu, Andrew J. Conley, Markus B. Linder, and Rima Menassa

29 Quality Assessment of Recombinant Proteins Produced in Plants 535
Giuliana Medrano, Maureen C. Dolan, Jose Condori, David N. Radin, and Carole L. Cramer

30 Cell-Free Protein Synthesis as a Promising Expression System for Recombinant Proteins 565
Xumeng Ge and Jianfeng Xu

PART VI ANIMALS AND ANIMAL CELLS

31 The Use of Bacterial Artificial Chromosomes for Recombinant Protein Production in Mammalian Cell Lines 581
Leander Blaas, Monica Musteanu, Beatrice Grabner, Robert Eferl, Anton Bauer, and Emilio Casanova

32 Engineering the Chaperone Network of CHO Cells for Optimal Recombinant Protein Production and Authenticity 595
Lyne Jossé, C. Mark Smales, and Mick F. Tuite

33 High-Throughput Baculovirus Expression in Insect Cells 609
Richard B. Hitchman, Robert D. Possee, and Linda A. King

34 Recombinant Protein Expression in Milk of Livestock Species 629
Zsuzsanna Bösze and László Hiripi

Index *643*

Contributors

SALMA ABDELMOULA-SOUSSI • *Centre de Biotechnologie de Sfax, Sfax, Tunisia*
INES YACOUBI-HADJ AMOR • *Centre de Biotechnologie de Sfax, Sfax, Tunisia*
KALINA ANDREEVA • *Department of Plant Pathology, UK-Advanced Genetic Technologies Center, University of Kentucky, Lexington, KY, USA*
MONTSERRAT ARGANDOÑA • *Department of Microbiology and Parasitology, University of Seville, Seville, Spain*
ANTON BAUER • *Ludwig Boltzmann Institute for Cancer Research, Vienna, Austria*
STEFAN BAUER • *Energy Biosciences Institute, University of California, Berkeley, CA, USA*
RENÉ BERGMANN • *Institute of Molecular Enzyme Technology, Heinrich-Heine-University, Juelich, Germany*
ZSUZSANNA BÖSZE • *Agricultural Biotechnology Center, Gödöll, Hungary*
LEANDER BLAAS • *Ludwig Boltzmann Institute for Cancer Research, Vienna, Austria*
ISTABRAK BORCHANI-CHABCHOUB • *Centre de Biotechnologie de Sfax, Sfax, Tunisia*
PAOLA BRANDUARDI • *Department of Biotechnology and Bioscience, University of Milano-Bicocca, Milan, Italy*
SCOTT BUSWELL • *Department of Plant Pathology, Physiology and Weed Science, Virginia Tech, Blacksburg, VA, USA*
EMILIO CASANOVA • *Ludwig Boltzmann Institute for Cancer Research, Vienna, Austria*
SEBASTIÁN CHAVEZ • *Departamento de Genética, Universidad de Sevilla, Sevilla, Andalucía, Spain*
JOSE CONDORI • *Arkansas Biosciences Institute, Arkansas State University and BioStrategies LC, State University, AR, USA*
ANDREW J. CONLEY • *VTT Technical Research Centre of Finland, Tietotie, Finland*
CAROLE L. CRAMER • *Arkansas Biosciences Institute, Arkansas State University and BioStrategies LC, State University, AR, USA*
LAURA DATO • *Department of Biotechnology and Biosciences, University of Milano-Bicocca, Milan, Italy*
JÓZSEF DOBÓ • *Institute of Enzymology, Hungarian Academy of Sciences, Budapest, Hungary*
MAUREEN C. DOLAN • *Arkansas Biosciences Institute, Arkansas State University, State University, AR, USA; Applied Food Technologies Inc., Alachua, FL, USA*
THOMAS DREPPER • *Institute of Molecular Enzyme Technology, Heinrich-Heine-University, Juelich, Germany*
ANGELA DUILIO • *Department of Organic Chemistry and Biochemistry, Università degli studi di Napoli Federico II, Naples, Italy*
ROBERT EFERL • *Ludwig Boltzmann Institute for Cancer Research, Vienna, Austria*
KARL-DIETER ENTIAN • *Department of Molecular Biosciences, Johann Wolfgang Goethe-University Frankfurt, Frankfurt, Germany*

Beatrix Fahnert • *Cardiff School of Biosciences, Cardiff University, Wales, UK*
Simona Florea • *Department of Plant Pathology, UK-Advanced Genetic Technologies Center, University of Kentucky, Lexington, KY, USA*
Sylvie Le Borgne • *Departamento de Procesos y Tecnología, Universidad Autónoma Metropolitana-Cuajimalpa, Mexico City, Mexico*
Ali Gargouri • *Centre de Biotechnologie de Sfax, Sfax, Tunisia*
Brigitte Gasser • *Department of Biotechnology, University of Natural Resources and Life Sciences, and Austrian Centre of Industrial Biotechnology, Vienna, Austria*
Xumeng Ge • *Arkansas Biosciences Institute, Arkansas State University, Jonesboro, AR, USA*
Maria Giuliani • *Department of Organic Chemistry and Biochemistry, Università degli studi di Napoli Federico II, Naples, Italy*
Beatrice Grabner • *Ludwig Boltzmann Institute for Cancer Research, Vienna, Austria*
Nadia Hadiji-Abbès • *Centre de Biotechnologie de Sfax, Sfax, Tunisia*
Zheng-Guo He • *Huazhong Agricultural University, Wuhan, China*
Laszló Hiripi • *Agricultural Biotechnology Center, Gödöll, Hungary*
Richard B. Hitchman • *Oxford Expression Technologies Ltd, Oxford Brookes University, Oxford, UK*
Elizabeth E. Hood • *Arkansas Biosciences Institute, Arkansas State University, Jonesboro, AR, USA*
Karl-Erich Jaeger • *Institute of Molecular Enzyme Technology, Heinrich-Heine-University, Juelich, Germany*
Jussi J. Joensuu • *VTT Technical Research Centre of Finland, Tietotie, Finland*
Lyne Jossé • *University of Kent, Kent, UK*
Nadine Katzke • *Institute of Molecular Enzyme Technology, Heinrich-Heine-University, Juelich, Germany*
Marcia J. Kieliszewski • *Department of Chemistry and Biochemistry, Ohio University, Athens, OH, USA*
Linda A. King • *School of Life Sciences, Oxford Brookes University, Oxford, UK*
Peter Kötter • *Department of Molecular Biosciences, Johann Wolfgang Goethe-University Frankfurt, Frankfurt, Germany*
Oscar Paul Kuipers • *Department of Molecular Genetics, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, The Netherlands; Kluyver Centre for Industrial Fermentation, Delft, The Netherlands*
Christos A. Kyratsous • *Regeneron Pharmaceuticals Inc., Tarrytown, NY, USA*
Alvaro R. Lara • *Departamento de Procesos y Tecnología, Universidad Autónoma Metropolitana-Cuajimalpa, Mexico*
Markus B. Linder • *VTT Technical Research Centre of Finland, Tietotie, Finland*
Caroline Machado • *Department of Plant Pathology, UK-Advanced Genetic Technologies Center, University of Kentucky, Lexington, KY, USA*
AriodeMarco • *University of Nova Gorica, Nova Gorica, Slovenia*
Gennaro Marino • *Department of Organic Chemistry and Biochemistry, Università degli studi di Napoli Federico II, Complesso Universitario M.S. Angelo, Naples, Italy*
Diethard Mattanovich • *Department of Biotechnology, University of Natural Resources and Life Sciences, and Austrian Centre of Industrial Biotechnology, Vienna, Austria*

Akihisa Matsuyama • *RIKEN Advanced Science Institute, Wako, Japan*
Anthony McMickle • *Arkansas Biosciences Institute, Arkansas State University, Jonesboro, AR, USA*
Fabricio Medina-Bolivar • *Arkansas Biosciences Institute, Arkansas State University, Jonesboro, AR, USA*
Giuliana Medrano • *Arkansas Biosciences Institute, Arkansas State University, State University, AR, USA*
Rima Menassa • *Agriculture and Agri-Food Canada, London, ON, Canada*
Gonzalo Millán-Zambrano • *Departamento de Genética, Universidad de Sevilla, Sevilla, Andalucía, Spain*
Raja Mokdad-Gargouri • *Centre de Biotechnologie de Sfax, Sfax, Tunisia*
Roberto Montes de Oca Luna • *Facultad de Medicina, Universidad Autónoma de Nuevo León, Monterrey, Mexico*
Maria De La Cruz Muñoz-Centeno • *Departamento de Genética, Universidad de Sevilla, Sevilla, Andalucía, Spain*
Monica Musteanu • *Ludwig Boltzmann Institute for Cancer Research, Vienna, Austria*
Joaquin J. Nieto • *Department of Microbiology and Parasitology, University of Sevilla, Seville, Spain*
Tomoaki Niimi • *Department of Bioengineering Sciences, Nagoya University, Nagoya, Japan*
Luis Ñopo • *Arkansas Biosciences Institute, Arkansas State University, Jonesboro, AR, USA*
Christos A. Panagiotidis • *Department of Pharmacy, Aristotle University, Thessaloniki, Greece*
Ermenegilda Parrilli • *Department of Organic Chemistry and Biochemistry, Università degli studi di Napoli Federico II, Naples, Italy*
Rosanna Papa • *Department of Organic Chemistry and Biochemistry, Università degli studi di Napoli Federico II, Naples, Italy*
Cinzia Pezzella • *Department of Organic Chemistry and Biochemistry, Università degli studi di Napoli Federico II, Naples, Italy*
Danilo Porro • *Department of Biotechnology and Biosciences, University of Milano-Bicocca, Milan, Italy*
Robert D. Possee • *Center for Hydrology and Ecology, NERC Institute of Virology and Environmental Microbiology, Oxford, UK*
David N. Radin • *BioStrategies LC, State University, AR, USA*
Octavio T. Ramirez • *Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Mexico*
Deborah G. Reed • *Department of Plant Pathology, Physiology and Weed Science, Virginia Tech, Blacksburg, VA, USA*
Mercedes Reina-Bueno • *Department of Microbiology and Parasitology, University of Seville, Seville, Spain*
Deborah V. Vicuña Requesens • *Arkansas Biosciences Institute, Arkansas State University, Jonesboro, AR, USA*
Valentina Rippa • *Department of Organic Chemistry and Biochemistry, Università degli studi di Napoli Federico II, Naples, Italy*

JAVIER RODRIGUEZ-MOYA • *Department of Microbiology and Parasitology, University of Seville, Seville, Spain*
ODILA SAUCEDO-CÁRDENAS • *Centro de Investigación Biomédica del Noreste, IMSS and Facultad de Medicina, Universidad Autónoma de Nuevo León, Monterrey, Mexico*
RÁCHEL SAJÓ • *Institute of Enzymology, Hungarian Academy of Sciences, Budapest, Hungary*
MANUEL SALVADOR • *Department of Microbiology and Parasitology, University of Seville, Seville, Spain*
MICHAEL SAUER • *Department of Biotechnology, University of Natural Resources and Life Sciences, and Austrian Centre of Industrial Biotechnology, Vienna, Austria*
BRETT J. SAVARY • *Arkansas Biosciences Institute, Arkansas State University, Jonesboro, AR, USA*
CHRISTOPHER L. SCHARDL • *Department of Plant Pathology, UK-Advanced Genetic Technologies Center, University of Kentucky, Lexington, KY, USA*
C. MARK SMALES • *School of Biosciences, University of Kent, Kent, UK*
BEATRIX SUESS • *Department of Molecular Biosciences, Johann Wolfgang Goethe-University Frankfurt, Frankfurt, Germany*
MARISA SUGINO • *Graduate School of Bioagricultural Sciences, Nagoya University, Nagoya, Japan*
TIINA TAMM • *Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia*
MICK F. TUITE • *School of Biosciences, University of Kent, Kent, UK*
MARIA LUISA TUTINO • *Department of Organic Chemistry and Biochemistry, Università degli studi di Napoli Federico II, Naples, Italy*
CARMEN VARGAS • *Department of Microbiology and Parasitology, University of Seville, Seville, Spain*
PRASANNA VASU • *Food Safety and Analytical Quality Department, Central Food Technological Research Institute (CSIR), Cheluvamba Mansion, Mysore, India*
JULIO VILLATORO-HERNÁNDEZ • *Department of Molecular Genetics, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, The Netherlands; Kluyver Centre for Industrial Fermentation, Delft, The Netherlands*
FERENC VONDERVISZT • *Research Institute of Chemical and Process Engineering, Faculty of Information Technology, University of Pannonia, Veszprém, Hungary*
JULIA E. WEIGAND • *Department of Molecular Biosciences, Johann Wolfgang Goethe-University Frankfurt, Frankfurt, Germany*
BONNIE J. WOFFENDEN • *Department of Plant Pathology, Physiology and Weed Science, Virginia Tech, Blacksburg, VA, USA*
JIANFENG XU • *Arkansas Biosciences Institute, Arkansas State University, State University, AR, USA*
MINORU YOSHIDA • *RIKEN Advanced Science Institute, Wako, Japan*
PÉTER ZÁVODSZKY • *Institute of Enzymology, Hungarian Academy of Sciences, Budapest, Hungary*
JUMEI ZENG • *Huazhong Agricultural University, Wuhan, China*
CHENMING ZHANG • *Department of Biological Systems Engineering, Virginia Tech, Blacksburg, VA, USA*

Part I

General Aspects

Chapter 1

Using Folding Promoting Agents in Recombinant Protein Production: A Review

Beatrix Fahnert

Abstract

Recombinant production has become an invaluable tool for supplying research and therapy with proteins of interest. The target proteins are not in every case soluble and/or correctly folded. That is why different production parameters such as host, cultivation conditions and co-expression of chaperones and foldases are applied in order to yield functional recombinant protein. There has been a constant increase and success in the use of folding promoting agents in recombinant protein production. Recent cases are reviewed and discussed in this chapter. Any impact of such strategies cannot be predicted and has to be analyzed and optimized for the corresponding target protein. The in vivo effects of the agents are at least partially comparable to their in vitro mode of action and have been studied by means of modern systems approaches and even in combination with folding/activity screening assays. Resulting data can be used directly for experimental planning or can be fed into knowledge-based modelling. An overview of such technologies is included in the chapter in order to facilitate a decision about the potential in vivo use of folding promoting agents.

Key words: Recombinant protein, Expression, Folding promoting agents, Chemical chaperones, Osmolytes, Solutes

1. Introduction

The first part of the chapter introduces the role and action of folding promoting agents in cells in general and in recombinant protein production in particular. The second part gives a detailed overview of their use and discusses findings of modern systems approaches as well as screening assays for protein folding/activity. Thus readers new to the field will easily find some information about how to get started. The experienced readers might appreciate the amount of data to choose from, while assessing and planning new experiments, and related suggestions to reduce trial and error.

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_1, © Springer Science+Business Media, LLC 2012

Over the last three decades recombinant protein production has become a pre-requisite for both basic and applied research by providing adequate amounts of the proteins for studying their structure and function. Moreover, scaled-up facilities produce recombinant vaccines, hormones, antibodies, growth factors, blood components and enzymes. Yet many target proteins are not soluble (every second recombinant protein on average when over-expressed in *Escherichia coli*) and/or correctly folded in the first attempt, and the strategies used to tackle this problem are contextually determined. We have protein structure centres and structural genomics projects on one hand with high throughput (HT) needs and shot gun approaches (e.g., "consensus" advice based on analyses of more than 10,000 different proteins). On the other hand the development of bespoke production schemes is warranted for specific products, crucial applications or profitable markets. Findings can be mutually informative (1–4).

Various strategies including host, induction conditions, temperature, compartment, proteinaceous fusion partners, co-expression of chaperones and foldases are embarked on in order to gain functional recombinant protein. As proteins vary physicochemically, a successful outcome is difficult to predict and based on case-by-case optimization. More comparative and interdisciplinary research has led to a better understanding of the complex networked effects on responsive elements in the hosts and the related impact on the target protein (e.g., heat shock protein involvement in different stress responses) (5–7). Applying the findings of systems biology, protein folding screening assays, in vitro studies and knowledge-based modelling will be invaluable for making recombinant protein production more predictable in the future.

Recently the use of fusion proteins to promote folding by adding predictable features to a target protein has been very popular. Despite constantly emerging fusion partners (some are screened for or designed to have certain features) they are not a "magic bullet" since they affect folding, have to be removed and do not cater well for HT needs (8–10). Thus it is worth changing the folding environment in the cell directly or indirectly to prevent aggregation of folding intermediates and to support them in reaching the stable native conformation. Adding folding promoting agents during cultivation has been successful in that respect.

Again the outcome of any such cultivation strategy is not predictable, and access to information about protocols, the folding of the target protein, the host cell and all interactions is crucial for a more informed experimental planning. Substances and protocols instrumental in in vitro refolding have been found worth testing in vivo with prokaryotes (e.g., *E. coli* as recombinant host (11)) and eukaryotes (e.g., animal cells in the context of diseases (12), yeast as recombinant host (13)). The outer membrane of *E. coli* is highly permeable for molecules smaller than 600 Da (14). That is why agents added

to the medium affect the prokaryotic periplasm more directly than the cytoplasm. The pH of the periplasmic space varies in response to the extra-cellular fluid (15). The gel-like consistency leads to slower diffusion (16) and concentration gradients can be used. So the periplasm can be seen as a test-tube of 50 nm width (about 30% of the cell volume) (17). Diffusion of small molecules further into the cytoplasm is possible and can be enhanced under osmotic stress, leading to changes in molecular interactions (18). Small, charged reagents can also passively enter the Endoplasmic Reticulum (ER) of eukaryotic hosts, even at low temperature (19). The added solutes can stabilize membranes under stress conditions, but can also lead to loss of viability (e.g., ectoines fluidize lipids) after a period of beneficially increased uptake (20–22).

Sufficient transporter expression (potentially over-expression) needs to be allowed as well as binding of the agent within the periplasm and/or further transport to avoid bottlenecks. Knowledge of regulatory processes is essential. External osmolality can be increased/adjusted to, e.g., regulate the activity of the glycine betaine (GB) carrier BetP of *Corynebacterium glutamicum*, which also responds to the internal potassium concentration (23, 24). The ProP transporter can be activated within a minute (25). *Bacillus subtilis* has five osmotically regulated transporters to allow uptake of various compatible solutes (26). Substances entering the cell can accumulate to concentrations comparable to in vitro studies. Reduced glutathione (GSH) (27) can enter the periplasm, whereas sucrose cannot (6). Glycerol can accumulate in the cell up to molar concentrations (28) and GB to 50 mM–1 M (29, 30). Not only does this allow comparison with their in vitro modes of action, but additionally leads to beneficial cellular protein-protective stress responses, such as the activation of molecular chaperones (28, 31) in case of sucrose (32) or ethanol (33). Host strain engineering can be used to further optimize the synergy, e.g., by over-production of endogenous trehalose to improve tolerance of *E. coli* to certain osmotic stresses (34).

The more direct effect of the additives on the recombinant protein and on the host proteins has to be seen in the physicochemical context of protein folding. Every change in the thermodynamic parameters of the environment of a protein potentially affects its configuration. Inter- and intra-molecular interactions between residues up to structural elements follow pathways leading to either the correct or incorrect conformation. In vivo, cellular proteins support nascent proteins in adopting their native state. Although there are differences between the folding in the crowded cytosol of a host cell and the optimized in vitro folding, the mechanisms seem to be comparable (35). Thus the effects of low molecular weight additives can be specified and will be discussed here.

1.1. Effects of Media

Manipulating the recombinant protein production starts with the culture medium. Either complex media or defined mineral salt media can be used. Defined media enable high cell density (HCD) cultivations, but the recombinant product yield is often low. Growth on glycerol (36) or complex medium (37–39) can be advantageous for solubility and folding of the recombinant product. The benefit of complex media may be caused by an increased expression of chaperones and foldases (40) or a change to osmotic pressure (41). T4-phage deoxycytidylate deaminase can be accumulated in a soluble active form to at least 20% of the cellular protein using a rich growth medium, whereas inclusion bodies are found in less rich media (37). Another cause can be the misincorporation of certain amino acids during growth in minimal media. Both the product quality and the folding can be affected, possibly leading to aggregation. Muramatsu et al. reported incorporation of β-methylnorleucine instead of isoleucine in recombinant hirudin produced by *E. coli* in a medium not enriched with amino acids (42). On the other hand the higher growth rates in complex media can impair secretion (43) and can lead to aggregation of the product (reviewed in ref. 44).

1.2. Effects of Alcohols

Alcohols are alkyl groups substituted with a hydroxyl group and influence the folding equilibrium of the protein. Monohydric alcohols mainly enhance the native structure and polyhydric alcohols (see Subheading 1.4) favour it (45). The addition of ethanol to the medium can be beneficial for the yield of functional product and was reported to increase the amount of soluble, native DsbA-proinsulin in *E. coli* about threefold (39).

Ethanol in the medium leads to changes in the lipid and fatty acids composition of the cell membrane (46). Moreover, peptidoglycan assembly is inhibited due to weakened hydrophobic interactions (47) being the main effect on protein structure in addition to hydrogen-bonding at low alcohol concentrations, thus causing a twist in the peptide backbone (48) and favouring the role of the peptide backbone over that of side chains (49). When using predominantly organic solutions the folding promoting influence can be dramatically increased by adding salts such as lithium chloride or bromide and sodium acetate or perchlorate, because the solubility of a charged species should be higher in the presence of another charged or polar species in solution (50).

In vivo, the addition of alcohols (e.g., ethanol) leads to an induction of heat shock proteins just like heat shock as such (51), which is used in practice to increase the solubility of recombinant proteins (52). As raising the cultivation temperature is not always feasible, ethanol can be added instead, e.g., in case of *E. coli* (53–55). The critical step for this is the interaction of the alcohols with the cell's lipophilic compounds (e.g., membranes, lipophilic core regions of proteins) (33). Ethanol-stressed cells also show changes in the

cellular fatty acid composition, which can be used in combination to allow other additives to enter the cells (56). However, ethanol at 10% can also impair translocation and lower the yield, if the target needs to be secreted. Unfolded precursor of periplasmic alkaline phosphatase remained in the cytoplasm of *E. coli*. Heat shock response was induced accordingly (57).

This could be beneficial to promote folding with the help of cytoplasmic chaperones, but then needs to be combined with a suitable translocation system, e.g., Tat-targeted.

1.3. Effects of Chemical Chaperones

The effects of compatible (protective only against extremes of temperature, dehydration, high salt; not affecting the biological activity of proteins) solutes have to be discussed with respect to their being osmoprotective towards the cells in the natural habitat. Such solutes (amino acids and their derivatives, sugars, polyols, quaternary amines and their sulphur analogues, sulphate esters, *N*-acetylated diamino acids, peptides) are uncharged at a neutral pH, highly water-soluble and can be accumulated to high amounts in the cell due to uptake or synthesis. They are compatible with macromolecular structure and function and do not interact with substrates and cofactors, and support protein folding due to the solvophobic effect (preferential exclusion from the protein surface (e.g., the exposed peptide backbone in the unfolded state) followed by preferential hydratation of the protein surface). Contrarily counteracting osmolytes reverse inactivation/destabilization by urea, inorganic ion inhibition and hydrostatic pressure and cause changes in protein function (58–64).

Chemical chaperones protect both prokaryotic and eukaryoytic cells (65) against environmental stresses and are therefore applicable to different recombinant hosts. For example proline is accumulated in bacteria and plants upon osmotic stress, whereas *Saccharomyces cerevisiae* induces glycerol or trehalose synthesis (66). Plants accumulate organic osmolytes generally by increased synthesis (67).

Polyols, amino acids and their derivatives affect protein stability and function, methylamines best support refolding and sugars stabilize proteins, but not necessarily the functional conformation (68, 69). Stabilizing osmolytes protect by reducing inactivation rate constants, by increasing transition free energy changes and/or by raising the midpoint of thermal denaturation as analyzed in vitro for trehalose, dimethyl sulfoxide (DMSO), glycine and proline, using guanidine-denatured yeast alcohol dehydrogenase as a model (70, 71). The eventual folding state of the protein depends on the sum of the osmolytes, the nature of other interacting macromolecules and any ligands involved (72, 73). Osmolytes also directly affect heat shock protein (e.g., Hsp90) conformation (74).

According to the compatibility hypothesis osmolytes are interchangeable but not in all respects due to some having unique roles (e.g., polyols are antioxidants) (75). It is worth mixing different

chemical chaperones, which can be successful at significantly lower concentrations than if used separately. Trying a certain additive in a new host can be considered, because the same one is often protecting a range of cells in nature (76). *Escherichia coli*, streptomyces, corynebacteria, mycobacteria and yeast synthesize trehalose in response to temperature, osmotic and oxidative stress (77).

1.4. Effects of Osmoprotectants and Osmotic Stress

Osmotic stress is exploited as a cultivation parameter in order to optimize the import of folding promoting agents into the cells (30). Exposing *E. coli* to 0.4 M NaCl causes within few minutes plasmolysis, respiration stop, cytoplasmic pH increase and up-regulation of GB and proline uptake to occur (78). Exogenously supplied GB is amassed up to 1 M then (29, 79). The cellular concentration of osmolytes depends on the organism, osmolyte availability in the medium and the type, severity and duration of the osmotic stress (28).

Due to water withdrawal, osmotic stress affects molecular crowding and leads to trapping (decreased diffusion) of proteins and thus accessibility of any interacting partners (chaperones, ligands, cofactors, redox agents), which highlights that the outcome of adding certain combinations/ concentrations of folding promoting agents at and for a certain time is diverse and cannot easily be predicted. 250 mM NaCl or 500 mM sorbitol decreased intracellular green fluorescent protein (GFP) diffusion in *E. coli* tenfold (80). Compatible osmolytes prevent phase separation (81).

Generally all effects must be considered systemically and not in isolation. The osmotic stress response also protects against oxidative and heat stresses. Adding 0.3 M NaCl is sufficient for maximal induction of thermotolerance even though the conditions are not hyperosmotic yet. Osmolytes (NaCl, KCl, proline or sucrose) affect ion channels and transporters in a way to achieve osmotic balance, and also contribute to cytoplasmic pH homeostasis in *E. coli*. K^+ can be substituted with NaCl, choline chloride, proline, or sucrose. The uptake of inorganic ions (mainly K^+) can compensate for osmolytes, but *E. coli* favours uptake of GB and proline over K^+ as well as synthesis of glutamate and trehalose (82–84). Van den Bogaart et al. showed that 10 mM K^+ ions, 1 mM GB and 1 mM proline restore protein mobility in osmotically stressed cells (80).

GB is the most efficient osmoprotectant at reversing protein aggregation, supporting correct folding, correcting folding of mutant proteins and countering the impact of denaturants in vivo and in vitro. It has been found to stabilize RNase-A over a broad range of pH values, but destabilizes α-lactalbumin and lysozyme at high and low pH (85, 86). So the effect cannot be predicted and the interrelation of all cultivation parameters such as pH is important. *Escherichia coli* can grow in media with pH 4.4–9.2 by means of considerable metabolic shifts depending on the actual media components. Periplasmic redox modulators are also pH-dependent. Thiol peroxidase is acid induced and the thiol:disulfide interchange protein DsbA is base induced (87).

The known synergy or interchangeability of agents can be beneficial in a mix-and-match of corresponding compensating parameters based on general host or process requirements. The combination of osmoprotectants and osmotic stress is often successfully applied in vivo. Using 0.5 M NaCl and 1 mM GB in a recombinant over-expression resulted in a higher soluble yield for six of nine targets (88).

Intracellular levels of DnaK are increased after addition of NaCl (32). Potassium glutamate and GB can activate ClpB (89). Other molecular chaperones are also activated in response to osmotic stress (28, 31). Pre-osmoadaptation of *E. coli* allows for osmolytes and heat shock proteins to prevent insoluble aggregates (28). Unsurprisingly salt-induced accumulation of osmolytes to improve yields of mis-folded recombinant proteins is also successfully used in combination with co-expression of DnaK-DnaJ-GrpE, ClpB, GroEL-GroES and IbpA/B. Yet similar high yields of target protein are achieved if only BA is added, which makes membranes more fluid and induces heat shock response and thus action of endogenous chaperones. Sufficient time has to be given to allow chaperone action. Fine-tuning the use of osmolyte addition, of endogenous or co-expressed chaperones, caters for different target proteins and process parameters (e.g., upscaling benzyl alcohol use to 500 mL can be a problem and requires cultivation optimization) (90). Even inhibitory osmotic stress conditions (NaCl and sorbitol) can be used to create an appropriate periplasmic microenvironment for the generation of high concentrations of correctly folded recombinant proteins (immunotoxins in this case), since adding GB also allows bacteria to grow (91). However, GB is only efficient up to conditions of 1 Osm via being excluded from the biopolymer surface. At higher osmolalities GB efficiency is decreased due to the low content of cytoplasmic water (92). Halophile organisms could be considered as hosts at higher osmolalities (93).

The additive effect allows sugar-methylamine mixtures to be very successful in enhancing protein stability and function (68). Using sorbitol/GB in rich medium did not only lead to absence of inclusion bodies, but the active yield of a target was increased up to 427-fold. In minimal medium the increase was 44-fold at least partially due to the lower osmotic pressure compared to rich medium (41). Again the effect of sorbitol/GB depends on amount and nature of protein production inducing agent, host strain and the target protein.

The yield of amorpha-4,11-diene synthase from *Artemisia annua* was sevenfold in *E. coli* since the critical aggregation level of the protein was not reached yet, whereas the yield of germacrene synthase from *Zingiber officinale* was only 2.5-fold, implying that lowering the rate of expression would be better to avoid critical aggregation level than preventing aggregation by using chemical chaperones (94).

The strategies can also be used with and optimized for other host systems. Salt stress prior to induction was successful to achieve higher scFv yields in *Pichia pastoris* (95), whereas salt supplementation impaired the secretion of an Fc fusion protein (96). In *Synechococcus spec*. salt stress inhibits the repair of photo-damaged photosystem II, but GB reversed this inhibition comparable to supporting correct folding of mis-folded proteins (97). Synthesis or uptake of ectoine and 5-hydroxyectoine is induced in *streptomyces* upon salt stress (98).

1.5. Effects of Amino Acids

The effect of adding amino acids has been studied both in vitro and in vivo and various modes of action were found. Twenty different amino acids were added to *E. coli* cultures and analyzed when producing periplasmic recombinant cytochrome b5.

Glycine and to a lesser extent histidine doubled the synthesis of the protein and its release into the medium (99). Moreover, glycine decreased the aggregation of ribonuclease-A (100). 1% Glycine supplemented during mid-exponential growth phase increases the yield of α-cyclodextrin glycosyltransferase in rich medium much more than if added at the beginning of the culture, where it impairs cell growth possibly due to increased membrane permeability (101). However, Ca^{2+} can remedy cell growth inhibition and increase the yield (102). The effect of glycine is not always beneficial and depends on the folding rate of the protein as well as on the concentration ratios.

The same goes for proline. Less than 0.5 M glycine and 1 M proline improved the in vitro refolding yields of creatine kinase, but higher concentrations decreased the recovery. In contrast to glycine, proline was able to inhibit aggregation of creatine kinase (103–105). Precipitation of hen egg white lysozyme is prevented in vitro by proline concentrations of more than 4.0 M due to an ordered supra molecular assembly (106). In vivo, proline reduces protein aggregation comparable to DnaK (107). Proline's role in osmoregulation needs to be considered when using it as a cultivation additive. At normal osmolarity proline is metabolized, but at high osmolarity 1 mM proline supports growth of *E. coli* and even better if combined with K^+ (108). So if proline is used it needs to be combined with salt. K^+ and PO_4^{3-} might be crucial for salt-induced proline uptake (109). Proline synthesis is rather independent of the availability of proline in enteric bacteria, not reaching high intracellular concentrations (0.05 mM). Elevated proline levels (90 mM) can be achieved by genetically modifying the proline biosynthesis pathway. Such strains could be used as hosts for recombinant protein production (107). The optimal salt and proline combination can be monitored in vivo in *E. coli* using an in-cell fluorescent-labelling assay (see Subheading 2.2).

Proline solubilizes the native state of an aggregation-prone model protein, but destabilizes its partially folded states and early

aggregates. 300 mM NaCl is best combined with 20 mM proline at the time of induction of recombinant production. Proline can support almost full solubility (25, 110). In *S. cerevisiae* the uptake of proline is induced by the addition of sorbitol or ethanol to the medium (66).

Arginine is preferentially excluded from protein–protein complexes, but not from the dissociated protein molecules. It acts as an aggregation suppressor but not as a protein-stabilizer (111–113). Due to this solubilising effect arginine increases the yield of native product (e.g., pro-insulin fused to the C-terminus of DsbA (39); native tissue-type plasminogen activator variant; scFV (11)). Furthermore it prevents death of animal cell cultures producing recombinant protein, or compensates the arginine consumption needed for recombinant protein production (114). Combining oxidized glutathione (GSSG) and L-arginine allows complete refolding of scFv in vitro (115). Arginine analogues such as L-argininamide enhance refolding yield due to their high positive net charge (116).

The addition of amino acids to cultures for recombinant protein production can both supplement limitations and directly affect the target proteins.

1.6. Effects of Sugars

Sugar supplementation establishes osmotic stress conditions, thereby indirectly affecting the recombinant protein. β-Lactamase aggregation is inhibited in *E. coli* by adding non-metabolizable sugars (e.g., sucrose) to the culture, which drive the in vivo folding pathway analogous to the well-analyzed effect of sugars in vitro (117, 118). Sucrose inhibits an increase in protein surface area exposure by affecting the hydrogen exchange rates of the amide protons with intermediate rates (119). In vivo, some stress-related sigma32- and sigmaE-dependent promoters are rapidly but transiently induced 15 min after adding sucrose to *E. coli* cultures following loss of turgor, which can be exploited during the production of recombinant proteins (32).

The disaccharide trehalose has a high affinity for water molecules and hence stabilizes partially unfolded protein molecules unspecifically (unrelated to any secondary structure motif) and inhibits protein aggregation (77, 120). In vivo, trehalose enhances heat shock protein mRNA levels in *S. cerevisiae* in response to stress (121) and stabilizes the native state of proteins during heat shock comparable to sucrose and maltose, and suppresses the aggregation of denatured proteins, maintaining them in a partially folded state from which they can be activated by molecular chaperones. Vice versa cold and heat shock activates trehalose synthesis in *E. coli* (122). Endogenously synthesized trehalose is not necessarily preventing protein aggregation in *E. coli* (25), but as an antioxidant it protects against free radicals produced by cells subjected to environmental stress (83). Trehalose reduces aggregate formation

also in mammalian cells (123). In vitro, trehalose protects membranes in addition to proteins both in solution and in a dry environment (13, 124). Thermo- as well as cryostability is increased by trehalose (58).

Xylose was proven to significantly increase the rate of folding and unfolding of staphylococcal nuclease. It stabilizes the folded state of proteins through surface tension effects (125).

1.7. Effects of Polyols

Glycerol and sorbitol are the polyols used mostly in recombinant protein production. The presence of glycerol in the culture medium improved the yield of wild type and mutant forms of the human recombinant phenylalanine hydroxylase and increased the specific activity of the purified enzymes in vitro. Also mannitol protected the enzyme from loss of activity (126, 127). The unfolded state of the target protein is made unfavourable relative to the folded state (65, 128). However, glycerol is not only an additive but also a substrate. Thus a potential decrease of the concentration has to be factored in. Glycerol is applicable to eukaryotic hosts (e.g., enhancing the yield and posttranslational stability of human P-glycoprotein expressed in yeast (129); restoring the mutant TP53-activated factor 1 to the wild type conformation in human cell lines (130)). Adding 1–5% (optimal at 5%; detrimental at 10%) glycerol to the culture medium reduced the aggregation of a recombinant human phenylalanine hydroxylase mutant with a hexa-histidine tag and allowed recovery of activity by stabilizing the nascent protein in *E. coli*. No change to the molecular chaperone expression was observed (131). There is also a pharmaceutical interest: antigen presentation is enhanced by glycerol (132). Even a mouse model for determining maximum serum glycerol concentrations has been established in order to test the in vivo efficacy of chemical chaperones for therapeutic use (133).

In vitro, polyols accelerate protein folding in general and disrupt hydrogen-bridged water bridges due to their ability to compete with water (134). Glycerol is also used as a cryo-protectant and for stabilizing proteins in aqueous solution (135).

Soluble expression of the *N* gene of the Chandipura virus could be achieved in *E. coli* by adding D-sorbitol. Without the chemical chaperone large insoluble aggregates are formed both in COS cells (derived from CV-1 cells, which are based on African green monkey kidney cells, and that were transformed by an origin of replication defective Simian Virus 40 mutant coding for wild type large T antigen) and in bacteria. The aggregation was significantly reduced in vitro as well (136).

Furthermore, polyols can support soluble expression of an scFv in *E. coli* when combined with over-expression of DsbABCD to catalyze disulphide bridge formation. Adding 0.5 M sorbitol to the culture medium allowed the best yield of correctly folded target. 0.5 M sucrose was much less effective, 4% NaCl was harmful to the host and 100 mM GB did not support the production (137).

1.8. Effects of Trimethylamine N-Oxide (TMAO)

TMAO can enhance protein folding and ligand binding, stabilizes both hydrophobic and hydrophilic proteins, and counteracts perturbations by urea and inorganic ions due to modifying water–amide interactions (58, 138). TMAO has an unfavourable interaction with the peptide backbone (139). The outcome of using TMAO depends on the protein and is not predictable. Trypsin activity is significantly enhanced in presence of TMAO, whereas chymotrypsin activity is not significantly changed (140). Activity of a recombinant human phenylalanine hydroxylase mutant with a hexa-histidine tag was recovered by adding 5 mM (detrimental at 50 mM) TMAO to the culture medium for *E. coli* (131).

Given the therapeutic interest in TMAO, tests with eukaryotic cell cultures were performed. Exposure of scrapie-infected mouse neuroblastoma cells to TMAO (or glycerol and DMSO) reduced the rate and extent of the formation of the pathogenic isoform of the cellular prion protein. The chemical chaperone interfered with the conversion of the α-helices into β-sheets in the newly synthesized proteins (141). TMAO restored up to 50% of the wild type activity in a deficient mitochondrial branched-chain α-ketoacid dehydrogenase complex, which causes maple syrup urine disease (142).

1.9. Effects of Reducing and Oxidizing Agents

Disulphide bridge isomerization is insufficient in the periplasm since DsbA is highly oxidizing and the isomerases cannot correct bridging sufficiently. Using a GSH/GSSG redox buffer in the medium allows modified formation and breakage of disulphide bridges in the periplasm of the bacterial host (143). Specific periplasmic redox states can be established by varying the redox potential of the medium. Thus comparable to already known in vitro effects (144), the addition of redox components allows reshuffling of incorrectly formed disulphide bridges in vivo in the periplasm. GSH can diffuse through outer membrane porins into the periplasm (145). The prokaryotic cytoplasm is not affected and the GSH and GSSG ratio of 50 to 200:1 maintained (146). In redox mutants this ratio is altered and hence (recombinant) protein folding by disulphide bridge formation is influenced.

Wunderlich and Glockshuber (27) reported a fivefold increase in correctly folded α-amylase/trypsin inhibitor from Ragi after adding GSH/GSSG to the medium. This successful outcome was even enhanced by co-expressing DsbA. This approach is confirmed successfully by others (147). GSH and also other agents (acetamide, ethylurea) added to the culture medium increased the yield of a native tissue-type plasminogen activator variant and an scFv up to 10- and 37-fold, respectively (11). The concentration of those supplements remained constant during cultivation until the ratio of GSH/GSSG changed after 20 h. Effects of the tested substances are comparable to those in vitro, and due to folding enhancing activities rather than secondary osmolyte effects. Nevertheless GSH/GSSG can also negatively affect the production of correctly folded target proteins (39).

An interesting alternative are diselenides such as selenocystamine, which can cross the outer bacterial membrane and can catalyze conversion of dithiols to disulphides in vivo, substituting for DsbA at micromolar concentrations, whilst the structurally analogous disulphides are only partially effective even at mM concentrations. Diselenides are more electrophilic than disulphides and are therefore also rapidly (re)oxidized to diselenides by atmospheric oxygen (148, 149).Successful use of reducing or oxidizing agents to the medium has also been reported for eukaryotic systems.

Oculopharyngeal muscular dystrophy is caused by the expansion of an alanine stretch in the intranuclear poly(A)-binding protein 2, leading to aggregations. The added chemical chaperone DMSO reduced aggregation of this protein and thus cell death (12).

Oxidative folding catalysts are even designed. One example was based on properties of protein disulphide isomerise (PDI) and found to improve the yield of a model protein up to three times compared to traditional in vitro protocols and to significantly improve in vivo folding (150).

1.10. Effects of Metal Ions

The addition of metal ions is beneficial due to their importance as cofactors for both recombinant and host proteins. The cofactors can specifically interact with the unfolded polypeptide. This is sometimes essential for the folding and thereby dramatically accelerates formation of the functional protein (151). Cofactors stabilize the native protein as well (152).

Cu^{2+} was reported to induce and enhance the formation of an α-helix conformation (153). Moreover copper catalyzes periplasmic disulphide bond formation under aerobic conditions, which needs to be considered when using it as a culture supplement (154).

The amount of an active, shortened form of the alkaline phosphatase was increased by addition of magnesium to the medium (155). Mg^{2+} ions increase the refolding yield of alkaline phosphatase and in combination with cyclodextrin to almost full, whereas other ions lead to low yields (156).

Zinc was not supportive in case of the alkaline phosphatase mutant but can inhibit proteases in the periplasm (157), which is also seen for calcium (38). The chaperone activity of TF is Zn^{2+}-dependent. Thus if TF is co-expressed, addition of Zn^{2+} could be beneficial (158). On the other hand eukaryotic PDI forms a dimer incorporating Zn^{2+}, which affects PDI activity in the ER and therefore needs to be considered when deciding upon additives (159).

The same goes for other proteins, where metal ions catalyze oligomerization as a means of post-translational regulation. This was discussed for the Ca^{2+}-dependent control of lipoprotein lipase dimerization (160). External free Ca^{2+} from the environment can

be concentrated to reach millimolar levels in the periplasm, but cytosolic free Ca^{2+} is controlled at micromolar range. So any supplementation can only support periplasmic recombinant yield (161). Iron is a major oxidative stress factor (87).

1.11. Effects of Artificial Chaperones

Positive effects of artificial chaperones on protein folding have been found both in vitro and in vivo. Cycloamylose supported the complete recovery of the enzymatic activity of chemically denatured citrate synthase, carbonic anhydrase B and a reduced form of lysozyme (162). The artificial chaperone system cetyltrimethylammonium bromide and β-cyclodextrin enhanced lysozyme renaturation in vitro (163). Cyclodextrins can simulate chaperone activity by controlling hydrophobic interactions (164) and are used for recombinant veterinary growth hormone production and storage (165). The combination of cetyltrimethylammonium bromide with dextrin 10 can also refold denatured protein (166).

The synthetic dithiol Vectrase P [(+/−)-trans-1,2-bis(2-mercaptoacetamido) cyclohexane] was successfully used as a folding promoting agent both in vivo and in vitro showing a partial isomerase-like function (167, 168). The agent can enter the eukaryotic ER and the prokaryotic periplasm. In vitro, Vectrase P catalyzed the activation of denatured bovine ribonuclease A. This dithiol added (0.4–1.5 mM) to the growth medium increased the heterologous secretion of *Schizosaccharomyces pombe* acid phosphatase from *S. cerevisiae* threefold (167), while 0.002–0.05 mM added at induction time increased the formation of native pro-insulin in the *E. coli* periplasm by about 60% (168).

New agents are continuously developed. In vitro, glucosylglycerol and glucosylglycerate stabilize enzymes partially better than trehalose during heat treatment and freeze drying (169).

1.12. Effects of Pharmacological Chaperones

Pharmacological chaperones are of recently emerging interest as folding promoting agents. They specifically support folding of a certain target protein by non-covalent interactions and are often ligands or antagonists of the target. Hence pharmacological chaperones are effective at lower concentrations than chemical chaperones and are not toxic (170, 171).

Given the potentially limited range of targets one such chaperone can be used for in recombinant production, might make them useful as bespoke solutions for profitable products and certain markets only. However, pharmaceutical chaperones could be more broadly used in combination with osmolytes (172).

RNAs are reported to enhance folding of their interacting proteins in vitro and thus might be used as folding promoting agents in a broader sense (173).

2. Experimental Approaches to Promoting the Folding of Recombinant Proteins

2.1. Overview

Many different protocols and variations thereof have been published over time. Any review of the literature shows that every target protein is different and one cannot predict the effect of a certain approach. More often than not one cannot even speculate. Therefore in the majority of cases it is necessary to adapt an entire protocol to the target protein. The procedure either has to be randomly tested or is chosen according to a certain rationale (e.g., if some protein parameters are known or there are in vitro findings). Methods for screening the additive's benefit have also been developed and established (see Subheading 2.2).

2.1.1. General Cultivation Strategies

The cultivation regime (environmental factors) ought to be planned depending on the most likely conducive medium, temperature, pH, osmolarity, timeline for additions and production. Those parameters depend not only on target protein, host and intrinsic factors such as codon usage, expression vector design and copy number but also on the context and purpose of the production and there are no empirical rules. Medium compounds, solubility, purity and monitoring of additives have to be considered with respect to the production scale (especially for HCD cultivation) and economic aspects (4, 44, 174–176).

There are four main approaches:

1. Both pre-cultures and main culture with additives at inoculation (41, 88, 94)
2. Pre-cultures without any additives, main culture with additives at inoculation (177)
3. Pre-cultures without any additives, main culture with additives after inoculation, induction of recombinant protein production after a certain adaptation time (90, 137, 178)
4. Pre-cultures without any additives, main culture with additives after inoculation, induction of recombinant protein production without any adaptation time (25, 131, 179).

A scheme of successful cultivation regimes cited in this chapter is depicted in Fig. 1. Overviews of used folding promoting agents (Table 1) and combinations thereof (Table 2) with corresponding concentrations might support the choice of strategy. The supplements can be used in rich media (88, 91), minimal media (41) or both (25, 79). The cultivation temperature has to be chosen according to the protein's aggregation tendency as well as economical growth and production rates. Lower temperatures often promote solubility due to slower protein production (8, 180). Thus the temperature ought to be as low as possible. In order to find a compromise in

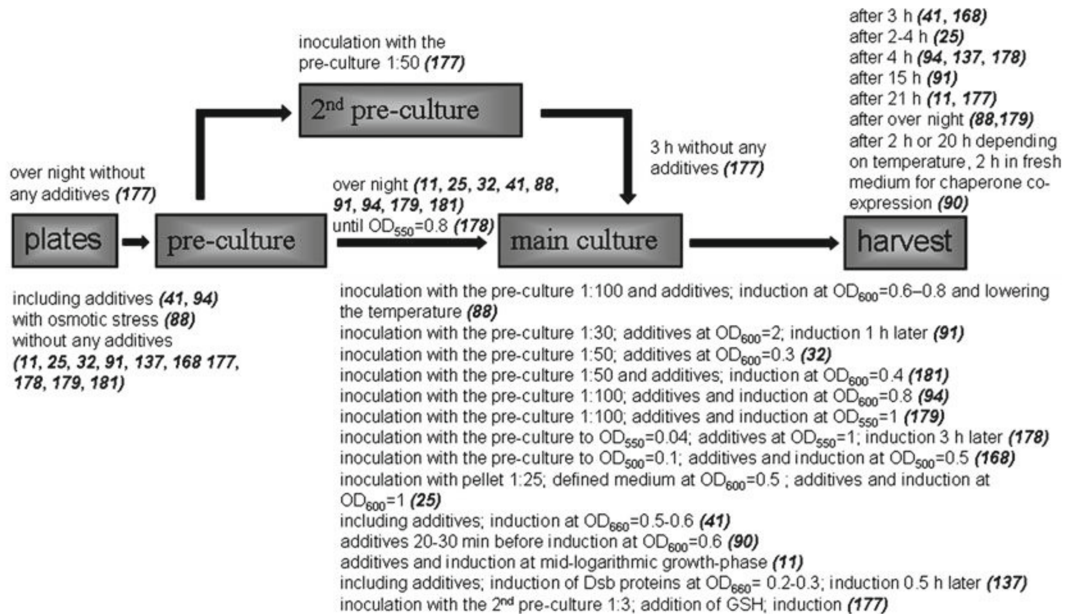


Fig. 1. Overview of successful cultivation regimes cited in this chapter. Adapted from ref. 243.

some cases the temperature is lowered at certain points such as the beginning of the second pre-culture (177) or at induction (41, 88). Nevertheless it might be more favourable to not change the temperature during the cultivation to avoid stressing the host cells additionally. Chosen temperatures in the references cited in this chapter are 20°C (88), 22°C (29, 168), 24°C (11, 177), 25°C (41), 26°C (179) and 30°C (32, 181). Thomas and Baneyx found 30°C to be more beneficial than 25, 20, 15 or 37°C (181). However, as in general lower cultivation temperatures favour solubility this is not a common standard.

2.1.2. Aspects to Be Considered

Certain agents affect growth and viability of the host cells, which has to be considered when planning cultivation schemes. More than 0.4 M of L-arginine (39) as well as 1% glycine (99, 101) lead to growth inhibition. Cytoplasmic water is a primary determinant of the growth rate of osmotically stressed cells. The cytoplasmic osmolality can be estimated by adding in vitro osmotic data of individual solutes and using a local bulk domain model for cytoplasmic water (92). 1 M NaCl inhibits proline uptake in *E. coli* (109). Depending on chain length and substitution, organic solvents can cause cell lysis (46, 182). Trehalose might be difficult to handle as a supplement, because it is mostly degraded in the periplasm as a carbon source. Only endogenously synthesized trehalose in response to a stressful environment has a recombinant protein protective function (30, 77). It is known that trehalose is degraded

Table 1
Overview of folding promoting agents with their corresponding concentrations for successful in vivo approaches as stated in the references cited in this chapter

Folding promoting agent	Concentration	Reference
Acetamide	1 M	(11)
L-Arginine	0.4 M 1.6–4.2 g/L	(11, 39) (114)
BA	10 mM	(90)
Dimethylsulfonioacetate	1 mM	(79)
Dimethylsulfoniopropionate	1 mM	(79)
Ectoine	0.5 mM 1 mM	(98) (79)
Ethanol	3%	(181)
Ethylurea	0.6 M	(11)
Formamide	1 M	(11)
Glycine betaine	1 mM 2.5 mM 5 mM 10 mM 20 mM	(79, 88, 94) (41, 94, 230) (90) (91) (25)
Glycerol	5% 10% 0.6–1.2 M	(131) (129) (130)
Glycine	1%	(99)
Glutathione	5 mM	(11, 177, 179)
Hydroxyectoine	0.5 mM 10 mM	(98) (91)
Methylformamide	0.6 M	(11)
Methylurea	0.6 M	(11)
NaCl	4% 0.3 M 0.5 M 0.6 M	(91) (25, 29, 110) (79, 88, 90, 98) (32)
$(NH_4)_2FeIISO_4$	0.1 mM	(131)
Polyethylene glycol	0.464 M	(32)
Potassium glutamate	5 mM 0.5 M	(90) (209)

(continued)

Table 1 (continued)

Folding promoting agent	Concentration	Reference
Proline	1 mM 20 mM	(79) (25, 110)
Raffinose	0.3–0.7 M	(179)
Sorbitol	6% 0.5 M 0.66 M 1 M	(230) (91, 137) (94) (41)
Sorbose	0.3–0.7 M	(179)
Sucrose	2% 0.9 M 0.464 M 0.5 M 0.3–0.7 M	(41) (79) (32) (137) (179)
TMAO	5 mM 1 M	(131) (142)
Vectrase P	0.4–1.5 mM 0.002–0.05 mM	(167) (168)
Zinc	0.5 mM	(91)

in the cells when recovering from stress to not inhibit reactivation of denatured proteins (13).

Thus additive concentrations need to be found in the range between minimum effectiveness and inhibiting the cell. Adaptation times are needed either for accumulation of the folding promoting agent in the host or for inducing a protein protective physiological response in the host. The heat shock response in *E. coli* peaks within 5–20 min, then declines and normal cell conditions are re-established within 2 h (183). Osmolytes can sustain heat shock-like conditions for longer. Sometimes adaptation is not necessary. In case of ethanol supplementation at less than 3% in the medium had no impact at all and more caused a growth inhibition. Ethanol added at induction time or 30 min before was less effective than added at inoculation of the main culture (181). 10% Ethanol impaired translocation (57).

The ratio of reducing and oxidizing thiols needs to be appropriate. Most efficient are 3–12 mM of one or various thiol reagents. Less than 0.1 mM is ineffective and more than 20 mM decrease the growth rate of the cells and the yield of the target protein. Both

Table 2
Overview of combinations of folding promoting agents with their corresponding concentrations for successful in vivo approaches as stated in the references cited in this chapter

Combination of folding promoting agents		Reference
1	300 mM NaCl 20 mM Proline 20 mM GB	(25)
2	0.5 M NaCl or 0.5 M KCl, 0.9 M sucrose 1 mM GB or ectoine, proline, dimethylsulfonioacetate, dimethylsulfoniopropionate	(79)
3	0.5 M NaCl 1 mM GB	(88)
4	0.5 M NaCl 5 mM GB 5 mM Potassium glutamate or 10 mM benzyl alcohol	(90)
5	4% NaCl 0.5 M Sorbitol 10 mM GB or hydroxyectoine	(91)
6	2.5 mM GB 660 mM Sorbitol	(94)
7	300 mM NaCl 20 mM Proline	(110)
8	0.1 mM $(NH_4)_2FeIISO_4$ 5% Glycerol or 5 mM TMAO	(131)
9	100 mM GB 0.5 M Sorbitol	(137)
10	Redox system of reducing and oxidizing thiols 5:1 to 10:1 GSH Cystein *N*-Acetylcystein Cysteamin β-Mercaptoethanol 0.3–0.7 M saccharides sorbose, sucrose, raffinose	(179)
11	6% Sorbitol 2.5 mM GB	(230)

the application of a single reducing thiol reagent and of combinations of reducing and oxidizing additives (at ratio 5:1 to 10:1) can improve the yield of correctly disulphide-bridged proteins (179). Selenocystine is very toxic to *E. coli* though (148, 149). Concentrations higher than 1 mM Vectrase P resulted in an adverse effect on the target protein (168).

2.1.3. Strategies Using Molecular Chaperones

Molecular chaperones assist in folding of newly synthesized proteins, prevent mis-folding and aggregation of proteins in the cells (reviewed in ref. (184, 185)). They are frequently applied in recombinant protein production (reviewed in ref. (186)). Here the expression of chaperones in the host cells as part of the physiological response to the addition of folding promoting agents is addressed briefly, as well as the use of folding promoting agents in combination with co-expressed chaperones. Proteins to be exported into the periplasm to then be subjected to molecular chaperones there, must not aggregate in the cytoplasm. As with all already discussed strategies, the outcome of including chaperones is unpredictable (187).

Chaperone over-production needs to be regulated and must not add to cellular stress (188). Cellular homeostasis requires a balance of intracellular concentrations of molecular chaperones. High concentrations of GrpE can inhibit DnaK-mediated refolding (189). Down-regulation or selective redirection of heat shock proteins could be better than over-expression or deletion, because manipulating the cellular properties based on the component interaction is better than manipulating the component (190). Unsuitable chaperone impairs not only host viability but also product stability (191), because *E. coli* quality control promotes solubility instead of conformational quality (192, 193). DnaK assists post- and co-translational folding (194), whereas GroEL/ES provides little or no co- or post-translational support (195). GroEL is the only chaperone in *E. coli* required under all growth conditions tested (196). However, only 85 host proteins actually require GroEL/ES for folding and use most of the GroEL capacity (197). Nevertheless, over-expression of GroEL/ ES enhances the thermotolerance of *E. coli* and thereby to conditions similar to heat shock (198).

Target proteins will be folded incorrectly for different reasons and thus require different chaperone activity for remedy (199). GroEL (being promoting) and DnaK (being preventing) are antagonistic in controlling inclusion body formation (200). DnaK mediates the solubilization or disaggregation of proteins (186, 201). Yet, over-expression of *dnaK* can lead to stabilization of inclusion bodies (202, 203). GroEL supports disaggregation, and IbpA and IbpB prevent aggregation (204, 205). Unsurprisingly increasing the yield of soluble protein via over-expressing chaperones does not always lead to correctly folded target (192, 206).

Given that the necessary molecular tools are widely (commercially and within the research community) available nowadays, combinatorial approaches for comparative screening the chaperone impact on the target protein can be used initially. The results can also be fed into models for experimental planning. De Marco provided a strategy where combinations of molecular chaperones are co-expressed with the target. The ratio of the individual chaperones reflects their ratio in *E. coli* cells under stress conditions. Then the system, which proved (most) successful, is scaled up.

As part of the cultivation scheme chloramphenicol is added to inhibit protein synthesis after sufficient target production to give time for disaggregation of precipitates and for folding (207).

However, combining chemical and molecular chaperones adds yet another level of complexity. Therefore the impact on the target protein cannot be deduced based on just analyzing the effect of molecular chaperones. Diamant et al. reported low physiological concentrations of proline, glycerol and especially GB to activate GroEL and DnaK. So the increased yield of native target protein might be due to local protein refolding within the chaperone and stabilization of the product (28). Osmolytes also impact on the chaperone–substrate interaction. Release and reactivation of proteins from the GroEL–target complex in the cell follows ATP binding, but the addition of glycerol (and other polyols such as sucrose, 1,2-propanediol or 1,3-propanediol, but not GB, sarcosine or high salt) had the same result even in the absence of the nucleotide. As dextran or Ficoll failed to reactivate a GroEL-bound model protein, the effect cannot be attributed to viscosity or molecular crowding. Thus glycerol may alter the chaperonin structure similar to ATP (134). Too high osmolyte concentrations on the other hand can be detrimental. Especially trehalose inhibits DnaK-dependent chaperone networks since high viscosity affects dynamic chaperone–substrate interactions and stabilizes protein aggregates. Different osmolytes affect chaperones differently, but the protection mechanism against aggregation is similar for all analyzed osmolytes. The nature of the osmolyte determines partitioning and commitment of the unfolded species to the proper refolding or improper mis-folding pathways. Furthermore the osmolyte's viscosity and protective effect (e.g., against heat, but not urea) are not always correlated (28). In order to maximize potential benefits some authors did additionally co-express certain molecular chaperones such as DnaK and J (181). DnaJ plus additives caused a dramatic increase in yield (11). The outcome might depend on the nature of the target protein, given that osmoprotectants accumulated in the cell lead to a low production of chaperones anyway. Thus the protection of destabilized proteins is mainly ensured by osmoprotectants (79, 193).

2.1.4. Recommended Basic Approach

As stated in Subheading 2.1.1 the cultivation regime depends on and is informed by the requirements of the protein. Every parameter has to be optimized. So far this has been mostly a time-consuming linear approach in shaking flasks. Additionally previously optimized parameters can require re-optimization later due to interrelations. This is depicted as a recommended basic approach in Fig. 2 with respect to medium, temperature, inducer, additives and timing. Furthermore host strain, target proteolysis, codon usage and involvement of molecular chaperones might have to be optimized in the same manner. It is impossible to prioritize any part of

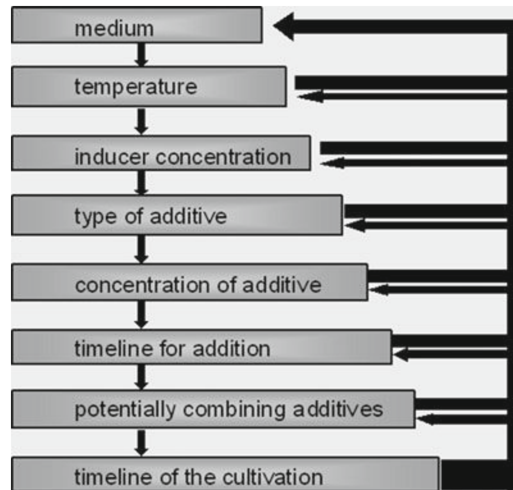


Fig. 2. Recommended basic approach to a cultivation with additives: The phases of optimization are interrelated and can require re-optimization. Adapted from ref. 243.

the optimization process, thus making it a sub-optimal successive approximation. An alternative is the use of the powerful statistic "design of experiment" approach. This methodology is implemented in modelling toolboxes. Even though it is a step forward, the next demanding task is the efficient and reliable analysis of the additive's effect, which can be studied by means of modern systems approaches and even in combination with folding/activity screening assays. Resulting data can be used directly for experimental planning or fed into knowledge-based modelling as an alternative to design of experiment.

2.2. Screening the Impact of Additives on the Folding of Recombinant Proteins

Finding an effective cultivation supplement, the appropriate concentration and adaptation conditions is in most cases a time-consuming process. That is why faster approaches are invaluable. Both in vivo and in vitro methods for screening the influences of osmolytes on the folding of the target protein have been developed. One strategy was published by Voziyan and co-workers. There the effects of a broad array of folding promoting agents (e.g., osmolytes, detergents, gradients of ionic strength and pH) can be conveniently analyzed in vitro by using a stable complex of the chaperone GroEL and the mis-folded target protein. As the target is stabilized, the impact of osmolytes on promoting the folding reactions towards the native conformation can be assessed. In some cases the initial aggregation reaction can be accelerated depending on the protein and the osmolyte and therefore leads to a negative readout. Using this screening technology, superior folding conditions for a truncation mutant of bacterial glutamine synthetase were rapidly identified: glycerol and sucrose were found to successfully promote folding. L-Proline or sarcosine supported folding of phosphoinositol

transfer protein and sucrose was beneficial in case of mitochondrial malate dehydrogenase. (134, 208–210) The GroEL-folding intermediates complexes can be immobilized on beads in multiwell plates, which are then used in HT screening of large chemical combinatorial libraries for specific ligands that prevent mis-folded intermediates of the target. Hits are read as an increase in the amount of soluble protein in solution. Identifying drugs against protein folding diseases is a recent application in addition to biotechnology contexts (211).

Other in vitro screening methods also apply chaperones. Stressed (mis-folded) protein is detected quickly by monitoring its binding to an immobilized chaperone by means of surface plasmon resonance. This sensor is actually used as an approach to detect stress, but one could equally imagine the opposite, i.e., adapting the screening to sense a reduction in stress upon exposure to folding promoting agents (212). An alternative could be a multiwell plate assay, where denatured firefly luciferase is treated with a mixture of DnaK and small molecules. The measured luminescence equates to refolded luciferase and thus the benefit of using the additive. However, this system is more applicable to finding new potential folding promoting agents to then try with the target of interest, since one cannot assume that what works for luciferase also works for the target (213).

Cell-free systems could be used for pre-screening or to pilot the concentration range of agents. Both translation initiation and elongation of globin mRNA were inhibited in a cell-free protein synthesis by unusually high concentrations of inorganic ions, but not by the compatible osmolytes GB and *myo*-inositol (214). Final concentrations inside the cells will obviously be different then depending on transport, but invaluable calibration curves of how the concentration in the cell-free system relates to the concentration when used as a cultivation additive with the same effect on the target protein, could be produced for further reference.

The data could be linked to known in vitro interaction of chemical chaperones with model proteins, which have been measured using various principles. Rösgen and colleagues gained protein solvation information from experimental thermodynamic data using Kirkwood–Buff theory. Kirkwood–Buff theory in combination with the transfer model provides information on the folding equilibrium of proteins in the presence of osmolytes. Hydration and osmolation of a protein was thermodynamically assessed based on measurements with model compounds (215, 216). Additionally Tanford's transfer model was used to predict equilibrium properties of proteins at certain concentration of osmolytes based on experiments (217). Experimentally derived static structure factors without and with co-solvents and salts can be analyzed based on the Derjaguin–Landau–Verwey–Overbeek potential (218). *M*-values for osmolytes in the presence and absence of a second

osmolyte are used to determine the effects of mixtures of organic osmolytes on protein stability (219). All measurements and findings could feed into knowledge-based models.

There are also in vitro screening methods based on cell component-free systems. The conformational change of a mis-folded polypeptide upon exposure to a test agent is measured by detecting the change in fluorescence intensity of the polypeptide in combination with a fluorescent compound. If the signal compares to the one of the correctly folded protein, the tested agent had a beneficial effect. Analyzing changes in the nuclear magnetic resonance spectrum or the circular dichroism as well as tracing the specific binding of an antibody to the peptide are other means of assessing the test agent (220).

If the target protein is a ligand its correct folding can also be efficiently screened by detecting successful binding to its receptor. Such a method was developed for recombinant human bone morphogenetic protein 2 to monitor the outcome of recombinant production strategies (193). The target is bound to a nitrocellulose membrane and then incubated with biotinylated receptor (here: ecto-domain of Alk3). Binding is detected colourimetrically. The signal is specific since two of the receptor molecules can bind only to native bone morphogenetic protein 2 dimer by contacting both monomers (221, 222). In vitro screening is only of limited use though, because the impact of additives does not merely depend on the folding pathway. A transcriptomic analysis (223) of the host (here: *E. coli*) response to 13 different cultivation strategies using either a single folding promoting agent or combinations was performed to complement the results of the in vitro assay of the recombinant product conformation (193). The analysis showed that producing the target fused to maltose binding protein without any additives led to an increased protein-protective stress response (increased expression of periplasmic factors and cytoplasmic *groEL, ES*), which was not the case if sorbitol and GB were added to the cultivation. This is in accordance with findings of Bourot et al. (79).

Meanwhile also other systems approaches were used to collect data about the changes in the host during recombinant production in an attempt to address (i.e., to correct detrimental states) them subsequently or to exploit them (i.e., the beneficial states) when developing cultivation strategies (224). The metabolic response to salt stress in *Streptomyces sp.* involves an up-regulation of amino acids and di- and tripeptides containing proline or glycine, which can inform supplementation (225). Proteomic information from stress response identified over-expressed molecular chaperones that can be used as bespoke fusion partners or for co-expression (226).

In case of HCD cultivation recombinant production of antibodies in *E. coli* the proteome showed an increase in the stress protein phage shock protein A, which can be co-produced to improve soluble periplasmic yield (227). Transcriptomic analysis

revealed that in dual stress conditions (heat shock and recombinant protein production) the response of *E. coli* cells is not a simple additive of the ones to individual stresses. Amino acid-tRNA genes were elevated for the dual stress, which again might be a folding supporting "side effect" also seen when using a cultivation parameter that contributes to triggering heat shock (228). Osmotically stressed *E. coli* were analyzed using DNA arrays. Stationary phase-specific genes and genes responsible for the de novo synthesis of GB and proline as well as for corresponding uptake (*proU* operon) were increased as was the transcriptional level of universal stress protein A (78).

Despite the groundbreaking benefits of combining in vitro screening for alternative/beneficial concentrations of agents or combinations, which could then be tried in vivo, direct in vivo screening is a much more straightforward approach. Several technologies were developed and established recently. Kraft et al. described in vivo monitoring of the formation of mis-folded proteins and their aggregates in the *E. coli* cytoplasm. Host cells are transformed with a plasmid encoding a luciferase reporter gene under control of a tandem promoter based on sigma32-dependent promoters *ibpAB* and *fxsA*. Not much light is produced in the absence of mis-folded proteins, whereas the stress response to mis-folded protein results in a significant signal, which can be detected both online (flow through luminometer) or offline (plate reader). The system is suitable for all scales from multiwell plate to bioreactor/fermenter (229). A periplasmic variant is available based on the sigma 24-regulated *degP* promoter. The on-line luciferase signal is inversely proportional to functional recombinant protein and was used to demonstrate the folding promoting activity of sorbitol and GB to commence very rapidly, while Skp over-expression prevents stress for some time with a later onset (230).

An in-cell fluorescent labelling method for monitoring protein stability and aggregation in vivo was developed by Ignatova et al. A tetra-Cys motif is inserted in a protein of interest or a model protein to allow binding of a fluorescein-based dye. Unfolding of the model protein leads to enhanced fluorescence. Hence direct and real time observation of protein production and aggregation in bacterial cells is possible, allowing for monitoring the effect of folding promoting agents. However, inserting the motif causes disruption of the target (potentially leading to inconclusive findings) and the method requires lysozyme treatment, which is not feasible for production conditions. Yet it is a very useful tool for screening and experimental planning (25, 231, 232). The same goes for a reporter-based monitoring system for protein folding in the *E. coli* periplasm described by Mansell et al. An N-terminal signal peptide is fused to a target protein, which in turn is N-terminally fused to β-lactamase. This model can be applied to general screening for periplasmic factors that support protein folding rather than to find

a solution for a certain target protein due to the unknown impact of the fusion (233).

All discussed in vivo and in vitro screening approaches in combination with analyzing the host responses are invaluable tools for using folding promoting agents in recombinant protein production.

3. Concluding remarks

It is well worth to optimize the production conditions for recombinant proteins and depending on the eventual use of the target, more or less effort, time and resources can be afforded. Initially substances and protocols proven instrumental in both in vitro refolding and in vivo approaches ought to be considered within the constraints of the production scale (laboratory, pilot, industry). This chapter was intended to guide through the plethora of recent applications of folding promoting agents in recombinant protein production in support of finding or choosing a suitable one since despite the many successful strategies some do not show any effect or are even disadvantageous.

Traditional shaking flask trial and error experiments only allow a linear optimization without considering interrelations directly. A multidimensional DOE procedure would be more efficient. Combining the cultivation strategies with systems approaches for analyzing the host response and with protein folding assays to screen for the impact on the target, and making the data available for knowledge-based modelling is trend setting and crucial, and will be characteristic of future strategic concepts. Any good modelling requires the combination of large amounts of data covering the target protein, the host, the folding promoting agents and any interactions. Moreover, suitable search criteria need to be used when applying the models to predict a feasible cultivation process.

It can be anticipated that eventually there will be repositories and searchable databases in support of such concerted schemes. Access might be free or limited to certain clients. Comparable developments were seen for in vitro refolding protocols. The REFOLD database (http://refold.med.monash.edu.au) allows browsing of hundreds of methods of refolding and purification of recombinant proteins and even includes molecular chaperone-assisted strategies (234–236).

Additionally HT technology including automation and miniaturization will contribute to more economic strategies. Recombinant protein expression and purification have already been subjected to HT approaches (237, 238). Identified successful folding promoting agents can be subsequently used in in vitro or in vivo screenings and in vivo optimizations. As an intermediate stage one might also use supplemented in vitro translation methods (239–242), which are even suitable for large-scale protein analysis (240, 241).

References

1. Chambers SP, Austen DA, Fulghum JR, et al., (2004) High throughput screening for soluble recombinant expressed kinases in *Escherichia coli* and insect cells. *Protein Expr. Purif.* **36**, 40–47.
2. Marblestone JG, Edavettal SC, Lim Y, et al., (2006) Comparison of SUMO fusion technology with traditional gene fusion systems: enhanced expression and solubility with SUMO. *Protein Sci.* **15**,182–189.
3. Chayen NE (2004) Turning protein crystallisation from an art into a science. *Curr.Opin. Struct. Biol.* **14**, 577–583.
4. Structural Genomics Consortium, China Structural Genomics Consortium, Northeast Structural Genomics Consortium *et al.* (2008) Protein production and purification. *Nat Methods* **5**, 135–146.
5. Baneyx F (1999) Recombinant protein expression in *Escherichia coli*. *Curr.Opin. Biotechnol.* **10**, 411–421.
6. Georgiou G, Valax P (1996) Expression of correctly folded proteins in *Escherichia coli*. *Curr. Opin. Biotechnol.* 7, 190–197.
7. de Marco A (2009) Strategies for successful recombinant expression of disulfide bond-dependent proteins in *Escherichia coli*. *Microb. Cell. Fact.* **8**, 26.
8. Esposito D, Chatterjee DK (2006) Enhancement of soluble protein expression through the use of fusion tags. *Curr. Opin. Biotechnol.* 7, 353–358.
9. Baneyx F, Mujacic M (2004) Recombinant protein folding and misfolding in *Escherichia coli*. *Nat. Biotechnol.* **22**, 1399–1408.
10. Waugh DS (2005) Making the most of affinity tags. *Trends Biotechnol.* **23**, 316–320.
11. Schaeffner J, Winter J, Rudolph R, et al., (2001) Cosecretion of chaperones and low-molecular-size medium additives increases the yield of recombinant disulfide-bridged proteins. *Appl. Environ. Microbiol.* **67**, 3994–4000.
12. Bao YP, Cook LJ, O'Donovan D, et al., (2002) Mammalian, yeast, bacterial, and chemical chaperones reduce aggregate formation and death in a cell model of oculopharyngeal muscular dystrophy. *J. Biol. Chem.* **277**, 12263–12272.
13. Singer MA, Lindquist S (1998) Multiple effects of trehalose on protein folding *in vitro* and *in vivo*. *Mol. Cell.* **1**, 639–648.
14. Rosenbusch JP (1990) Structural and functional properties of porin channels in *Escherichia coli* outer membranes. *Experientia* **46**, 167–173.
15. Ostermeier M, Georgiou G (1994) The folding of bovine pancreatic trypsin inhibitor in the *Escherichia coli* periplasm. *J. Biol. Chem.* **269**, 21072–21077.
16. Brass JM, Higgins CF, Foley M, et al., (1986) Lateral diffusion of proteins in the periplasm of *Escherichia coli*. *J. Bacteriol.* **165**, 787–795.
17. Van Wielink JE, Duine JA (1990) How big is the periplasmic space? *Trends Biochem. Sci.* **15**, 136–137.
18. Mika JT, van den Bogaart G, Veenhoff L, et al., (2010) Molecular sieving properties of the cytoplasm of *Escherichia coli* and consequences of osmotic stress. *Mol. Microbiol.* **77**, 200–207.
19. Le Gall S, Neuhof A, Rapoport T (2004) The endoplasmic reticulum membrane is permeable to small molecules. *Mol. Biol. Cell.* **15**, 447–455.
20. Hincha DK, Hagemann M (2004) Stabilization of model membranes during drying by compatible solutes involved in the stress tolerance of plants and microorganisms. *Biochem J.* **383**, 277–283.
21. Hincha DK (2006) High concentrations of the compatible solute glycinebetaine destabilize model membranes under stress conditions. *Cryobiology* **53**, 58–68.
22. Harishchandra RK, Wulff S, Lentzen G, et al., (2010) The effect of compatible solute ectoines on the structural organization of lipid monolayer and bilayer membranes. *Biophys. Chem.* **150**, 37–46.
23. Ott V, Koch J, Späte K, et al., (2008) Regulatory properties and interaction of the C- and N-terminal domains of BetP, an osmoregulated betaine transporter from *Corynebacterium glutamicum*. *Biochemistry* **47**, 12208–12218.
24. Morbach S, Krämer RJ (2005) Structure and function of the betaine uptake system BetP of *Corynebacterium glutamicum*: strategies to sense osmotic and chill stress. *Mol. Microbiol. Biotechnol.* **10**, 143–153.
25. Ignatova Z, Gierasch LM (2007) Effects of osmolytes on protein folding and aggregation in cells. *Methods Enzymol.* **428**, 355–372.
26. Holtmann G, Bremer EJ (2004) Thermoprotection of *Bacillus subtilis* by exogenously provided glycine betaine and structurally related compatible solutes: involvement of Opu transporters. *Bacteriol.* **186**, 1683–1693.
27. Wunderlich M, Glockshuber R (1993) *In vivo* control of redox potential during protein folding catalyzed by bacterial protein disulfide-

isomerase (DsbA). *J. Biol. Chem.* **268**, 24547–24550.

28. Diamant S, Eliahu N, Rosenthal D, et al., (2001) Chemical chaperones regulate molecular chaperones *in vitro* and in cells under combined salt and heat stresses. *J. Biol. Chem.* **276**, 39586–39591.
29. Caldas T, Demont-Caulet N, Ghazi A, et al., (1999) Thermoprotection by glycine betaine and choline. *Microbiology* **145**, 2543–2548.
30. Kempf B, Bremer E (1998) Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. *Arch. Microbiol.* **170**, 319–330.
31. Gill RT, DeLisa MP, Valdes JJ, et al., (2001) Genomic analysis of high-cell-density recombinant *Escherichia coli* fermentation and "cell conditioning" for improved recombinant protein yield. *Biotechnol. Bioeng.* **72**, 85–95.
32. Bianchi AA, Baneyx F (1999) Hyperosmotic shock induces the sigma32 and sigmaE stress regulons of *Escherichia coli. Mol. Microbiol.* **34**, 1029–1038.
33. Neuhaus-Steinmetz U, Rensing L (1997) Heat shock protein induction by certain chemical stressors is correlated with their cytotoxicity, lipophilicity and protein-denaturing capacity. *Toxicology* **123**, 185–195.
34. Miller EN, Ingram LO (2007) Combined effect of betaine and trehalose on osmotic tolerance of *Escherichia coli* in mineral salts medium. *Biotechnol. Lett.* **29**, 213–217.
35. Jaenicke R (1998) Protein self-organization *in vitro* and *in vivo*: partitioning between physical biochemistry and cell biology. *Biol. Chem.* **379**, 237–243.
36. Kopetzki E, Schumacher G, Buckel P (1989) Control of formation of active soluble or inactive insoluble baker's yeast alpha-glucosidase PI in *Escherichia coli* by induction and growth conditions. *Mol. Gen. Genet.* **216**, 149–155.
37. Moore JT, Uppal A, Maley F, et al., (1993) Overcoming inclusion body formation in a high-level expression system. *Protein Expr. Purif.* **4**, 160–163.
38. Kurokawa Y, Yanagi H, Yura T (2000) Overexpression of protein disulfide isomerase DsbC stabilizes multiple-disulfide-bonded recombinant protein produced and transported to the periplasm in *Escherichia coli. Appl. Environ. Microbiol.* **66**, 3960.
39. Winter J, Neubauer P, Glockshuber R, et al., (2000) Increased production of human proinsulin in the periplasmic space of *Escherichia coli* by fusion to DsbA *J. Biotechnol.***84**, 175–185.
40. Wei Y, Lee JM, Richmond C, et al., (2001) High-density microarray-mediated gene expression profiling of *Escherichia coli. J. Bacteriol.***183**, 545–556.
41. Blackwell JR, Horgan R (1991) A novel strategy for production of a highly expressed recombinant protein in an active form. *FEBS Lett.* **295**, 10–12.
42. Muramatsu R, Negishi T, Mimoto T, et al., (2002) Existence of beta-methylnorleucine in recombinant hirudin produced by *Escherichia coli. J. Biotechnol.* **93**, 131–142.
43. Mergulhão FJ, Monteiro GA, Larsson G, et al., (2003) Medium and copy number effects on the secretion of human proinsulin in *Escherichia coli* using the universal stress promoters *uspA* and *uspB. Appl. Microbiol. Biotechnol.* **61**, 495–501.
44. Fahnert B, Lilie H, Neubauer P (2004) Inclusion bodies: formation and utilisation. *Adv. Biochem. Eng. Biotechnol.* **89**, 93–142.
45. Usha R, Maheshwari R, Dhathathreyan A, et al., (2006) Structural influence of mono and polyhydric alcohols on the stabilization of collagen. *Colloids Surf B. Biointerfaces* **48**, 101–105.
46. Ingram LO, Dickens BF, Buttke TM (1980) Reversible effects of ethanol on *Escherichia coli. Adv. Exp. Med. Biol.* **126**, 299–337.
47. Ingram LO (1981) Mechanism of lysis of *Escherichia coli* by ethanol and other chaotropic agents. *J. Bacteriol.* **146**, 331–336.
48. Dwyer D S (1999) Molecular simulation of the effects of alcohols on peptide structure. *Biopolymers* **49**, 635–645.
49. Wang A, Bolen DW (1997) A naturally occurring protective system in urea-rich cells: mechanism of osmolyte protection of proteins against urea denaturation. *Biochemistry* **36**, 9101–9108.
50. Rariy RV, Klibanov AM (1999) Protein refolding in predominantly organic media markedly enhanced by common salts. *Biotechnol. Bioeng.* **62**, 704–710.
51. Chiang ML, Ho WL, Yu RC, et al., (2008) Protein expression in *Vibrio parahaemolyticus* 690 subjected to sublethal heat and ethanol shock treatments. *J. Food Prot.* **71**, 2289–2294.
52. Chen J, Acton TB, Basu SK, et al., (2002) Enhancement of the solubility of proteins overexpressed in *Escherichia coli* by heat shock. *J Mol. Microbiol. Biotechnol.* **4**, 519–524.
53. Thomas J, Baneyx F (1995) Protein misfolding and inclusion body formation in recombinant *Escherichia coli* cells overexpressing heat shock proteins. *J. Biol. Chem.* **271**, 11141–11147.
54. Kusano K, Waterman MR, Sakaguchi M, et al., (1999) Protein synthesis inhibitors and ethanol

selectively enhance heterologous expression of P450s and related proteins in *Escherichia coli*. *Arch. Biochem. Biophys.* **367**, 129–136.

55. Barosso JF, Elholm M, Flatmark T (2003) Tight binding of deoxyribonucleotide triphosphates to human thymidine kinase 2 expressed in *Escherichia coli*. *Biochemistry* **42**, 15158–15169.
56. Chiou RY, Phillips RD, Zhao P, et al., (2004) Ethanol-mediated variations in cellular fatty acid composition and protein profiles of two genotypically different strains of *Escherichia coli* O157:H7. *Appl. Environ. Microbiol.* **70**, 2204–2210.
57. Chaudhuri S, Jana B, Basu T (2006) Why does ethanol induce cellular heat-shock response? *Cell Biol. Toxicol.* **22**, 29–37.
58. Yancey PH (2005) Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. *J. Exp. Biol.* **208**, 2819–2830.
59. Schiefner A, Breed J, Bösser L et al., (2004) Cation-pi interactions as determinants for binding of the compatible solutes glycine betaine and proline betaine by the periplasmic ligand-binding protein ProX from *Escherichia coli*. *J. Biol. Chem.* **279**, 5588–5596.
60. Rösgen J (2007) Molecular basis of osmolyte effects on protein and metabolites. *Methods Enzymol.* **428**, 459–486.
61. Empadinhas N, da Costa MS (2008) Osmoadaptation mechanisms in prokaryotes: distribution of compatible solutes. *Int. Microbiol.* **11**, 151–161.
62. Arakawa T, Tsumoto K, Kita Y, et al., (2007) Biotechnology applications of amino acids in protein purification and formulations. *Amino Acids* **33**, 587–605.
63. Yancey PH (2004) Compatible and counteracting solutes: protecting cells from the Dead Sea to the deep sea. *Sci. Prog.* **87**, 1–24.
64. Yancey PH (2003) Proteins and counteracting osmolytes. *Biologist* **50**, 126–131.
65. Burg MG, Ferraris JD (2008) Intracellular organic osmolytes: function and regulation. *J. Biol. Chem.* **283**, 7309–7313.
66. Kaino T, Takagi H (2008) Gene expression profiles and intracellular contents of stress protectants in *Saccharomyces cerevisiae* under ethanol and sorbitol stresses. *Appl. Microbiol. Biotechnol.* **79**, 273–283.
67. Parida AK, Das AB (2005) Salt tolerance and salinity effects on plants: A review. *Ecotoxicol. Environ. Saf.* **60**, 324–349.
68. Jamal S, Poddar NK, Singh LR, et al., (2009) Relationship between functional activity and protein stability in the presence of all classes of stabilizing osmolytes. *FEBS J.* **276**, 6024–6032.
69. Dong XY, Huang Y, Sun Y (2004) Refolding kinetics of denatured-reduced lysozyme in the presence of folding aids. *J. Biotechnol.* **114**, 135–142.
70. Han HY, Yao ZG, Gong CL, et al., (2010) The Protective Effects of Osmolytes on Yeast Alcohol Dehydrogenase Conformational Stability and Aggregation. *Protein Pept. Lett.* **17**, 1058–1066.
71. Yancey PH, Clark ME, Hand SC, et al., (1982) Living with water stress: evolution of osmolyte system. *Science* **217**, 1212–1222.
72. Kumar R (2009) Role of naturally occurring osmolytes in protein folding and stability. *Arch. Biochem. Biophys* **491**, 1–6.
73. Olsen SN, Ramløv H, Westh P (2007) Effects of osmolytes on hexokinase kinetics combined with macromolecular crowding: test of the osmolyte compatibility hypothesis towards crowded systems.*Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **148**, 339–345.
74. Street TO, Krukenberg KA, Rosgen J, et al., (2010) Osmolyte-induced conformational changes in the Hsp90 molecular chaperone. *Protein Sci.* **19**, 57–65.
75. Bolen DW (2004) Macromolecular Crystallization Effects of naturally occurring osmolytes on protein stability and solubility: issues important in protein crystallization. *Methods* **34**, 312–322.
76. Singh LR, Chen X, Kozich V, et al., (2007) Chemical chaperone rescue of mutant human cystathionine beta-synthase. *Mol. Genet. Metab.* **91**, 335–342.
77. Jain NK, Roy I (2009) Effect of trehalose on protein structure. *Protein Sci* **18**, 24–36.
78. Weber A, Jung KJ (2002) Profiling early osmostress-dependent gene expression in *Escherichia coli* using DNA macroarrays. *Bacteriol* **184**, 5502–5507.
79. Bourot S, Sire O, Trautwetter A, et al., (2000) Glycine betaine-assisted protein folding in a lysA mutant of *Escherichia coli*. *J. Biol. Chem.* **275**, 1050–1056.
80. van den Bogaart G, Hermans N, Krasnikov V, et al., (2007) Protein mobility and diffusive barriers in *Escherichia coli*: consequences of osmotic stress. *Mol. Microbiol.* **64**, 858–671.
81. Braun FN (2006) Salting-in the microbial cytoplasm. *FEBS Lett.* **580**, 720–722.
82. Kitko RD, Wilks JC, Garduque GM, et al., (2010) Osmolytes contribute to pH homeostasis of *Escherichia coli*. *PLoS One* **5**, e10078.
83. Gunasekera TS, Csonka LN, Paliy O (2008) Genome-wide transcriptional responses of

Escherichia coli K-12 to continuous osmotic and heat stresses. *J Bacteriol.* **190**, 3712–3720.

84. Shabala L (2009) Organic vs inorganic: What makes the major contribution to osmotic adjustment in bacteria? *Commun. Integr. Biol.* **2**, 74–75.
85. Singh LR, Dar TA, Rahman S, et al., (2009) Glycine betaine may have opposite effects on protein stability at high and low pH values. *Biochim. Biophys Acta* **1794**, 929–935.
86. Natalello A, Liu J, Ami D, et al., (2009) The osmolyte betaine promotes protein misfolding and disruption of protein aggregates. *Proteins* **75**, 509–517.
87. Stancik LM, Stancik DM, Schmidt B, et al., (2002) pH-dependent expression of periplasmic proteins and amino acid catabolism in *Escherichia coli. J. Bacteriol.* **184**, 4246–4258.
88. Oganesyan N, Ankoudinova I, Kim SH, et al., (2007) Effect of osmotic stress and heat shock in recombinant protein overexpression and crystallization. *Protein Expr. Purif.* **52**, 280–285.
89. Diamant S, Rosenthal D, Azem A, et al., (2003) Dicarboxylic amino acids and glycine-betaine regulate chaperone-mediated protein-disaggregation under stress. *Mol. Microbiol.* **49**, 401–410.
90. de Marco A, Vigh L, Diamant S, et al., (2005) Native folding of aggregation-prone recombinant proteins in *Escherichia coli* by osmolytes, plasmid- or benzyl alcohol-overexpressed molecular chaperones. *Cell Stress Chaperones* **10**, 329–339.
91. Barth S, Huhn M, Matthey B, et al., (2000) Compatible-solute-supported periplasmic expression of functional recombinant proteins under stress conditions. *Appl. Environ. Microbiol.* **66**, 1572–1579.
92. Cayley S, Record MT Jr (2003) Roles of cytoplasmic osmolytes, water, and crowding in the response of *Escherichia coli* to osmotic stress: biophysical basis of osmoprotection by glycine betaine. *Biochemistry* **42**, 12596–12609.
93. Tokunaga M, Arakawa T, Tokunaga H (2010) Recombinant expression in moderate halophiles. *Curr. Pharm. Biotechnol.* **11**, 259–266.
94. Picaud S, Olsson ME, Brodelius PE (2007) Improved conditions for production of recombinant plant sesquiterpene synthases in *Escherichia coli. Protein Expr. Purif.* **51**, 71–79.
95. Shi X, Karkut T, Chamankhah M, et al., (2003) Optimal conditions for the expression of a singlechain antibody (scFv) gene in *Pichia pastoris. Protein Expr. Purif.* **28**, 321–330.
96. Lin H, Kim T, Xiong F, et al., (2007) Enhancing the production of Fc fusion protein in fed-batch fermentation of *Pichia pastoris* by design of experiments. *Biotechnol. Prog.* **23**, 621–625.
97. Ohnishi N, Murata N (2006) Glycinebetaine counteracts the inhibitory effects of salt stress on the degradation and synthesis of D1 protein during photoinhibition in *Synechococcus sp* PCC 7942. *Plant Physiol.* **141**, 758–765.
98. Bursy J, Kuhlmann AU, Pittelkow M, et al., (2008) Synthesis and uptake of the compatible solutes ectoine and 5-hydroxyectoine by *Streptomyces coelicolor* A3(2) in response to salt and heat stresses. *Appl. Environ. Microbiol.* **74**, 7286–7296.
99. Kaderbhai N, Karim A, Hankey W, et al., (1997) Glycine-induced extracellular secretion of a recombinant cytochrome expressed in *Escherichia coli. Biotechnol. Appl. Biochem.* **25**, 53–61.
100. Kita Y, Arakawa T (2002) Salts and glycine increase reversibility and decrease aggregation during thermal unfolding of ribonuclease-A. *Biosci. Biotechnol. Biochem.* **66**, 880–882.
101. Li Z, Gu Z, Wang M, et al., (2010) Delayed supplementation of glycine enhances extracellular secretion of the recombinant alpha-cyclodextrin glycosyltransferase in *Escherichia coli. Appl. Microbiol. Biotechnol.* **85**, 553–561.
102. Li ZF, Li B, Liu ZG, et al., (2009) Calcium leads to further increase in glycine-enhanced extracellular secretion of recombinant alpha-cyclodextrin glycosyltransferase in *Escherichia coli. J. Agric. Food Chem.* **57**, 6231–6237.
103. Ou WB, Park YD, Zhou HM (2002) Effect of osmolytes as folding aids on creatine kinase refolding pathway. *Int. J. Biochem. Cell Biol.* **34**,136–147.
104. Meng F, Park Y, Zhou, H (2001) Role of proline, glycerol, and heparin as protein folding aids during refolding of rabbit muscle creatine kinase. *Int. J. Biochem. Cell Biol.* **33**, 701–709.
105. Osváth S, Gruebele M (2003) Proline can have opposite effects on fast and slow protein folding phases. *Biophys. J.* **85**, 1215–1222.
106. Samuel D, Kumar TK, Jayaraman G, et al., (1997) Proline is a protein solubilizing solute. *Biochem. Mol. Biol. Int.* **41**, 235–242.
107. Chattopadhyay MK, Kern R, Mistou MY et al., (2004) The chemical chaperone proline relieves the thermosensitivity of a dnaK deletion mutant at 42 degrees C. *J. Bacteriol.* **186**, 8149–8152.
108. Nagata S, Sasaki H, Oshima A, et al., (2005) Effect of proline and K+on the stimulation of cellular activities in *Escherichia coli* K-12 under high salinity. *Biosci. Biotechnol. Biochem.* **69**, 740–746.

109. Nagata S, Wang Y, Zhang H, et al., (2009) Effect of moderate salinity stress treatment on the stimulation of proline uptake and growth in *Escherichia coli* CSH4 and its mutants under high salinity. *J. Biosci. Bioeng.* **108**, 205–210.
110. Ignatova Z, Gierasch LM (2006) Inhibition of protein aggregation *in vitro* and *in vivo* by a natural osmoprotectant. *Proc. Natl. Acad. Sci. USA.* **103**,13357–13361.
111. Baynes BM, Wang DI, Trout BL (2005) Role of arginine in the stabilization of proteins against aggregation. *Biochemistry* **44**, 4919–4925.
112. Tsumoto K, Ejima D, Kita Y, et al., (2005) Review: Why is arginine effective in suppressing aggregation? *Protein Pept. Lett.* **12**, 613–619.
113. Lange C, Rudolph R (2009) Suppression of protein aggregation by L-arginine. *Curr. Pharm. Biotechnol.* **10**, 408–414.
114. Kim TK, Chung JY, Lee GM, et al., (2001) Arginine-enriched medium composition used for mass-producing recombinant protein in animal cell culture. *patent WO0144442.*
115. Umetsu M, Tsumoto K, Hara M, et al., (2003) How additives influence the refolding of immunoglobulin-folded proteins in a stepwise dialysis system. Spectroscopic evidence for highly efficient refolding of a single-chain Fv fragment. *J. Biol. Chem.* **278**, 8979–8987.
116. Hamada H, Shiraki K (2007) L-argininamide improves the refolding more effectively than L-arginine. *J. Biotechnol.* **130**, 153–160.
117. Georgiou G, Valax P, Ostermeier M, et al., (1994) Folding and aggregation of TEM beta-lactamase: analogies with the formation of inclusion bodies in *Escherichia coli*. *Protein. Sci.* **3**, 1953–1960.
118. Bowden GA, Georgiou G (1990) Folding and aggregation of beta-lactamase in the periplasmic space of *Escherichia coli*. *J. Biol. Chem.* **265**, 16760–16766.
119. Wang A, Robertson AD, Bolen DW (1995) Effects of a naturally occurring compatible osmolyte on the internal dynamics of ribonuclease A. *Biochemistry* **34**, 15096–15104.
120. Jain NK, Roy I (2010) Trehalose and protein stability. *Curr. Protoc. Protein. Sci.* **4**,U49.
121. Conlin LK, Nelson HC (2007) The natural osmolyte trehalose is a positive regulator of the heat-induced activity of yeast heat shock transcription factor. *Mol. Cell Biol.* **27**, 1505–1515.
122. Kandror O, DeLeon A, Goldberg AL (2002) Trehalose synthesis is induced upon exposure of *Escherichia coli* to cold and is essential for viability at low temperatures. *Proc. Natl. Acad. Sci. USA.* **99**, 9727–9732.
123. Davies JE, Sarkar S, Rubinsztein DC (2006) Trehalose reduces aggregate formation and delays pathology in a transgenic mouse model of oculopharyngeal muscular dystrophy. *Hum. Mol. Genet.* **15**, 23–31.
124. Sola-Penna M, Ferreira-Pereira A, Lemos AP, et al., (1997) Carbohydrate protection of enzyme structure and function against guanidinium chloride treatment depends on the nature of carbohydrate and enzyme. *Eur. J. Biochem.* **248**, 24–29.
125. Frye KJ, Royer CA (1997) The kinetic basis for the stabilization of staphylococcal nuclease by xylose. *Protein Sci.* **6**, 789–793.
126. Leandro P, Lechner MC, Tavares de Almeida I, et al., (2001) Glycerol increases the yield and activity of human phenylalanine hydroxylase mutant enzymes produced in a prokaryotic expression system. *Mol. Genet. Metab.* **73**, 173–178.
127. Nascimento C, Leandro J, Lino PR, et al., (2010) Polyol additives modulate the *in vitro* stability and activity of recombinant human phenylalanine hydroxylase. *Appl. Biochem. Biotechnol.* **162**, 192–207.
128. Hohmann S (2002) Osmotic stress signaling and osmoadaptation in yeasts. *Microbiol. Mol. Biol. Rev.* **66**, 300–372.
129. Figler RA, Omote H, Nakamoto RK, et al., (2000) Use of chemical chaperones in the yeast Saccharomyces cerevisiae to enhance heterologous membrane protein expression: high-yield expression and purification of human P-glycoprotein. *Arch. Biochem. Biophys.* **376**, 34–46.
130. Ohnishi T, Ohnishi K, Wang X, et al., (1999) Restoration of mutant TP53 to normal TP53 function by glycerol as a chemical chaperone. *Radiat. Res.* **151**, 498–500.
131. Nascimento C, Leandro J, Tavares de Almeida I, et al., (2008) Modulation of the activity of newly synthesized human phenylalanine hydroxylase mutant proteins by low-molecular-weight compounds. *Protein J.* **27**, 392–400.
132. Ghumman B, Bertram EM, Watts TH (1998) Chemical chaperones enhance superantigen and conventional antigen presentation by HLA-DM-deficient as well as HLA-DM-sufficient antigen-presenting cells and enhance IgG2a production *in vivo*. *J. Immunol.* **161**, 3262–3270.
133. Bai C, Biwersi J, Verkman AS, et al., (1998) A mouse model to test the *in vivo* efficacy of chemical chaperones. *J. Pharmacol. Toxicol. Methods* **40**, 39–45.
134. Voziyan PA, Fisher MT (2002) Polyols induce ATP-independent folding of GroEL-bound bacterial glutamine synthetase. *Arch. Biochem. Biophys.* **397**, 293–297.

135. Sedgwick H, Cameron JE, Poon WC, et al., (2007) Protein phase behavior and crystallization: effect of glycerol. *J. Chem. Phys.* **127**, 125102.
136. Majumder A, Basak S, Raha T, et al., (2001) Effect of osmolytes and chaperone-like action of P-protein on folding of nucleocapsid protein of Chandipura virus. *J. Biol. Chem.* **276**, 30948–30955.
137. Sandee D, Tungpradabkul S, Kurokawa Y, et al., (2005) Combination of Dsb coexpression and an addition of sorbitol markedly enhanced soluble expression of single-chain Fv in *Escherichia coli. Biotechnol. Bioeng.* **91**, 418–424.
138. Wei H, Fan Y, Gao YQ (2010) Effects of urea, tetramethyl urea, and trimethylamine N-oxide on aqueous solution structure and solvation of protein backbones: a molecular dynamics simulation study. *J. Phys. Chem. B* **114**, 557–568.
139. Hu CY, Lynch GC, Kokubo H, et al., (2010) Trimethylamine N-oxide influence on the backbone of proteins: an oligoglycine model. *Proteins* **78**, 695–704.
140. Kumar R, Serrette JM, Thompson EB (2005) Osmolyte-induced folding enhances tryptic enzyme activity *Arch. Biochem. Biophys.* **436**, 78–82.
141. Tatzelt J, Prusiner SB, Welch WJ (1996) Chemical chaperones interfere with the formation of scrapie prion protein. *EMBO J.* **15**, 6363–6373.
142. Song JL, Chuang DT (2001) Natural osmolyte trimethylamine N-oxide corrects assembly defects of mutant branched-chain alphaketoacid decarboxylase in maple syrup urine disease. *J. Biol. Chem.***276**, 40241–40246.
143. Samuelsson E, Jonasson P, Viklund F, et al., (1996) Affinity-assisted *in vivo* folding of a secreted human peptide hormone in *Escherichia coli. Nat. Biotechnol.* **14**, 751–755.
144. Buchner J, Pastan I, Brinkmann U (1992) A method for increasing the yield of properly folded recombinant fusion proteins: single-chain immunotoxins from renaturation of bacterial inclusion bodies. *Anal. Biochem.* **205**, 263–270.
145. Inaba K (2009) Disulfide bond formation system in *Escherichia coli. J. Biochem.* **146**, 591–597.
146. Hwang C, Sinskey AJ, Lodish HF (1992) Oxidized redox state of glutathione in the endoplasmic reticulum. *Science* **257**, 1496–1502.
147. Bardwell JC (1994) Building bridges: disulphide bond formation in the cell. *Mol. Microbiol.* **14**, 199–205.
148. Beld J, Woycechowsky KJ, Hilvert D (2010) Small-molecule diselenides catalyze oxidative protein folding *in vivo. ACS Chem. Biol.* **5**, 177–182.
149. Beld J, Woycechowsky KJ, Hilvert D (2007) Selenoglutathione: Efficient oxidative protein folding by a diselenide. *Biochemistry* **46**, 5382–5390.
150. Lees WJ (2008) Small-molecule catalysts of oxidative protein folding. *Curr. Opin. Chem. Biol.* **12**, 740–745.
151. Bushmarina NA, Blanchet CE, Vernier G, et al., (2006) Cofactor effects on the protein folding reaction: acceleration of alpha-lactalbumin refolding by metal ions. *Protein Sci.* **15**, 659–671.
152. Wittung-Stafshede P (2002) Role of cofactors in protein folding. *ACC Chem. Res.* **35**, 201–208.
153. Zou J, Sugimoto N (2000) Complexation of peptide with Cu^{2+} responsible to inducing and enhancing the formation of alpha-helix conformation. *Biometals* **13**, 349–359.
154. Hiniker A, Collet JF, Bardwell JC (2005) Copper stress causes an *in vivo* requirement for the *Escherichia coli* disulfide isomerase DsbC. *J. Biol. Chem.* **280**, 33785–33791.
155. Beck R, Burtscher H (1994) Expression of human placental alkaline phosphatase in *Escherichia coli. Protein Expr. Purif.* **5**, 192–197.
156. Khodagholi F, Yazdanparast R (2008) Designing a highly efficient refolding system for alkaline phosphatase using combination of cyclodextrin and Mg^{2+} ion. *Protein J.* **27**, 1–6.
157. Baneyx F, Ayling A, Palumbo T, et al., (1991) Optimization of growth conditions for the production of proteolytically-sensitive proteins in the periplasmic space of *Escherichia coli. Appl. Microbiol. Biotechnol.* **36**, 14–20.
158. Suno R, Taguchi H, Masui R, et al., (2004) Trigger factor from Thermus thermophilus is a Zn^{2+}-dependent chaperone. *J. Biol. Chem.* **279**, 6380–6384.
159. Solovyov A, Gilbert HF (2004) Zinc-dependent dimerization of the folding catalyst, protein disulfide isomerase. *Protein Sci.* **13**, 1902–1907.
160. Zhang L, Lookene A, Wu G, et al., (2005) Calcium triggers folding of lipoprotein lipase into active dimers. *J. Biol. Chem.* **280**, 42580–42591.
161. Jones HE, Holland IB, Campbell AK (2002) Direct measurement of free Ca^{2+} shows different regulation of Ca^{2+} between the periplasm and the cytosol of *Escherichia coli. Cell Calcium* **32**, 183–192.

162. Machida S, Ogawa S, Xiaohua S, et al., (2000) Cycloamylose as an efficient artificial chaperone for protein refolding. *FEBS Lett.* **486**, 131–135.
163. Dong XY, Shi JH, Sun Y (2002) Cooperative effect of artificial chaperones and guanidinium chloride on lysozyme renaturation at high concentrations. *Biotechnol. Prog.* **18**, 663–665.
164. Sasaki Y, Akiyoshi K (2010) Development of an artificial chaperone system based on cyclodextrin. *Curr. Pharm. Biotechnol.* **11**, 300–305.
165. Bajorunaite E, Cirkovas A, Radzevicius K, et al., (2009) Anti-aggregatory effect of cyclodextrins in the refolding process of recombinant growth hormones from *Escherichia coli* inclusion bodies. *Int. J. Biol. Macromol.* **44**, 428–434.
166. Laslo AC, Ganea E, Obinger C (2009) Refolding of hexameric porcine leucine aminopeptidase using a cationic detergent and dextrin-10 as artificial chaperones. *J. Biotechnol.* **140**, 162–168.
167. Woycechowsky KJ, Wittrup KD, Raines RT (1999) A small-molecule catalyst of protein folding *in vitro* and *in vivo*. *Chem. Biol.* **6**, 871–879.
168. Winter J, Lilie H, Rudolph R (2002) Recombinant expression *in vitro* folding of proinsulin are stimulated by the synthetic dithiol Vectrase-P. *FEMS Microbiol. Lett.* **213**, 225–230.
169. Sawangwan T, Goedl C, Nidetzky B (2010) Glucosylglycerol and glucosylglycerate as enzyme stabilizers. *Biotechnol. J.* **5**, 187–191.
170. Sauer T, Patel M, Chan CC, et al., (2008) Unfolding the Therapeutic Potential of Chemical Chaperones for Age-related Macular Degeneration. *Expert Rev. Ophthalmol.* **3**, 29–42.
171. Loo TW, Clarke DM (2007) Chemical and pharmacological chaperones as new therapeutic agents. *Expert. Rev. Mol. Med.* **9**, 1–18.
172. Arakawa T, Ejima D, Kita Y, et al., (2006) Small molecule pharmacological chaperones: From thermodynamic stabilization to pharmaceutical drugs. *Biochim. Biophys. Acta* **1764**, 1677–1687.
173. Kim HK, Choi SI, Seong BL (2010) 5 S rRNA-assisted DnaK refolding. *Biochem. Biophys. Res. Commun.* **391**, 1177–1181.
174. Jana S, Deb JK (2005) Strategies for efficient production of heterologous proteins in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **67**, 289–298.
175. Neubauer P, Fahnert B, Lilie H, et al., (2006) Protein Inclusion Bodies in Recombinant Bacteria. In: Shively JM (ed) Inclusions in Prokaryotes, *Microbiology Monographs*, **1**, Springer, Berlin.
176. Gasser B, Saloheimo M, Rinas U et al., (2008) Protein folding and conformational stress in microbial cells producing recombinant proteins: a host comparative overview. *Microb. Cell Fact.* **7**, 11.
177. Schwarz E, Rudolph R, Ambrosius D, et al., (2001) Process for the production of naturally folded and secreted proteins by co-secretion of molecular chaperones. *patent EP1077262*.
178. Schroeckh V, Hortschansky P, Fricke S, et al., (2000) Expression of soluble, recombinant alphavbeta3 integrin fragments in *Escherichia coli*. *Microbiol. Res.* **155**, 165–177.
179. Glockshuber R, Skerra A, Rudolph R, et al., (1992) Improvement of the secretion yield of proteins with a disulfide bridge. *Patent EP0510658*.
180. Berrow NS, Büssow K, Coutard B, et al., (2006) Recombinant protein expression and solubility screening in *Escherichia coli*: a comparative study. *Acta Crystallogr. D. Biol. Crystallogr.* **62**, 1218–1226.
181. Thomas JG, Baneyx F (1997) Divergent effects of chaperone overexpression and ethanol supplementation on inclusion body formation in recombinant *Escherichia coli*. *Protein Expr. Purif.* **11**, 289–296.
182. Gustafson C, Tagesson C (1985) Influence of organic solvent mixtures on biological membranes. *Br. J. Ind. Med.* **42**, 591–595.
183. Hasan CM, Shimizu K (2008) Effect of temperature up-shift on fermentation and metabolic characteristics in view of gene expressions in *Escherichia coli*. *Microb. Cell Fact.* **7**, 35.
184. Hartl FU, Hayer-Hartl M (2002) Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* **295**, 1852–1858.
185. Hartl FU, Hayer-Hartl M (2009) Converging concepts of protein folding *in vitro* and *in vivo*. *Nat. Struct. Mol. Biol.* **16**, 574–581.
186. Schlieker C, Bukau B, Mogk A (2002) Prevention and reversion of protein aggregation by molecular chaperones in the *Escherichia coli* cytosol: implications for their applicability in biotechnology. *J. Biotechnol.* **96**, 13–21.
187. Kolaj O, Spada S, Robin S, et al., (2009) Use of folding modulators to improve heterologous protein production in *Escherichia coli*. *Microb. Cell Fact.* **8**, 9.
188. Rinas U, Hoffmann F, Betiku E, et al., (2007) Inclusion body anatomy and functioning of chaperone-mediated *in vivo* inclusion body disassembly during high-level recombinant protein production in *Escherichia coli*. *J. Biotechnol.* **127**, 244–257.
189. Sugimoto S, Saruwatari K, Higashi C, et al., (2008) The proper ratio of GrpE to DnaK is

important for protein quality control by the DnaK-DnaJ-GrpE chaperone system and for cell division. *Microbiology* **154**, 1876–1885.

190. Balagurunathan B, Jayaraman G (2008) Theoretical and experimental investigation of chaperone effects on soluble recombinant proteins in *Escherichia coli*: effect of free DnaK level on temperature-induced recombinant streptokinase production. *Syst. Synth. Biol.* **2**, 27–48.
191. Hoffmann F, Rinas U (2004) Roles of heat-shock chaperones in the production of recombinant proteins in *Escherichia coli. Adv. Biochem. Eng. Biotechnol.* **89**, 143–161.
192. García-Fruitós E, Martínez-Alonso M, Gonzàlez-Montalbán N, et al., (2007) Divergent genetic control of protein solubility and conformational quality in *Escherichia coli. J. Mol. Biol.* **374**, 195–205.
193. Fahnert B (2001) Rekombinantes humanes BMP-2 aus *Escherichia coli* – Strategien zur Expression und Funktionalisierung. *Doctorate Thesis.* Friedrich-Schiller-University Jena (Germany).
194. Bonomo J, Welsh JP, Manthiram K, et al., (2010) Comparing the functional properties of the Hsp70 chaperones, DnaK and BiP. *Biophys. Chem.* **149**, 58–66.
195. Ying BW, Taguchi H, Ueda H, et al., (2004) Chaperone-assisted folding of a single-chain antibody in a reconstituted translation system. *Biochem. Biophys. Res. Commun.* **320**, 1359–1364.
196. Ellis RJ (2005) Chaperomics: *in vivo* GroEL function defined. *Curr. Biol.* **15**, 661–663.
197. Kerner MJ, Naylor DJ, Ishihama Y, et al., (2005) Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli. Cell* **122**, 209–220.
198. Kim SY, Ayyadurai N, Heo MA, et al., (2009) Improving the productivity of recombinant protein in *Escherichia coli* under thermal stress by coexpressing GroELS chaperone system. *J. Microbiol. Biotechnol.* **19**, 72–77.
199. Naletova IN, Muronetz VI, Schmalhausen EV (2006) Unfolded, oxidized, and thermoinactivated forms of glyceraldehyde-3-phosphate dehydrogenase interact with the chaperonin GroEL in different ways. *Biochim. Biophys. Acta* **1764**, 831–838.
200. Carrió MM, Villaverde A (2003) Role of molecular chaperones in inclusion body formation. *FEBS Lett.* **537**, 215–221.
201. Mogk A, Schlieker C, Friedrich KL, et al., (2003) Refolding of substrates bound to small Hsps relies on a diaggregation reaction mediated most efficiently by ClpB/ DnaK. *J. Biol. Chem.* **278**, 31033–31034.
202. Petersson L, Carrió MM, Vera A, et al., (2004) The impact of *dnaKJ* overexpression on recombinant protein solubility results from antagonistic effects on the control of protein quality. *Biotechnol. Lett.* **26**, 595–596.
203. González-Montálban N, Natalello A, García-Fruitós E, et al., (2008) In situ protein folding and activation in bacterial inclusion bodies. *Biotechnol. Bioeng.* **100**, 797–802.
204. Kitagawa M, Miyakawa M, Matsumura Y, et al., (2002) *Escherichia coli* small heat shock proteins, IbpA and IbpB, protect enzymes from inactivation by heat and oxidants. *Eur. J. Biochem.* **269**, 2907–2911.
205. Sorensen HP, Mortensen KK (2005) Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli. Microb. Cell. Fact.* **4**, 1.
206. González-Montálban N, García-Fruitós E, Villaverde A (2007) Recombinant protein solubility – does more mean better? *Nat. Biotechnol.* **25**, 718–720.
207. de Marco A (2007) Protocol for preparing proteins with improved solubility by co-expressing with molecular chaperones in *Escherichia coli. Nat. Protoc.* **2**, 2632–2639.
208. Voziyan PA, Jadhav L, Fisher MT (2000) Refolding a glutamine synthetase truncation mutant *in vitro*: identifying superior conditions using a combination of chaperonins and osmolytes. *J. Pharm. Sci.* **89**, 1036–1045.
209. Fisher MT, Voziyan P (2002) Chaperonin and osmolyte protein folding and related screening methods. *Patent US2002006636.*
210. Voziyan PA, Johnston M, Chao A, et al., (2005) Designing a high throughput refolding array using a combination of the GroEL chaperonin and osmolytes. *J. Struct. Funct. Genomics* **6**, 183–188.
211. Naik S, Haque I, Degner N, et al., (2010) Identifying protein stabilizing ligands using GroEL. *Biopolymers* **93**, 237–251.
212. George DF, Bilek MM, McKenzie DR (2008) Detecting and exploring partially unfolded states of proteins using a sensor with chaperone bound to its surface. *Biosens. Bioelectron.* **24**, 969–975.
213. Wisén S, Gestwicki JE (2008) Identification of small molecules that modify the protein folding activity of heat shock protein 70. *Anal. Biochem.* **374**, 371–377.
214. Brigotti M, Petronini PG, Carnicelli D, et al., (2003) Effects of osmolarity, ions and compatible osmolytes on cell-free protein synthesis. *Biochem. J.* **369**, 369–374.
215. Rösgen J, Pettitt BM, Bolen DW (2007) An analysis of the molecular origin of osmolyte-

dependent protein stability. *Protein Sci.* **16**, 733–743.

216. Auton M, Bolen DW, Rösgen J (2008) Structural thermodynamics of protein preferential solvation: osmolyte solvation of proteins, aminoacids, and peptides. *Proteins* **73**, 802–813.
217. O'Brien EP, Ziv G, Haran G, et al., (2008) Effects of denaturants and osmolytes on proteins are accurately predicted by the molecular transfer model. *Proc. Natl. Acad. Sci. USA.* **105**, 13403–13408.
218. Javid N, Vogtt K, Krywka C, et al., (2007) Protein–protein interactions in complex cosolvent solutions. *Chem. Phys. Chem.* **8**, 679–689.
219. Holthauzen LM, Bolen DW (2007) Mixed osmolytes: the degree to which one osmolyte affects the protein stabilizing ability of another. *Protein Sci.* **16**, 293–298.
220. Ko YH, Pedersen P (2001) Methods for identifying an agent that corrects defective protein folding. *Patent WO0121652.*
221. Kirsch T, Sebald W, Dreyer MK (2000) Crystal structure of the BMP-2-BRIA ectodomain complex. *Nat. Struct. Biol.* **7**, 492–496.
222. Kirsch T, Nickel J, Sebald W (2000) Isolation of recombinant BMP receptor IA ectodomain and its 2:1 complex with BMP-2. *FEBS Lett.* **468**, 215–219.
223. Fahnert B, Hahn D, Guthke R (2002) Knowledge-based assessment of gene expression data from chemiluminescence detection. *J. Biotechnol.* **94**, 23–35.
224. Smith HE (2007) The transcriptional response of *Escherichia coli* to recombinant protein insolubility. *J. Struct. Funct. Genomics* **8**, 27–35.
225. Kol S, Merlo ME, Scheltema RA, et al., (2010) Metabolomic characterization of the salt stress response in *Streptomyces coelicolor. Appl. Environ. Microbiol.* **76**, 2574–2581.
226. Han MJ, Lee SY, Koh ST, et al., (2010) Biotechnological applications of microbial proteomes. *J. Biotechnol.* **145**, 341–349.
227. Aldor IS, Krawitz DC, Forrest W, et al., (2005) Proteomic profiling of recombinant *Escherichia coli* in high-cell-density fermentations for improved production of an antibody fragment biopharmaceutical. *Appl. Environ. Microbiol.* **71**, 1717–1728.
228. Harcum SW, Haddadin FT (2006) Global transcriptome response of recombinant *Escherichia coli* to heat-shock and dual heat-shock recombinant protein induction. *J. Ind. Microbiol. Biotechnol.* **33**, 801–814.
229. Kraft M, Knüpfer U, Wenderoth R, et al., (2007) An online monitoring system based on a synthetic sigma32-dependent tandem promoter for visualization of insoluble proteins in the cytoplasm of *Escherichia coli. Appl. Microbiol. Biotechnol.* **75**, 397–406.
230. Kraft M, Knüpfer U, Wenderoth R, et al., (2007) A dual expression platform to optimise the soluble production of heterologous proteins in the periplasm of *Escherichia coli. Appl. Microbiol. Biotechnol.* **76**, 1413–1422.
231. Ignatova Z, Gierasch LM (2004) Monitoring protein stability and aggregation *in vivo* by real-time fluorescent labeling. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 523–528.
232. Ignatova Z (2005) Monitoring protein stability *in vivo. Microb. Cell. Fact.* **4**, 23.
233. Mansell TJ, Linderman SW, Fisher AC, et al., (2010) A rapid protein folding assay for the bacterial periplasm. *Protein Sci.* **19**, 1079–1090.
234. Chow MK, Amin AA, Fulton KF, et al., (2006) REFOLD: an analytical database of protein refolding methods. *Protein Expr. Purif.* **46**, 166–171.
235. Buckle AM, Devlin GL, Jodun RA, et al., (2005) The matrix refolded. *Nat. Methods* **2**, 3.
236. Chow MK, Amin AA, Fulton KF, et al., (2006) The REFOLD database: a tool for the optimization of protein expression and refolding. *Nucleic Acids Res.* **34**, D207–D212.
237. Braun P, Hu Y, Shen B, et al., (2002) Proteome-scale purification of human proteins from bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 2654–2659.
238. Dieckman LJ, Hanly WC, Collart ER (2006) Strategies for high-throughput gene cloning and expression. *Genet. Eng.* **27**,179–190.
239. Park CG, Kim TW, Oh IS, et al., (2009) Expression of functional *Candida antarctica* lipase B in a cell-free protein synthesis system derived from *Escherichia coli. Biotechnol. Prog.* **25**, 589–593.
240. Frey S, Haslbeck M, Hainzl O, et al., (2008) Synthesis and characterization of a functional intact IgG in a prokaryotic cell-free expression system. *J. Biol. Chem.* **389**, 37–45.
241. Kang SH, Kim DM, Kim HJ, et al., (2005) Cell-free production of aggregation-prone proteins in soluble and active forms. *Biotechnol. Prog.* **21**, 1412–1419.
242. He M (2008) Cell-free protein synthesis: applications in proteomics and biotechnology. *Nat. Biotechnol.* **25**, 126–132.
243. Fahnert B (2004) Folding promoting agents in recombinant protein production. *Methods Mol. Biol.* **267**, 53–74.

Chapter 2

Routine Identity Confirmation of Recombinant Proteins by MALDI-TOF Mass Spectrometry

Brett J. Savary and Prasanna Vasu

Abstract

Peptide mass fingerprinting (PMF) by matrix-assisted laser-desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) provides a simple and direct means to unequivocally confirm identity of recombinant proteins based on predicted peptide profiles. Many universities or research institutions now carry mass spectrometry instrumentation as part of their core bioanalytical facilities or provide public service to outside investigators. This chapter provides methods we have used to generate routinely high quality samples for MALDI-TOF MS analysis. Following resolution of protein preparations by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), we easily process sets of 12 samples manually for MS analysis. Target bands are alkylated and digested in-gel with trypsin, followed by extraction of peptides and desalting with a C18 adsorbent resin (e.g., a "ZipTips"). Acquisition of PMF data on MALDI-TOF mass spectrometers is fast, and with on-site instrumentation, the entire process can be completed within 2 days.

Key words: Peptide mass fingerprint, Protein identification, MALDI-TOF MS, Trypsin digestion, SDS-PAGE

1. Introduction

Preparation of a recombinant protein generally involves separating the protein of interest from the bulk of host cell proteins, purification by means of sequential chromatography (typically at least three steps), followed by biochemical and structural analysis. If the recombinant protein is a fusion protein containing an affinity tag such as a C-terminal poly-His sequence, it may be directly purified by one-step metal chelation affinity chromatography to get milligram quantities of pure proteins. The purity of the recombinant protein is usually evaluated by SDS-PAGE, which also provides critical information

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_2, © Springer Science+Business Media, LLC 2012

for structural analysis and identification (1). The identity of putative recombinant protein band(s) may be confirmed by electroblotting to a suitable membrane, then probing by immunodetection (2) or by N-terminal amino acid sequence analysis (3). Immunodetection requires a suitably selective antiserum be available for the target protein. Matching the N-terminal amino acid sequence requires the N-terminal amine not be chemically modified, which blocks the Edman sequencing chemistry and results in no sequence information. In cases like this, more complex processing of samples is required to obtain internal peptide sequences for matching (4). This can take considerable time and expense. With the proliferation of high resolution biomolecule mass spectrometers over the past 10 years or so, a relatively simple and direct means for identifying proteins is available. This is based on the highly reproducible digestion of proteins with trypsin and acquiring highly accurate mass spectra of peptides (5), a process termed as peptide mass fingerprinting (PMF) (6). PMF by matrix-assisted laser-desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) provides direct structural identification by matching highly accurate masses determined from peptide ions predicted peptide profiles determined from the known amino acid sequence. Mass spectrometry is now the method of choice for direct and unequivocal identification of a protein and it can readily distinguish between closely related isoforms (7).

Universities and research institutions now commonly carry a MALDI-TOF MS instrument as part of proteomic or core analytical facilities. Many also provide public service to outside investigators for routine protein identification (e.g., see Association of Biomolecular Resource Facilities, http://www.abrf.org/index.cfm/dir.yp). Commercial bioanalytical service companies can also be readily located via an internet search. MALDI-TOF mass spectrometers are highly accurate (10–50 ppm) with resolution (FWHM, full width half maximum) better than 12,000 in reflectron mode with sensitivity for detecting sub-femtomole levels of peptides. Owing to this high sensitivity, it is critical that high quality (clean) samples be prepared to enable ready detection and matching of protein-specific peptides – i.e., samples must be free of exogenously introduced proteins. Particularly, contaminating peptides from proteins such as keratins (from skin) introduced by direct contact or dust particles, or from microbial contamination in old buffer solutions, can overwhelm the MS spectrum, obscuring the peptides from the target protein.

This chapter describes procedures we have used routinely to generate high quality samples for MALDI-TOF MS identification of recombinant proteins (see Chapter 21). Trichloroacetic acid (TCA) precipitation of soluble protein in extracts (e.g., from plant tissues, yeast cells, or culture media) is used to remove salts and other interfering substances prior to separation by SDS-PAGE. Sample loading on PAGE is calibrated to provide ca. 1 μg of a target

protein resolved as a gel band. Individual bands are excised from a PAGE gel and then transferred to a microcentrifuge tube for alkylation, digestion with trypsin, and recovery of the resultant peptides. Extracted peptides may be immediately desalted and spotted with matrix solution onto target plates for immediate on-site MALDI-TOF MS analysis, or the extracted peptides may be stored in dry form until sending to a bioanalytical facility for final processing and analysis. We routinely process (manually) up to 12 protein samples at a time from a standard SDS-PAGE gel, and the sample processing and analysis can be completed in 2 days with a local instrument. The chapter concludes with a description for using PMF data to confirm identification of a protein.

2. Materials

2.1. General Considerations

Minimally detectable protein bands (at least 0.1 μg) taken from SDS-PAGE gels stained with (any) Coomassie Blue stain generally provide good yields for strong peptide ion signals by MALDI-TOF MS. We describe here our use of Invitrogen's NuPAGE gels with Bis-Tris buffer system, but any manufacturer's system for SDS-PAGE (or original Laemmli buffer system) can perform equally well. Similarly, proteins spots from two-dimensional PAGE gels can be used. What is critical for MS identification is to use ultraclean water in preparing solutions, prevent inadvertent dust or protein contamination, or to introduce chemicals that suppress peptide ionization (detergents, non-volatile buffers, plasticizers, salts) (see Note 1).

2.2. General

1. 0.15% Sodium deoxycholate: 10 mL working solution.
2. 75% TCA: 10 mL working solution from 100% reagent; (e.g., Sigma-Aldrich).
3. Acetone: pre-chilled at –20°C.
4. SpeedVac® or similar vacuum concentrator.
5. Sonication bath (optimal, but not essential).
6. Water bath (temperature at 70°C or 37°C).
7. NuPAGE Novex BisTris gel: 12% monomer (Invitrogen).
8. 4× NuPAGE LDS sample loading buffer (Invitrogen).
9. 10× Reducing agent (Invitrogen).
10. Antioxidant (Invitrogen).
11. 20× NuPAGE MOPS-SDS running buffer (Invitrogen).
12. XCell Surelock mini electrophoresis cell (Invitrogen).
13. Gel loading tips.
14. Electrophoresis power pack.

15. Coomassie-Blue G-250 stain: a pre-mixed colloidal dye solution such as SimplyBlue™SafeStain (Invitrogen).
16. Acetonitrile (ACN), LC grade or better.
17. 50% ACN with 0.1% trifluoracetic acid (TFA): 100 mL ACN, 0.2 mL TFA, and 99.8 mL water.
18. 0.1% TFA: 0.1 mL TFA and 99.9 mL water.
19. 1 M ammonium bicarbonate (NH_4HCO_3) stock: 0.7906 g NH_4HCO_3 dissolved in 10 mL, pH 7.8.
20. 100 mM NH_4HCO_3 in 50% ACN: 1.0 mL 1 M stock, 5 mL ACN, and 4 mL water.
21. 40 mM NH_4HCO_3 in 10% ACN: 0.40 mL 1 M stock, 1 mL ACN, and 8.6 mL water.
22. 5% TFA in 50% ACN: 0.50 mL TFA, 5 mL ACN, and 4.5 mL water.
23. Trypsin Gold (Promega) or other mass spectrometry grade TCKP-treated trypsin.
24. 50 mM acetic acid: 0.0287 mL glacial acetic acid into 9.97 mL water.
25. 10 mM dithiothreitol (DTT) in 25 mM NH_4HCO_3: 15.4 mg/10 mL, prepared ahead and stored frozen in 0.5 mL aliquots.
26. 25 mM NH_4HCO_3: 19.8 mg NH_4HCO_3 in 10 mL water. Important: always make up fresh.
27. 30 mM iodoacetamide (IAA) in 25 mM NH_4HCO_3: 11.1 mg in 2.0 mL (always prepare fresh).
28. C_{18} ZipTips (Millipore, ZTC18S096).
29. Stainless steel MALDI target plate (matches local instrument).

2.3. Gel Electrophoresis Buffer Preparation

1. 1× NuPAGE® SDS Running Buffer: prepare 800 mL by adding 40 mL of 20× MOPS NuPAGE® SDS Running Buffer to 760 mL of deionized water (dH_2O). Mix thoroughly and set aside 600 mL of this solution for use in the lower (*outer*) buffer chamber of the XCell SureLock Mini-Cell.
2. Immediately, prior to electrophoresis, mix 500 µL of NuPAGE® Antioxidant into the remaining 200 mL of 1× Running Buffer for use in the upper (*inner*) buffer chamber of the XCell SureLock Mini-Cell.

2.4. Matrix and Calibration Standards Preparation for MALDI-TOF MS Analysis

1. CHCA (matrix) solution: dissolve 10 mg of recrystallized α-cyano-4-hydroxy-cinnamic acid (CHCA, Sigma Aldrich) into 1 mL of solution containing 49.5% ethanol, 49.5% ACN, and 1% aqueous 0.1% TFA. Prepare and use as freshly made solution.

2. Sodiated PEG (polyethylene glycol) for instrument mass calibration standards: Prepare 10 mg/mL stock solutions for each PEG in 50% ACN, then for final working solution, mix 1 μL PEG 1000, 2 μL PEG 2000, and 6 μL PEG 3000 with 9 μL of CHCA solution, then add 3 μL of 2 mg/mL of sodium iodide.
3. Adrenocorticotropic hormone (ACTH; 18–39 clip, 2,465. 1989 Da, Sigma-Aldrich) for instrument resolution and sensitivity testing: Mix 1 μL 1 mg/mL ACTH with 39.5 μL of 0.1% TFA to give 10 pmol/μL. Mix 1 μL of 10 pmol/μL solution with 9 μL of 0.1% TFA to give 1 pmol/μL. Mix 1:1 with CHCA solution to give 500 fmol/μL.
4. GluFib peptide B (GFPB; 1,570.6774 Da, Sigma-Aldrich) for instrument lock mass: Mix 1 μL of 1 mg/mL GFPB with 62.7 μL of 0.1% TFA to give 10 pmol/μL. Mix 1 μL of 10 pmol/μL solution with 9 μL of 0.1% TFA to give 1 pmol/μL. Mix 1:1 with CHCA solution to give 500 fmol/μL.

3. Methods

Protein samples are initially prepared for SDS-PAGE by TCA precipitation in the presence of sodium dioxycholate (see Subheading 3.1). This step is used to stabilize and concentrate proteins in samples extracted from cells and tissues. Recombinant enzymes secreted into culture media, as described in Chapter 21 of this book, were prepared and analyzed by the methods provided here. Your target protein(s) is assumed to be expressed in a soluble form and in sufficient yield that it is readily observable by SDS-PAGE (see Note 2). If SDS-PAGE methods (Subheading 3.2) are routinely used in your laboratory, and your protein of interest is observable as a band on a Commassie Blue stained gel, then, proceed to Subheading 3.3 for trypsin digestion. The dried peptides recovered from the trypsin digestion can be submitted directly to a MALDI-TOF MS facility, or if you have access to a local instrument, Subheading 3.4 describes sample clean up. The process for sample preparation is summarized in Fig. 1. Subheading 3.5 introduces use of the Mascot search engine with PMF data acquired from your samples to established protein identity.

3.1. Trichloroacetic Acid Precipitation of Protein Samples (8)

1. Pipet a volume from samples determined to contain 10–100 μg of total protein into a 0.5-mL microcentrifuge tube and add water or extraction buffer to provide a final volume of 100 μL (see Note 3).
2. Add 10 μL of 0.15% sodium deoxycholate, mix well, and place on ice for 10 min.
3. Add 10 μL of 75% TCA, mix well, and place on ice for 20 min.

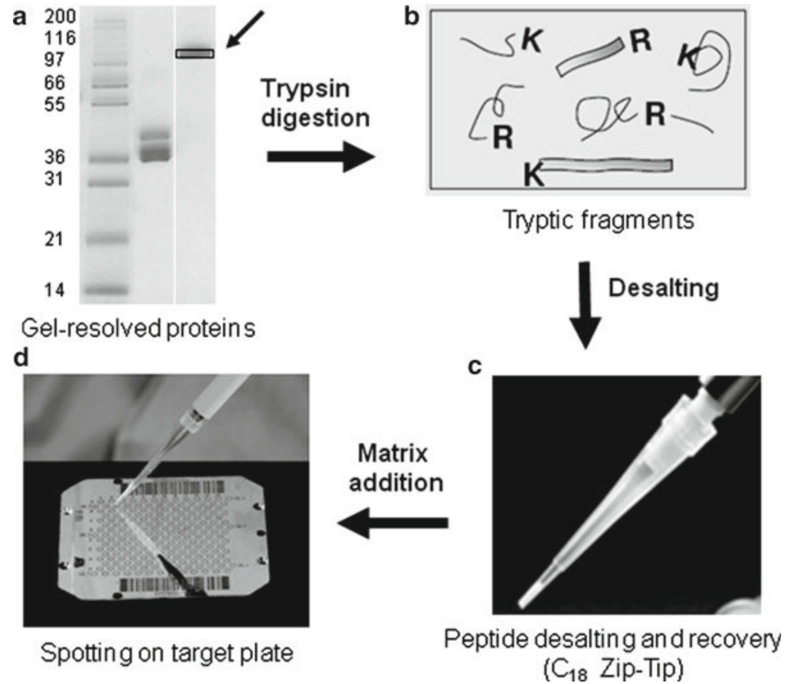


Fig. 1. Sample preparation for MALDI-TOF MS analysis. The selected protein band (*indicated by small arrow*) is excised from the SDS-PAGE gel (**a**) and transferred to a microcentrifuge tube to be destained and digested with trypsin. The digested fragments (**b**) are desalted using a ZipTip (**c**) and are directly eluted into the CHCA matrix solution, which is then spotted onto the target plate. (**d**) Lanes in the SDS-PAGE gel are molecular weight markers and two recombinant xylanolytic enzymes (β-xylanase C, AN1818.2; and β-xylosidase, AN2359.2; which are described in Chapter 21).

4. Centrifuge in a microcentrifuge at ca. 17,000×*g* for 10 min and immediately decant or siphon off the supernatant (carefully draw from the side opposite of the pellet). The pellet will appear as a slight white residue along the outer side of the microcentrifuge tube.
5. Wash pellet with 0.25 mL of cold acetone. Incubate at –20°C for 20 min (see Note 4).
6. Centrifuge again for 10 min and aspirate acetone from the pellet. Apply vacuum in a SpeedVac briefly to remove traces of solvent, but do not dry excessively. At this stage, you can cap and store samples at –20°C until ready to run electrophoresis.
7. Dissolve the dried pellet in 50 μL of 1× LDS sample loading buffer using vigorous vortexing or sonication – the buffer solution should be blue in color (see Note 5).
8. Heat samples at 70°C for 10 min, then place on ice. Centrifuge again before loading on to an electrophoresis gel.

3.2. SDS-PAGE for Separating Proteins (9)

1. Prepare precast BisTris page gel and assemble electrophoresis cell (see Note 6).
2. Fill fully the upper (inner) buffer chamber with 1× running buffer containing antioxidant. Confirm there is a good seal – no immediate leakage is observed from the chamber.
3. Fill the lower (outer) buffer chamber with 600 mL of 1× running buffer.
4. Using a pipet with gel loading tips, load protein samples, control proteins, and molecular mass markers (see Note 7).
5. Run power with 200 V constant for 55 min.
6. Rinse the gel three times for 5 min each with 100 mL of dH_2O (discard each rinse). Be sure gloves are on and rinsed before handling gel!
7. Stain the gel with *ca.* 25 mL with Coomassie Brilliant Blue for up to 1 h at room temperature (RT) with gentle mixing.
8. Discard the staining solution, rinse the gel briefly in dH_2O and then continue destaining for 1 h with gentle mixing. Repeat with a second 100 mL volume to further clear gel background.
9. Transfer gel to a Zip-lock bag and obtain a gel image, such as by flatbed scanner (600 dpi) (Fig. 1). Gel can be stored in temporarily at 4°C (1 or 2 days) before further processing by trypsin digestion in the next section.

3.3. In Gel Trypsin Digestion of Protein Samples (10)

1. Place the gel on a clean glass cutting surface. Cut out the target protein bands using a new razor blade. Transfer the gel slices (see Note 8) to a prewashed 0.5-mL microcentrifuge tubes (see Note 9).
2. Destain gel pieces by adding 200 μL of 100 mM NH_4HCO_3 in 50% ACN, shake well, cap the tubes, and incubate at 37°C for 45 min. Spin down briefly; then pipet off (or use vacuum siphon) supernatant and discard. Repeat this step once.
3. Add 100 μL of 10 mM DTT in 25 mM NH_4HCO_3 at 37°C for 30 min to reduce proteins in gel pieces. Remove the supernatant.
4. Add 100 μL of 30 mM IAA in 25 mM NH_4HCO_3 and incubate *in the dark* at 37°C for 30 min to alkylate cysteine residues. Remove the supernatant.
5. Add 200 μL of 25 mM NH_4HCO_3 for 15 min at RT with occasional mixing to wash gel pieces. Spin down the gel slices and discard the supernatant. Repeat this twice.
6. Add 100 μL of 100% ACN for 10 min at RT to dehydrate gel pieces. The gel pieces will shrink and become "whitish." Briefly spin down and remove the solvent, then dry gel pieces in the SpeedVac (for 10–15 min – no heat). Samples can be stored temporarily in this dried form (longer term storage with desiccant in the freezer).

7. Add 25 μL of trypsin working solution (or about 1.5× the original gel slice volume) to each tube, cap tightly, and incubate for 1 h at RT (see Note 10). The slices will rehydrate during this time period promoting trypsin diffusion into the gel slice. If the gel slices appear white or opaque after 1 h, add an additional 10 μL of trypsin working solution and further incubate 30 min.
8. Remove excess trypsin solution that is not absorbed into the gel piece, then add 50 μL of 40 mM NH_4HCO_3 in 10% ACN. Cap the tubes tightly and incubate overnight at 37°C.
9. The next day, add 150 μL of ultrapure water to the gel slice containing tryptic digests and incubate for 15 min with frequent vortex mixing (use sonication bath, if available). Spin down briefly and transfer the supernatant containing the extracted peptides to a new pre-washed and labeled microcentrifuge tube.
10. Add 50 μL of 5% TFA in 50% ACN to the gel pieces and incubate at 37°C for 30 min (vortex or sonicate intermittently) to further extract peptide products. Spin down the gel slices and pool with the supernatant from step 9. Repeat this step once.
11. Dry the pooled supernatants in the SpeedVac. Do not use heat. Samples can be stored in this dry form until ready for desalting and MS data acquisition (see Note 11).

3.4. C_{18} ZipTip Treatment of Samples and Spotting on the MALDI Target Plate (see Note 12)

1. Reconstitute dried peptide samples by adding 10 μL of 0.1% TFA – vortex or sonicate to dissolve. Briefly centrifuge to bring fluid to bottom of the tubes and set aside.
2. Place *ca*. 1.0 mL each of 100% ACN, 50% ACN containing 0.1% TFA, and 0.1% TFA into cleaned microcentrifuge tubes. These are working solutions for ZipTip activation.
3. Prepare a C_{18} ZipTip by pipetting in 10 μL of 100% ACN and expelling it to waste, three times, then repeating with three times with 10 μL of 50% ACN containing 0.1% TFA, then finally treating three times with 10 μL of 0.1% TFA. Do not draw air into the resin bed between volumes.
4. Bind the peptides to the ZipTip resin by pipetting in and out the 10 μL peptide sample five times.
5. Wash the ZipTip by pipetting in 10 μL of fresh 0.1% TFA and discharging the solution to waste. Repeat this at least three times.
6. Elute peptides by pipetting in 10 μL of freshly prepared CHCA solution, then expelling into a pre-washed and labeled microcentrifuge tube. Re-pipet the sample volume in and out three cycles, being careful to not introduce air into the resin packing. Expel the last volume of solvent from the tip by raising the tip just above the solution in the tube and pushing out with air.

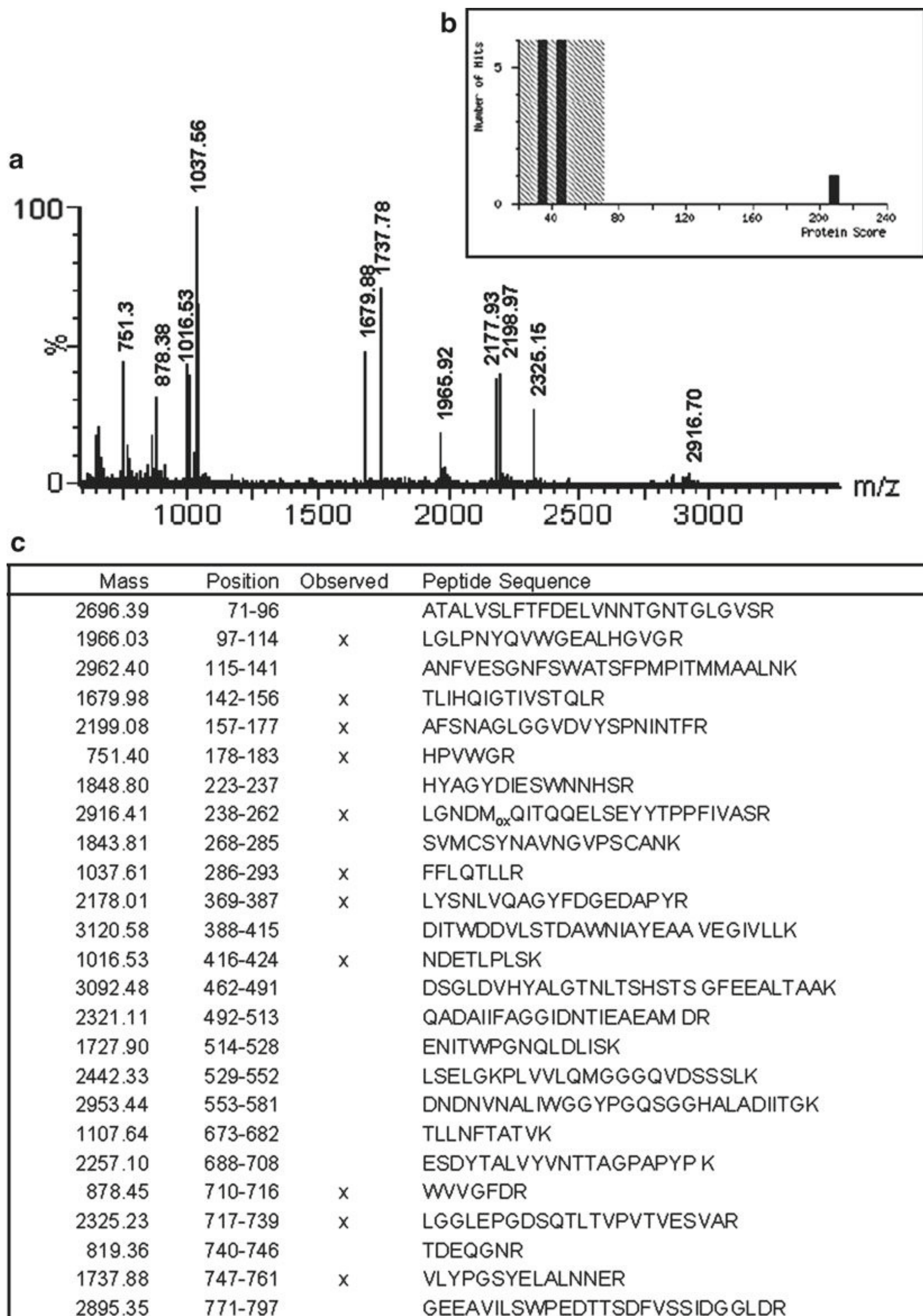
7. Pipet 1 μL of each target protein sample and control standards (gel blank and BSA or ADH digestions) onto wells of a MALDI target plate (see Note 13). Make sure you record on a log sheet the well positions of all samples and standards.
8. Pipet 1 μL each of a resolution check standard working solution (e.g., ACTH) and calibration standard working solution (e.g., sodiated PEG) onto wells of the target plate and finally 1 μL of GFPB solution onto calibration (a.k.a. lock mass) wells of the target plate.
9. Place target plate in a covered desiccator for 20 min at RT. Do not warm the plate during solvent evaporation. The MALDI target plate is now spotted and ready for loading into the MALDI-TOF MS instrument for data acquisition (see Note 14).

3.5. Identification of Recombinant Proteins with Mascot (see Note 15)

1. Go to http://www.matrixscience.com to open *Mascot* search engine. Click on *Peptide Mass Fingerprinting* to open the program for information entry (see Note 16).
2. Enter the all necessary information and search parameters. Provide *Name*, *Email*, and *Search Title*. Select: *Database(s)* (*SwissProt and/or NCBInr*), *Enzyme* (*Trypsin*), and *Allow up to* (1) missed cleavages. For *Taxonomy*, select appropriate classification to limit the taxonomic search to related groups, e.g., (*Viridiplantae*) for all green plants. For *Fixed modifications*, select (*carbamidomethyl*) for alkylated cysteine residues and for *Variable modification*, select (*Oxidation − M*) for variable oxidation of methionine residues. The field *Protein mass* is left blank and set *Peptide tol.* ±to (*50 ppm*), and finally select for *Mass values* (*M+H+*) and *Monoisotopic* ions.
3. Select *Browse* to activate window to select file path to the saved MS data file, launch by activating *Start Search* button.
4. Your search result will be displayed with a histogram indicating best "hits" (protein scores with significance $p<0.05$) (Fig. 2b).
5. If your protein is not indicated in the Protein Summary Report, manually inspect the MS spectrum using the list of peptide ions generated by in silico trypsin digestion of your recombinant protein (see Note 16).

4. Notes

1. Use powder-free nitrile gloves (rinsed with dH_2O) at all times for handling gels, tips, vials, etc. Use "low-protein-binding" polypropylene tubes that are cleaned of plasticizers, as well as new vials, disposable plasticware, and new razor blades. We recommend using high quality bottled water, such as a Spectro- or



Mass	Position	Observed	Peptide Sequence
2696.39	71-96		ATALVSLFTFDELVNNTGNTGLGVSR
1966.03	97-114	x	LGLPNYQVWGEALHGVGR
2962.40	115-141		ANFVESGNFSWATSFPMPITMMAALNK
1679.98	142-156	x	TLIHQIGTIVSTQLR
2199.08	157-177	x	AFSNAGLGGVDVYSPNINTFR
751.40	178-183	x	HPVWGR
1848.80	223-237		HYAGYDIESWNNHSR
2916.41	238-262	x	LGNDM$_{ox}$QITQQELSEYYTPPFIVASR
1843.81	268-285		SVMCSYNAVNGVPSCANK
1037.61	286-293	x	FFLQTLLR
2178.01	369-387	x	LYSNLVQAGYFDGEDAPYR
3120.58	388-415		DITWDDVLSTDAWNIAYEAA VEGIVLLK
1016.53	416-424	x	NDETLPLSK
3092.48	462-491		DSGLDVHYALGTNLTSHSTS GFEEALTAAK
2321.11	492-513		QADAIIFAGGIDNTIEAEAM DR
1727.90	514-528		ENITWPGNQLDLISK
2442.33	529-552		LSELGKPLVVLQMGGGQVDSSSLK
2953.44	553-581		DNDNVNALIWGGYPGQSGGHALADIITGK
1107.64	673-682		TLLNFTATVK
2257.10	688-708		ESDYTALVYVNTTAGPAPYP K
878.45	710-716	x	WVVGFDR
2325.23	717-739	x	LGGLEPGDSQTLTVPVTVESVAR
819.36	740-746		TDEQGNR
1737.88	747-761	x	VLYPGSYELALNNER
2895.35	771-797		GEEAVILSWPEDTTSDFVSSIDGGLDR

Fig. 2. The MALDI-TOF mass spectrum (“peptide mass fingerprint”) (**a**) from a trypsin digest of *Aspergillus nidulans* β-xylosidase, AN2359.2 (see Chapter 21). The inserted histogram (**b**) shows results from a protein database search with the MS spectrum peak list data using Mascot. The top score (209) matches the correct *A. nidulans* accession. Theoretical peptide ion list from the protein sequence (**c**) (range for 750–3,500 *m*/*z*) is used for direct manual inspection of MS data. Peptides observed in the spectrum are indicated. Sequence of coverage is commonly 30–60%.

LC-grade water for preparing MS reagent solutions. Laboratory water deionization systems (e.g., Nanopure or MilliQ) that are regularly sanitized may provide suitable in place of bottle ultrapure water. Always use fresh solutions, including electrophoresis solutions and reagent buffers. Minimize introduction of dust by storing materials in sealed dust-free containers. If available, use a laminar-flow transfer hood during all steps for processing samples following the electrophoresis stage. Clean staining containers and all other equipment that comes into contact with the gel or gel slices with laboratory detergent (e.g., Alconox), rinse thoroughly with dH_2O, and finally wipe with alcohol or 50% aqueous acetonitrile immediately prior to use.

2. In depth discussion and details for extraction, solubilization, and quantification of protein preparations are readily found in manuals and reviews (11–13). For concentration of dilute proteins (such as from culture medium), microcentrifuge ultrafiltration devices can also be used (e.g., Millipore Microcon-3).

3. Approximate protein determination is needed to calibrate sample protein loadings on SDS-PAGE gels. We use Coomassie Brilliant Blue G-250 dye ("Bradford's reagent") for routine protein determination. Prepared solutions are available from vendors such as Pierce Chemical or Bio-Rad.

4. If a large amount of protein is precipitated resulting in a thick pellet, vigorously vortex or sonicate it to break it up, then incubate at –20°C, centrifuge, and repeat pellet washing with cold acetone.

5. A greenish-yellow sample indicates presence of residual TCA. This may be corrected with 1 μL aliquots of 1 M Tris solution. If the sample does not then readily turn blue, you should repeat the precipitation step by adding 4 volumes of cold acetone, incubate, and centrifuge as before.

6. The assembly of electrophoresis cells and running gels are taken from the Invitrogen instruction guide (9). Be sure to rinse out the samples wells in the BisTris precast gel with 1× NuPAGE running buffer and to remove the sealing tape on the bottom. Clamp the gel into XCell SureLock Mini-Cell according, using the plastic buffer dam if you are using only one gel.

7. Generally, load a volume of protein samples containing between 2.5 and 25 μg of total protein, depending on anticipated relative purity of the target protein(s). Ideally, minimal bands (at least 0.5 μg) stained by Coomassie Brilliant Blue are suited for processing for MALDI-TOF MS analysis. Loading samples at multiple volumes (e.g., 5, 10, and 20 μL) allow for variance from estimated protein contents. Be sure to use at least one known protein such as BSA or ADH loaded in separate lanes. These are used as positive controls to assess the effectiveness of the trypsin

digestion, peptide recovery, and peptide ion signal intensity. Leave at least one lane empty to use for negative control gel pieces in Subheading 3.3, step 1.

8. Gel bands should be cut at *ca.* 0.5–0.8 mm wide, eliminating as much unstained polyacrylamide as possible. For heavy protein bands, cut gel piece from the upper edge of the band (multiple bands can be excised and treated from large bands). From empty gel lanes, cut similarly sized gel pieces for use in negative control digestions – these samples are used to assess for background peptide ions introduced by systemic contamination.

9. Plasticizer residues present in microcentrifuge tube will interfere with MALDI-TOF MS analysis by contributing polymer peaks and suppressing peptide ionization by affecting matrix crystallization. Microcentrifuge tubes must be prewashed thoroughly with a solvent solution such as 50% ACN in 0.1% TFA to remove these residues. Fill microcentrifuge tubes with solution for *ca.* 10 min and then rinse with an alcohol or ACN. Store washed dry microcentrifuge tubes in a sealed container.

10. Trypsin Gold (V5280) from Promega (10) is prepared by adding 100 μL of 50 mM acetic acid directly to the vial containing 100 μg trypsin. After dissolving completely, dispense 5 μL (5 μg) aliquots into microcentrifuge tube (0.5 mL, prewashed) and store at −20°C freezer. To prepare working trypsin solution, add 45 μL of 50 mM acetic acid to an aliquot, then dilute with 200 μL of 40 mM NH_4HCO_3 in 10% ACN. This provides a final working solution with concentration of 20 μg/mL, which is sufficient for performing for 10–12 digestions. Unused trypsin working solution can be stored at −20°C; discard after five freeze–thaw cycles. Trypsin digestion of proteins can be accelerated and peptide recovery improved with 0.1% w/v 1-*O*-*n*-octyl-β-d-glucopyranoside or the acid-labile surfactant RapiGest SF (0.05%) added to the working trypsin solution (14). These are particularly useful for hydrophobic proteins. Trypsin digestion is complete in 2 h, providing an alternative to overnight digestion.

11. Dried peptide samples are stable for short-term storage (refrigerated or in freezer), and you should submit your samples to a biomolecule mass spectrometry facility in this form. Contact the facility you will be sending your samples to for specific shipping instructions. To support effective analysis of a sample, send an image of the SDS-PAGE gel along with the samples and the expected amino acid sequence (best as a FASTA file). Include contiguous sequence(s) for any inserted epitope or other affinity tag (e.g., poly-His peptides). The facility should be able to do basic matching of peptide ions from the sequence with the MS spectrum and provide a summary of results. If you have direct local access to a MALDI-TOF MS instrument, proceed with ZipTip sample cleanup as described in Subheading 3.4.

12. This section describes final treatment to samples for analysis with a local MS instrument. C_{18} ZipTips (Millipore; 15) are micropipet tips packed with 0.6 μL of C_{18} reversed-phase resin. They are used to bind peptides from aqueous solution and wash them free of salts and contaminates introduced during gel digestion and extraction, resulting in improved spectral data from MALDI-TOF MS. ZipTip treatment is best done immediately before acquisition of MS data, because the CHCA used in the ZipTip elution solution will start to decompose once samples are spotted onto a target plate, resulting in degraded spectral data.

13. Stainless steel MALDI targets plates must match the instrument. Samples may be further diluted with matrix solution at 1:3 and 1:10 (v/v) before spotting, which sometimes improves signal intensities.

14. An experienced user must perform operation of the local mass spectrometer. Skill in instrument tuning and calibration is as critical as high quality sample preparation. MALDI-TOF MS spectrum peak list data files acquired from protein digests (i.e., the PMF) will be evaluated with integrated software for search and matching, or the data files are saved for export and/or later analysis. Be sure to obtain peak list data files for samples submitted to a core or commercial facility. How to use the Mascot search engine to identify a protein from MS spectrum peak list data file is described in the next section.

15. The high mass accuracy of current generation MALDI-TOF mass spectrometers enables confident identification of proteins from PMFs using a search engine such as Mascot (16). However, identification of an unknown protein requires that the sequence for it already exist in a protein database (e.g., NCBInr or SwissProt). A recombinant protein whose sequence that may not be publically disclosed can be identified simply from manual inspection of the MS spectrum for predicted peptide ions. There will be multiple peptide ions present that match masses calculated from a theoretical peptide list (Fig. 2). Software programs and related tools for protein identification and characterization are accessed through the EXPASy Proteomics Server (http://expasy.org). Links to various search engines are found there as well as prediction tools such as PeptideMass for in silico trypsin digestion.

16. To generate a list of theoretical peptide ions for manual inspection of an MS spectrum (Fig. 2) open PeptideMass at http://www.expasy.ch/tools/peptide-mass.html and paste in the sequence. The sequence of any signal peptide should be removed first using the SignalP 3.0 program found in the ExPASy Proteomic Tools. Select for peptide masses: cysteines treated with (*Iodoacetamide*) and check or activate "with methionine oxidized," "$[M+H]^+$," and "monoisotopic." Select an enzyme

(*Trypsin*) and set for display mass bigger than (*750*) and smaller than (*3,000*) Da. You can select to "sort by" peptide masses or in chronological (sequential) order in the protein. Click on *Perform* button to obtain the peak list.

Acknowledgments

This work was supported by funding from the DOE Mid-South/Southeast Biofuels Consortium (DE-FG36-08BO88036) and the NSF EPSCoR, RII: Arkansas ASSET Initiative Plant Protein Production Center (#EPS-0701890).

References

1. Shi, Q. and Jackowski, G. (1998) One-dimensional polyacrylamide gel electrophoresis. *IN*: Gel Electrophoresis of Proteins – A Practical Approach, third ed. (Hames, B. D., editor) Oxford Univ. Press, NY, pp 1–52.
2. Magi, B. and Liberatori, S. (2005) Immunoblotting techniques. *Methods Mol. Biol.* **295**, 224–254.
3. LeGendre, N. and Matsudaira, P. T. (1989) Purification of proteins and peptides by SDS-PAGE. *IN*: A Practical Guide to Protein and Peptide Purification for Microsequencing (Matsudaira, P. T., editor) Academic Press, NY, pp 52–69.
4. Rosenfeld, J., Capdevielle, J., Guillemot, J. C. and Ferrara, P. (1992) *In-gel* digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. *Anal. Biochem.* **203**, 173–179.
5. Aebersold, R. and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* **422**, 198–207.
6. Henzel, W. J. and Watanabe, C. (2003) Protein identification: the origins of peptide mass fingerprinting. *J. Am. Soc. Mass. Spectrom.* **14**, 931–942.
7. Savary, B. J., Prasanna V., Nuñez, A., and Cameron, R. G. (2010) Identification of thermolabile pectin methylesterases from sweet orange fruit by peptide mass fingerprinting. *J. Agric. Food Chem.* **58**, 12462–12468.
8. Peterson, G. L (1977) A simplification of the protein assay method of Lowry *et al.*, which is more generally applicable. *Anal. Biochem.* **83**, 346–356.
9. Invitrogen Life Technologies (2003) Instruction Manual. NuPAGE Technical Guide – General information and protocols for using the NuPAGE electrophoresis system. (http://tools.invitrogen.com/content/sfs/manuals/nupage_tech_man.pdf)
10. Promega (2006) Technical bulletin: Trypson Gold, mass spectrometry grade. (http://www.promega.com.cn/techserv/tbs/TB241-550/tb309.pdf)
11. Burgess, R. R., and Deutscher, M. P., editors (2009) Methods in Enzymology, Vol. 463, Guide to Protein Purification, 2nd Edition. Academic Press.
12. Wang, W. Tai, F., and Chen, S. (2008) Optimizing protein extraction from plant tissues for enhanced proteomics analysis. *J. Sep. Sci.* **31**, 2032–2039.
13. Michaud, D., and Asselin, A. (1995) Application to plant proteins of gel electrophoretic methods. *J. Chromatog. A* **698**, 263–279.
14. Waters (2007) Care and use manual, Rapigest SF surfactant. (http://wwwp1.waters.com/webassets/cms/support/docs/715000122.pdf)
15. Millipore (2007) User guide for reversed-phase ZipTip pipette tips for sample preparation. (http://www.millipore.com/userguides/tech1/pr02358)
16. Perkins, D. N., Pappin, D. J., Creasy, D. M. and Cottrell, J. S. (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–3567.

Chapter 3

A Matter of Packaging: Influence of Nucleosome Positioning on Heterologous Gene Expression

María de la Cruz Muñoz-Centeno, Gonzalo Millán-Zambrano, and Sebastián Chávez

Abstract

The organization of DNA into the various levels of chromatin compaction is the main obstacle that restricts the access of transcriptional machinery to genes. Genome-wide chromatin analyses have shown that there are common chromatin organization patterns for most genes but have also revealed important differences in nucleosome positioning throughout the genome. Such chromatin heterogeneity is one of the reasons why recombinant gene expression is highly dependent on integration sites. Different solutions have been tested for this problem, including artificial targeting of chromatin-modifying factors or the addition of DNA elements, which efficiently counteract the influence of the chromatin environment.

An influence of the chromatin configuration of the recombinant gene itself on its transcriptional behavior has also been established. This view is especially important for heterologous genes since the general parameters of chromatin organization change from one species to another. The chromatin organization of bacterial DNA proves particularly dramatic when introduced into eukaryotes. The nucleosome positioning of recombinant genes is the result of the interaction between the machinery of the hosting cell and the sequences of both the recombinant genes and the promoter regions. We discuss the key aspects of this phenomenon from the heterologous gene expression perspective.

Key words: Chromatin reassembly, Cryptic transcription, Recombinant gene expression, Insulator, Nucleosome positioning, +1 Nucleosome

1. Introduction

Gene expression is a complex phenomenon involving a broad set of integrated steps that helps decipher the information encoded in the DNA. The natural configuration of DNA in living cells is chromatin, a highly ordered structure in which the nucleic acid packages itself when interacting with a number of structural proteins.

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_3, © Springer Science+Business Media, LLC 2012

The main protein components of chromatin are histones H2A, H2B, H3, and H4. These four histones are the core of the nucleosome, the basic element contributing to folding DNA into chromatin.

Nucleosomes are the main obstacle restricting the access of the transcriptional machinery to genes. Both in vitro and in vivo experiments have clearly documented the repressive role of nucleosomes in transcription. Nucleosomes can hinder the access of activator proteins to their binding sites within enhancers and upstream activating sequences, they preclude the assembly of preinitiation complexes on the core promoters, and play an inhibitory role during RNA polymerase II-dependent transcription elongation (reviewed in ref. 1).

In spite of this general negative function of chromatin in transcription, positive contributions of nucleosomes to gene expression have been described (2–7). In these cases, the precise location of these regulatory nucleosomes is critical. Translational positioning (high preference of a 146-bp sequence to be wrapped around the histone octamer in comparison to the adjacent linker regions) can be required for a nucleosome to bring two distant binding sites together (2, 4). In other cases, a fix rotational setting (the orientation of the sequence in relationship to the surface of the histone octamer due to the 10-bp periodicity of the DNA helix) is needed for the stable binding of a transcription factor to the positioned nucleosome (5) or to allow the optimal synergism between the factors regulating a transcriptional promoter (8).

Genome-wide chromatin analyses have shown that there are common chromatin organization patterns for most genes, but have also revealed important differences in nucleosome positioning throughout the genome (9, 10). Genome-wide analyses have also described the distribution of covalent histone modifications, such as H3K4, H3K9, and H3K36 methylation, H4K16 acetylation, or H2BK123 ubiquitination, which directly relate to the euchromatic or heterochromatic state of a chromosomal region and, consequently, to its transcriptional activity (11).

These differences are particularly marked in mammals where specialized genome domains are numerous and relate to the epigenetic processes that influence cell fate. This chromatin heterogeneity of the mammalian genome is one of the reasons why the expression of transgenes is highly dependent on the integration sites, and is one of the main fields of research to improve mammalian cell factories (12, 13). Different solutions have been tested for this problem, including the creation of hot spots for integration, the artificial targeting of chromatin-modifying factors to the transgene, or the addition of DNA elements which efficiently counteract the influence of the chromatin environment. The first part of this chapter reviews the experimental results obtained by following the last cited strategy.

The second part analyzes some cases in which an influence of the chromatin configuration of the recombinant gene itself on its transcriptional behavior has been established. This perspective is especially important for genes of a heterologous origin since the general parameters of chromatin organization change from one species to another (14, 15), or even between different cell-types from the same organism (16).

2. Influence of Chromatin Environments on the Expression of Recombinant Genes

Integration of transgenes in the recipient genome normally occurs at random. Every clone exhibits a different integration site, while the expression levels of the transgene and its transcriptional properties are highly influenced by the locus of integration (17). In contrast, endogenous genes in their natural context can be transcribed from their promoters and be regulated by their enhancers without any interference of the regulatory elements (enhancers and silencers) of their neighboring loci. According to this view, natural genes are isolated from their chromosomal context. The elements responsible for this isolation are called insulators which are separable from the other elements that drive transcriptional regulation (18). After the functional description of transposable elements in Drosophila harboring insulating functions (19), the first insulator from vertebrates was found in the chicken β-globin locus (20), and many others have since been described (reviewed in ref. 21).

Insulators, also called boundary elements, show at least one of the following functions: enhancer-blocking and silencing barriers. When an enhancer-blocking insulator is located between the target promoter and an enhancer, the promoter does not respond to the regulatory inputs mediated by the enhancer, although it maintains its capacity to enhance the transcriptional activity of a different promoter located on the other side of the insulator (Fig. 1).

Although the mechanism of action of enhancer-blocking insulators is not completely clear, two main models have been proposed. The first is based on the formation of chromosomal loops, which would prevent the interaction of the target promoter with an enhancer located outside the loop. In the second model, enhancer-blocking insulators act as decoy promoters, which are able to directly interact with the enhancer and, therefore, to impede the interaction between the enhancer and the target promoter. The binding of different transcription factors and the transcriptional complex, including RNA polymerase II, to enhancer-blocking elements, supports the second model (22). The first model is supported by the enhancer-blocking insulators' ability to interact with each other (23). Based on the genome-wide mapping of factors

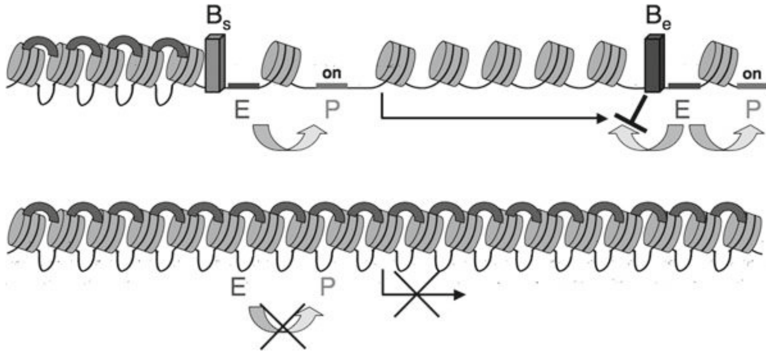


Fig. 1. Boundary elements (B) insulate recombinant genes from the genomic environment. Enhancer blocking insulators (B_e) prevent the influence of external enhancers (E) on the internal promoter (P); Silencing barriers (B_s) counteract surrounding heterochromatin. The recombinant gene is silenced in the absence of silencing barriers.

involved in enhancer-blocking insulation, a new dynamic model has been recently proposed. According to this, an insulator–insulator loop, which allows a productive enhancer–promoter interaction, would switch into an enhancer-blocking situation by favoring insulator–enhancer contacts (24).

Silencing barriers are the other type of insulator elements. They prevent silencing heterochromatin spreading across an active locus (Fig. 1). Silencing barriers, therefore, mark the borders between open and condensed chromatin. The mechanisms by which this kind of insulation works remain unclear, but several experimental results suggest that chromatin modification is the key molecular aspect of these elements. The well-known HS4 element, which insulates the chicken β-globin locus from neighboring heterochromatin, is not only bound by factors that recruit histone acetylases and methylases, but impedes the spread of DNA methylation (25, 26). These molecular features resemble those of CpG island promoters, and they allowed Raab and Kamakaka (24) to propose how silencing barriers and promoters also share common mechanisms. In fact, the spread of heterochromatin in *Saccharomyces cerevisiae* yeast is efficiently restricted merely by active regular promoters driven by either RNA polymerase II (*TEF1* and *TEF2*) or RNA polymerase III (HMR $tRNA^{Thr}$) (reviewed in ref. 27). In fact, the recruitment of the activation domain of the powerful transcriptional regulator Rap1 is sufficient for barrier activity in yeast (28).

According to this perspective, transcriptional isolation is the result of combining functions that (1) keep chromatin free of silencing marks and (2) counteract the action of the active enhancers present in the neighboring loci. Do all boundary elements exhibit these two functions? Several reported insulators are simultaneously barriers and enhancer-blockers: the aforementioned HS4 from the chicken ß-globin, the boundary element associated with the 5 -end of the mouse tyrosinase locus (21), the SINE2B element

present in the murine growth hormone locus (29) and the *Drosophila SF1* element (30).

Other elements that are unable to act as a barrier confer enhancer-blocking activity. For instance, the CCCTC-binding factor (CTCF) operates as an enhancer-blocking factor in HS4, but the barrier function of this insulator is not mediated by CTCF (31). However, the activity of CTCF is also required for the D4Z4 microsatellite repeat's antisilencing activity (32), which suggests that the contribution of CTCF to barrier insulation depends on the genomic context.

The insulators flanking the mouse whey acidic protein (WAP) gene operate in a CTCF-independent manner by not only acting as enhancer blockers, but also avoiding the interference of the mammary gland-specific regulation of this gene by the two neighboring ubiquitously expressed loci (33). An essential feature of these elements is the presence of matrix-attachment regions (MARs) which induce a chromatin loop organization that efficiently insulates the WAP gene (34).

The common objective of many of the experimental attempts performed so far has been to identify the sequences capable of ensuring a sustained transcription of the transgene independently of the integration site. The simplest approach is to use an artificial chromosome containing a euchromatic locus as the vector backbone. The expression of porcine WAP proved efficient and strictly specific to the mammary gland of lactating transgenic mice when a long genomic fragment, containing the transgene flanked by 145 and 5 kb of the 5′ and 3′ sequences, respectively, was expressed from a bacterial artificial chromosome (BAC) (35). BACs have also been successfully used to improve recombinant protein production. For instance, the constant region of human IgG1 was stably produced over time when the transgene was expressed in human HEK 293 cells from a BAC containing the euchromatic murine *Rosa26* locus as the vector backbone, and protein production increased by a factor of 10 when compared with a conventional vector (36).

In general, YACs (yeast artificial chromosomes) or BACs based on large genomic sequences are the best solution for optimizing the expression of a transgene (37). However, huge efforts have been made to construct universal vectors that are capable of ensuring the accurate expression of any recombinant gene in all possible genomic environments (12, 37). For instance, MARs have been incorporated into viral and nonviral vectors to increase the transgene expression in recombinant protein production and experimental approaches to gene therapy (38, 39); nonetheless, the effects of MARs on the recombinant gene expression are not copy number-dependent (38).

The accumulated knowledge on insulators has been also applied to the expression of heterologous recombinant constructs in a

position-independent manner. The use of the HS4 chicken β-globin insulator has improved the performance of integrative vectors for gene therapy. For instance, the addition of this insulator element to retroviral and lentiviral vectors significantly increases the likelihood of the expression of γ- and β-globins after mouse bone marrow transduction and transplantation (40, 41). In contrast with the improved efficiency observed in gene-therapy vectors, the addition of insulator elements to transgene constructs does not guarantee the full level of expression in all the transgenic animal lines produced (reviewed in ref. 21), and is unable to confer a copy number-dependent expression of transgenes (42).

The HS4 insulator also improves the recombinant protein production mediated by the nonintegrative lytic virus. This insulator enhances the production of green fluorescent protein (GFP), alkaline phosphatase, and luciferase when inserted downstream to the polyhedrin promoter-directed gene expression cassette of the AcMNPV baculovirus (43). The effect of the HS4 insulator on the expression of the target genes was mediated by its influence on the neighboring chromatin since it is orientation-dependent, sensitive to inhibitors of histone acetyltransferases and increases the sensitivity of the surrounding DNA to DNAse I (43). This is a good example of the function of an insulator from vertebrates in insect cells. However, the function of heterologous insulators among metazoa is not always optimal (44).

Other elements that are able to protect transgenes from silencing are derived from CpG islands or promoter regions of ubiquitously expressed housekeeping genes. One of these so-called ubiquitous chromatin-opening elements (UCOEs) has been successfully incorporated into mammalian expression vectors to increase the average expression of the transgene from the hCMV promoter/enhancer by up to 16-fold (45).

Nondirected screening technologies have also been successfully used to identify stabilizing and antirepressor (STAR) elements, which confer a stable position-independent expression to a transgene and allow the efficient production of recombinant proteins in human cells, in a copy number-dependent manner (46). STARs can be incorporated into expression vectors in combination with other antisilencing elements to enhance the recombinant protein expression in mammalian cells. The targeting of p300 histone acetyl transferase to strong viral or cellular promoters is sufficient to increase the expression of the target gene; when this strategy was applied to recombinant genes flanked by STARs, a significantly longer stability of the enhanced expression was obtained (47). No enhancement in stable protein expression was achieved by targeting the chromatin-remodeling factor Brahma to the same vectors and target genes; this indicates that histone acetylation is a better tool for counteracting silencing chromatin environments in mammalian cells.

3. Chromatin Interference from Inside: Nucleosome Positioning of Recombinant Genes

As we have shown, the genomic environment strongly affects the chromatin organization of recombinant genes, and several strategies have been put forward to avoid the negative consequences of this influence on their expression; however, chromatin interferences can also originate from inside. The chromatin organization of the recombinant gene is the result of the combination of several components, some of which are inherent to either the sequence of the recombinant gene or the interaction between its sequence and the hosting cell.

One question that firstly arises from this perspective is: just how strong is the influence of the sequence of a recombinant gene on its chromatin organization? This is the particular case of a broader question: what controls nucleosome positions? This has been a hot field of discussion over the last 20 years. First of all, nucleosome *positioning*, defined as the probability of a nucleosome being located in a given position, should be distinguished from nucleosome *occupancy*, the probability of a base pair being covered by nucleosomes irrespectively of their exact positions (see ref. 48 for a detailed discussion of these definitions).

A large number of in vitro reconstitution studies have shown that DNA sequences display different intrinsic tendencies to be organized in chromatin, and that the distributions of nucleosomes along some DNA fragments are not random and resemble in vivo distribution (49–51). According to these results, nucleosome positioning would be determined in *cis*. Correlation between good in vitro positioning sequences and DNA bendability has been well established using models in which positioning in vivo would be mainly determined by the sequence of every positioned nucleosome (52, 53). In other cases, however, there was no clear correspondence between the in vitro and the in vivo results. There are two possible explanations: either *trans* factors were involved which were not present in the in vitro experiments, or the *cis* elements controlling the positioning in vivo were distant and, therefore, were not included in the fragment analyzed in vitro. Both explanations are compatible with the statistical positioning concept (54). According to this model, barriers are located at specific sites along the genome which prevent nucleosome binding; in this scenario, positioning would be the consequence of nucleosomes being packed between the barriers. The barriers would be either intrinsically reluctant sequences to nucleosome positioning or factors capable of displacing nucleosomes by binding specific sequences with a very high affinity.

In the last 5 years, genome-wide studies have produced complete nucleosome positioning maps in yeast, worms, *Drosophila*, and human cells (reviewed in ref. 55). The picture that emerges

from these studies is a genome in which very extensive regions are covered by positioned nucleosomes and where transcriptional promoters are key elements in the organization of genomic chromatin. In both yeast and metazoa, 160–180-bp nucleosome-free regions (NFRs) precede the transcriptional start sites. The sequences of these NFRs are rich in polyA/T, a composition that in vitro studies have demonstrated to be somewhat incompatible with nucleosome binding. The genomic picture also displays that the two nucleosomes flanking these NFRs (+1 and −1) are highly positioned, and that the degree of positioning is also high for the 5′ half of the transcribed region, although it decreases progressively toward the midpart of the gene. In addition, nucleosome reconstitution experiments have been performed with total genomic DNA and also analyzed by massive DNA sequencing to allow genome-wide comparisons between in vitro and in vivo nucleosomes (56, 57). The general conclusion drawn from this huge amount of information is that DNA sequence plays an important role in establishing nucleosome positioning in the promoter regions, mainly through NFRs. However, the direct influence of the rest of the genomic sequence on nucleosome positioning is poor, with the exception of promoter-proximal and enhancer sequences (57, 58).

This conclusion supports a new synthetic model that reconciles the opposite views of DNA-directed and statistical positioning. The equilibrium model by Segal and Widom (48) proposes that nucleosome positioning is a dynamic phenomenon in which the probability of chromatin configuration would be the result of integrating the concentration and affinities for the DNA of nucleosomes and sequence-specific binding factors. Since promoters and transcriptional enhancers are the richest sites of those factors binding across the genome, their sequences are key modulators of chromatin structure and have been shaped during evolution to deal with nucleosome positioning (15, 59).

According to this knowledge, the potential positioning-related problems of recombinant genes may originate from undesired interactions between the sequences present in the construct and *trans*-acting factors of the hosting cell. These spurious interactions may alter the distribution of positioned nucleosomes, either directly by competing with nucleosomes on their binding sites or indirectly by recruiting chromatin remodelers, histone-modifying complexes, or DNA methylases to the transgenes (48). Other potential problems may arise when the recombinant construct includes sequences that can act as NFRs or strong nucleosome positioners.

All these problems can be minimized if the recombinant construct is the chimeric combination of a homologous promoter and a heterologous gene body. If we choose a promoter, together with its regulatory sequences, from the genome of the hosting species, we prevent most of these undesired events. However, the body of the recombinant gene can still bring sequence information which

could either be recognized by the *trans*-acting factors of the hosting cell or have extremely low or high affinities for nucleosomes (NFRs or positioning sequences).

The probability of this kind of event increases with the evolutionary distance between the hosting species and the transgene, and is maximal when we introduce bacterial DNA into a eukaryotic host, in which case, nucleosome positioning simply does not take place (Fig. 2a). Experiments performed in *S. cerevisiae* using the strongly positioned and well-characterized *GAL1-10* promoter have demonstrated that the *E. coli lacZ* gene is quite unable to adopt a positioned nucleosomal distribution, whereas similar chimeric constructs between the *GAL1* promoter and other eukaryotic transgenes exhibit positioned patterns (60, 61). This phenomenon may explain why *lacZ* reporters show strong variegation when they are included in mammalian transgenic constructs, whereas similar constructs lacking bacterial DNA behave much more uniformly (62, 63). These observations fully agree with the results of in vitro reconstitution experiments, where the nucleosomes assembled by salt dialysis were almost ten times more likely to form on yeast than *E. coli* DNA (57). These results demonstrate that eukaryotic DNA has evolved to favor nucleosome formation. They also partially contradict the statistical positioning model because not every DNA can be positioned when located adjacent to a barrier.

One likely consequence of its reluctance to position nucleosomes stably is that the bacterial DNA in eukaryotes promotes the transcription from cryptic promoters (Fig. 2a). The multiple transcripts initiating in the bacterial backbone of conventional yeast vectors have been mapped (64), and the DNA motive present in *E. coli* pBR322 or pUC19 serves as upstream activating sequences in *S. cerevisiae* (65, 66).

In light of these phenomena, it is rather surprising that the expression of bacterial transgenes is still possible in eukaryotes. We may wonder whether the eukaryotic transcriptional machinery somehow counteracts any intrinsic chromatin alterations induced by the sequence of recombinant genes. The answer to this question stems from the genetic and molecular analyses carried out by Winston and coworkers. They described how a set of co-transcriptionally recruited factors is responsible for the reassembly of nucleosomes on the passage of elongating RNA pol II (67). The list of factors involved in preventing cryptic transcription from inside transcribed genes includes several histone chaperons and general elongation factors (68). The malfunction of these factors dramatically alters chromatin through the action of RNA pol II, originates an accumulation of free histones (69), and leads to the activation of cryptic promoters sequences which are usually present in the body of eukaryotic genes (Fig. 2b).

Mutations impairing some of these factors, such as Spt6 or the FACT complex, disrupt the nucleosome positioning of recombinant

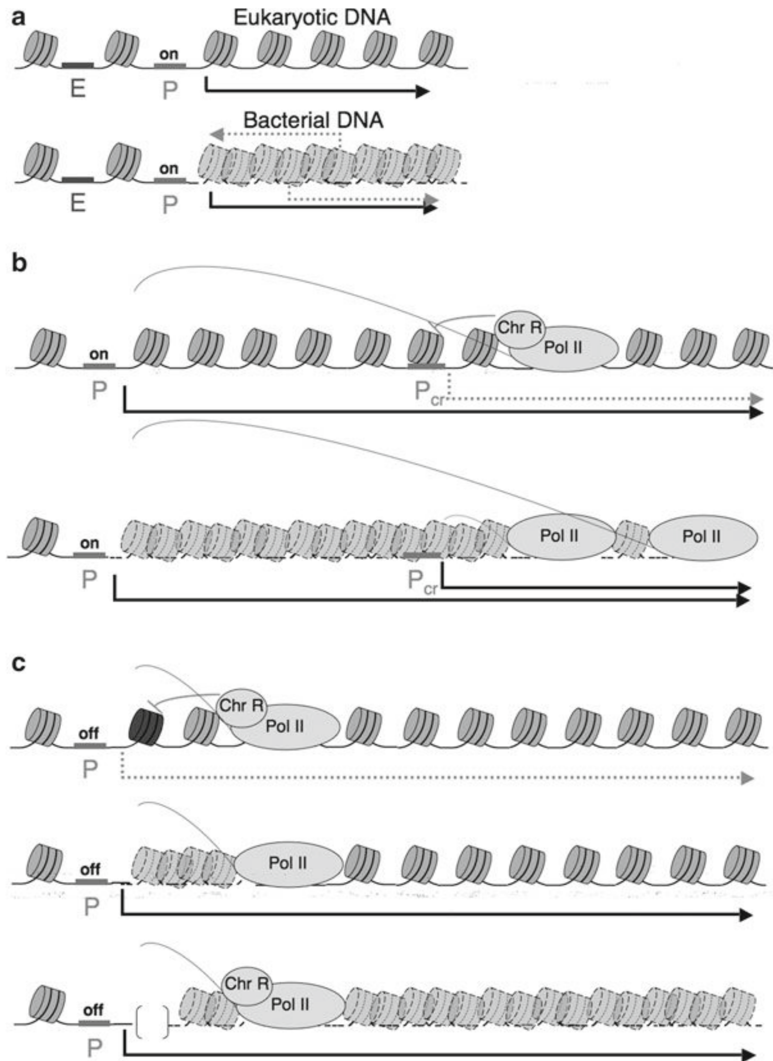


Fig. 2. The chromatin influence on the expression of recombinant genes also comes from inside. (**a**) When expressed in eukaryotes, bacterial DNA never displays positioned nucleosomes, but favors the appearance of cryptic transcripts. (**b**) Chromatin reassembly factors (Chr R) act cotranscriptionally with elongating RNA polymerase II (Pol II) to ensure nucleosome rebuilding after the passage of the polymerase. In the absence of chromatin reassembly factors, chromatin organization is impaired and cryptic transcripts arise from the body of the gene. (**c**) Chromatin reassembly factors are required for repressing the basal transcription of recombinant genes; in their absence, basal transcription increases. The presence of a positioned +1 nucleosome (*dark*) is required for the effective repression of the recombinant gene.

genes in yeast under nonactivating conditions. A chimeric construct formed by the yeast Ty1 promoter and the 5′ end of the HIV transcribed region shows a strong delocalization of the nucleosomes in the mutant strains under conditions that do not induce the Ty1 promoter (70). This chromatin distortion is accompanied by a

threefold increase in the basal expression of the recombinant gene, which allows a predicted function of the chromatin reassembly factors in HIV latency (70). These results indicate that chromatin reassembly machinery plays a role in the control of basal transcription by preserving nucleosome positioning (Fig. 2c). One intriguing finding is the deletion of the 5′ end of the recombinant gene, corresponding to the sequences holding the +1 nucleosome of HIV, which causes a similar phenomenon (increased basal expression and absence of nucleosome positioning) in the wild type (70).

According to these results, the sequences that are able to stabilize the +1 nucleosome would be essential for the function of the chromatin reassembly factors in controlling the basal expression (Fig. 2c). In a different work, our team found a second functional interaction between the FACT complex and the nucleosome positioning sequences located in the 5′ end of recombinant genes which, on this occasion, were driven by an activated yeast promoter (61). The functional importance of the positioned +1 nucleosomes explains why it has been so evolutionary preserved (15) and why their sequences usually direct nucleosome positioning (55) either directly or in collaboration with "an early step in the transcription process," as previously proposed (57).

Controlling the expression "noise" of recombinant genes due to *inside* chromatin distortion is, therefore, facilitated by chromatin reassembly machinery, which operates co-transcriptionally. Nevertheless, it is convenient to add an optimal sequence in the 5′ end of the transcribed region to favor the positioning of the +1 nucleosome. Whenever possible, the best solution would be to express our favorite recombinant gene under the control of an endogenous promoter, including its +1 nucleosome sequence.

Acknowledgments

This work has been supported by the Spanish Ministry of Education and Science (grant BFU2007-67575-C03-02/BMC), by the Andalusian Government (grants P07-CVI02623 and P08-CVI03508) and by the European Union (FEDER). We thank Helen Warburton for English corrections.

References

1. Li, B., Carey, M., and Workman, J. L. (2007) The role of chromatin during transcription. *Cell* **128**, 707–719.
2. Schild, C., Claret, F. X., Wahli, W., et al., (1993) A nucleosome-dependent static loop potentiates estrogen-regulated transcription from the *Xenopus vitellogenin* B1 promoter in vitro. *EMBO J* **12**, 423–433.
3. Chavez, S., and Beato, M. (1997) Nucleosome-mediated synergism between transcription factors on the mouse mammary tumor virus promoter. *Proc Natl Acad Sci USA* **94**, 2885–2890.
4. Stunkel, W., Kober, I., and Seifart, K. H. (1997) A nucleosome positioned in the distal promoter region activates transcription of the

human U6 gene. *Mol Cell Biol* **17**, 4397–4405.

5. Cirillo, L. A., and Zaret, K. S. (1999) An early developmental transcription factor complex that is more stable on nucleosome core particles than on free DNA. *Mol Cell* **4**, 961–969.
6. Wyrick, J. J., Holstege, F. C., Jennings, E. G., et al., (1999) Chromosomal landscape of nucleosome-dependent gene expression and silencing in yeast. *Nature* **402**, 418–421.
7. Koop, R., Di Croce, L., and Beato, M. (2003) Histone H1 enhances synergistic activation of the MMTV promoter in chromatin. *EMBO J* **22**, 588–599.
8. Chavez, S., Candau, R., Truss, M., et al., (1995) Constitutive repression and nuclear factor I-dependent hormone activation of the mouse mammary tumor virus promoter in *Saccharomyces cerevisiae*. *Mol Cell Biol* **15**, 6987–6998.
9. Rando, O. J., and Chang, H. Y. (2009) Genome-wide views of chromatin structure. *Annu Rev Biochem* **78**, 245–271.
10. Jiang, C., and Pugh, B. F. (2009) A compiled and systematic reference map of nucleosome positions across the *Saccharomyces cerevisiae* genome. *Genome Biol* **10**, R109.
11. Campos, E. I., and Reinberg, D. (2009) Histones: Annotating chromatin. *Annu Rev Genet* **43**, 559–599.
12. Kwaks, T. H., and Otte, A. P. (2006) Employing epigenetics to augment the expression of therapeutic proteins in mammalian cells. *Trends Biotechnol* **24**, 137–142.
13. Barnes, L. M., and Dickson, A. J. (2006) Mammalian cell factories for efficient and stable protein expression. *Curr Opin Biotechnol* **17**, 381–386.
14. Bernardi, F., Zatchej, M., and Thoma, F. (1992) Species specific protein–DNA interactions may determine the chromatin units of genes in *S. cerevisiae* and in *S. pombe*. *EMBO J* **11**, 1177–1185.
15. Tsankov, A. M., Thompson, D. A., Socha, A., et al., The role of nucleosome positioning in the evolution of gene regulation. *PLoS Biol* **8**, e1000414.
16. Prunell, A., and Kornberg, R. D. (1982) Variable center to center distance of nucleosomes in chromatin. *J Mol Biol* **154**, 515–523.
17. Wilson, C., Bellen, H. J., and Gehring, W. J. (1990) Position effects on eukaryotic gene expression. *Annu Rev Cell Biol* **6**, 679–714.
18. Ortiz, B. D., Cado, D., Chen, V., et al., (1997) Adjacent DNA elements dominantly restrict the ubiquitous activity of a novel chromatin-opening region to specific tissues. *EMBO J* **16**, 5037–5045.
19. Harrison, D. A., Geyer, P. K., Spana, C., et al., (1989) The gypsy retrotransposon of *Drosophila melanogaster*: Mechanisms of mutagenesis and interaction with the suppressor of Hairy-wing locus. *Dev Genet* **10**, 239–248.
20. Chung, J. H., Whiteley, M., and Felsenfeld, G. (1993) A 5′ element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. *Cell* **74**, 505–514.
21. Molto, E., Fernandez, A., and Montoliu, L. (2009) Boundaries in vertebrate genomes: different solutions to adequately insulate gene expression domains. *Brief Funct Genomic Proteomic* **8**, 283–296.
22. Chopra, V. S., Cande, J., Hong, J. W., et al., (2009) Stalled Hox promoters as chromosomal boundaries. *Genes Dev* **23**, 1505–1509.
23. Gerasimova, T. I., and Corces, V. G. (2001) Chromatin insulators and boundaries: Effects on transcription and nuclear organization. *Annu Rev Genet* **35**, 193–208.
24. Raab, J. R., and Kamakaka, R. T. (2010) Insulators and promoters: Closer than we think. *Nat Rev Genet* **11**, 439–446.
25. Huang, S., Li, X., Yusufzai, T. M., et al., (2007) USF1 recruits histone modification complexes and is critical for maintenance of a chromatin barrier. *Mol Cell Biol* **27**, 7991–8002.
26. Dickson, J., Gowher, H., Strogantsev, R., et al., (2010) VEZF1 elements mediate protection from DNA methylation. *PLoS Genet* **6**, e1000804.
27. Bi, X., and Broach, J. R. (2001) Chromosomal boundaries in *S. cerevisiae*. *Curr Opin Genet Dev* **11**, 199–204.
28. Yu, Q., Qiu, R., Foland, T. B., et al., (2003) *Rap1p* and other transcriptional regulators can function in defining distinct domains of gene expression. *Nucleic Acids Res* **31**, 1224–1233.
29. Lunyak, V. V., Prefontaine, G. G., Nunez, E., et al., (2007) Developmentally regulated activation of a SINE B2 repeat as a domain boundary in organogenesis. *Science* **317**, 248–251.
30. Majumder, P., Roy, S., Belozerov, V. E., et al., (2009) Diverse transcription influences can be insulated by the Drosophila SF1 chromatin boundary. *Nucleic Acids Res* **37**, 4227–4233.
31. Recillas-Targa, F., Pikaart, M. J., Burgess-Beusse, B., et al., (2002) Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities. *Proc Natl Acad Sci USA* **99**, 6883–6888.
32. Ottaviani, A., Rival-Gervier, S., Boussouar, A., et al., (2009) The D4Z4 macrosatellite repeat acts as a CTCF and A-type lamins-dependent insulator in facio-scapulo-humeral dystrophy. *PLoS Genet* **5**, e1000394.

33. Millot, B., Montoliu, L., Fontaine, M. L., et al., (2003) Hormone-induced modifications of the chromatin structure surrounding upstream regulatory regions conserved between the mouse and rabbit whey acidic protein genes. *Biochem J* **372**, 41–52.
34. Montazer-Torbati, M. B., Hue-Beauvais, C., Droineau, S., et al., (2008) Epigenetic modifications and chromatin loop organization explain the different expression profiles of the Tbrg4, WAP and Ramp3 genes. *Exp Cell Res* **314**, 975–987.
35. Rival-Gervier, S., Viglietta, C., Maeder, C., et al., (2002) Position-independent and tissue-specific expression of porcine whey acidic protein gene from a bacterial artificial chromosome in transgenic mice. *Mol Reprod Dev* **63**, 161–167.
36. Blaas, L., Musteanu, M., Eferl, R., et al., (2009) Bacterial artificial chromosomes improve recombinant protein production in mammalian cells. *BMC Biotechnol* **9**, 3.
37. Montoliu, L., Roy, R., Regales, L., et al., (2009) Design of vectors for transgene expression: The use of genomic comparative approaches. *Comp Immunol Microbiol Infect Dis* **32**, 81–90.
38. Kim, J. M., Kim, J. S., Park, D. H., et al., (2004) Improved recombinant gene expression in CHO cells using matrix attachment regions. *J Biotechnol* **107**, 95–105.
39. Harraghy, N., Gaussin, A., and Mermod, N. (2008) Sustained transgene expression using MAR elements. *Curr Gene Ther* **8**, 353–366.
40. Emery, D. W., Yannaki, E., Tubb, J., et al., (2002) Development of virus vectors for gene therapy of beta chain hemoglobinopathies: flanking with a chromatin insulator reduces gamma-globin gene silencing in vivo. *Blood* **100**, 2012–2019.
41. Malik, P., Arumugam, P. I., Yee, J. K., et al., (2005) Successful correction of the human Cooley's anemia beta-thalassemia major phenotype using a lentiviral vector flanked by the chicken hypersensitive site 4 chromatin insulator. *Ann N Y Acad Sci* **1054**, 238–249.
42. Truffinet, V., Guglielmi, L., Cogne, M., et al., (2005) The chicken beta-globin HS4 insulator is not a silver bullet to obtain copy-number dependent expression of transgenes in stable B cell transfectants. *Immunol Lett* **96**, 303–304.
43. Wang, X., Li, L., Ding, S., et al., (2009) Chicken HS4 insulator significantly improves baculovirus-mediated foreign gene expression in insect cells by modifying the structure of neighbouring chromatin in virus minichromosome. *J Biotechnol* **142**, 193–199.
44. Takada, T., Iida, K., Akasaka, K., et al., (2000) Evaluation of heterologous insulator function with regard to chromosomal position effect in the mouse blastocyst and fetus. *Mol Reprod Dev* **57**, 232–237.
45. Williams, S., Mustoe, T., Mulcahy, T., et al., (2005) CpG-island fragments from the HNRPA2B1/CBX3 genomic locus reduce silencing and enhance transgene expression from the hCMV promoter/enhancer in mammalian cells. *BMC Biotechnol* **5**, 17.
46. Kwaks, T. H., Barnett, P., Hemrika, W., et al., (2003) Identification of anti-repressor elements that confer high and stable protein production in mammalian cells. *Nat Biotechnol* **21**, 553–558.
47. Kwaks, T. H., Sewalt, R. G., van Blokland, R., et al., (2005) Targeting of a histone acetyltransferase domain to a promoter enhances protein expression levels in mammalian cells. *J Biotechnol* **115**, 35–46.
48. Segal, E., and Widom, J. (2009) What controls nucleosome positions? *Trends Genet* **25**, 335–343.
49. Pina, B., Barettino, D., Truss, M., et al., (1990) Structural features of a regulatory nucleosome. *J Mol Biol* **216**, 975–990.
50. Lowary, P. T., and Widom, J. (1998) New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J Mol Biol* **276**, 19–42.
51. Gencheva, M., Boa, S., Fraser, R., et al., (2006) In vitro and in vivo nucleosome positioning on the ovine beta-lactoglobulin gene are related. *J Mol Biol* **361**, 216–230.
52. Trifonov, E. N. (1980) Sequence-dependent deformational anisotropy of chromatin DNA. *Nucleic Acids Res* **8**, 4041–4053.
53. Thastrom, A., Lowary, P. T., Widlund, H. R., et al., (1999) Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. *J Mol Biol* **288**, 213–229.
54. Kornberg, R. D., and Stryer, L. (1988) Statistical distributions of nucleosomes: Nonrandom locations by a stochastic mechanism. *Nucleic Acids Res* **16**, 6677–6690.
55. Radman-Livaja, M., and Rando, O. J. (2010) Nucleosome positioning: How is it established, and why does it matter? *Dev Biol* **339**, 258–266.
56. Kaplan, N., Moore, I. K., Fondufe-Mittendorf, Y., et al., (2009) The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458**, 362–366.
57. Zhang, Y., Moqtaderi, Z., Rattner, B. P., et al., (2009) Intrinsic histone-DNA interactions are

not the major determinant of nucleosome positions in vivo. *Nat Struct Mol Biol* **16**, 847–852.

58. Tillo, D., Kaplan, N., Moore, I. K., et al., (2010) High nucleosome occupancy is encoded at human regulatory sequences. *PLoS One* **5**, e9129.
59. Field, Y., Fondufe-Mittendorf, Y., Moore, I. K., et al., (2009) Gene expression divergence in yeast is coupled to evolution of DNA-encoded nucleosome organization. *Nat Genet* **41**, 438–445.
60. Chavez, S., Garcia-Rubio, M., Prado, F., et al., (2001) Hpr1 is preferentially required for transcription of either long or G+C-rich DNA sequences in *Saccharomyces cerevisiae*. *Mol Cell Biol* **21**, 7054–7064.
61. Jimeno-Gonzalez, S., Gomez-Herreros, F., Alepuz, P. M., et al., (2006) A gene-specific requirement for FACT during transcription is related to the chromatin organization of the transcribed region. *Mol Cell Biol* **26**, 8710–8721.
62. Montoliu, L., Chavez, S., and Vidal, M. (2000) Variegation associated with *lacZ* in transgenic animals: a warning note. *Transgenic Res* **9**, 237–239.
63. Ramirez, A., Milot, E., Ponsa, I., et al., (2001) Sequence and chromosomal context effects on variegated expression of *keratin 5/lacZ* constructs in stratified epithelia of transgenic mice. *Genetics* **158**, 341–350.
64. Prado, F., Piruat, J. I., and Aguilera, A. (1997) Recombination between DNA repeats in yeast hpr1delta cells is linked to transcription elongation. *EMBO J* **16**, 2826–2835.
65. Sidhu, R. S., and Bollon, A. P. (1990) Bacterial plasmid pBR322 sequences serve as upstream activating sequences in *Saccharomyces cerevisiae*. *Yeast* **6**, 221–229.
66. Melcher, K., Sharma, B., Ding, W. V., et al., (2000) Zero background yeast reporter plasmids. *Gene* **247**, 53–61.
67. Kaplan, C. D., Laprade, L., and Winston, F. (2003) Transcription elongation factors repress transcription initiation from cryptic sites. *Science* **301**, 1096–1099.
68. Cheung, V., Chua, G., Batada, N. N., et al., (2008) Chromatin- and transcription-related factors repress transcription from within coding regions throughout the *Saccharomyces cerevisiae* genome. *PLoS Biol* **6**, e277.
69. Morillo-Huesca, M., Maya, D., Muñoz-Centeno, M. C., et al., (2010) FACT prevents the accumulation of free histones evicted from transcribed chromatin and a subsequent cell cycle delay in G1. *PLoS Genet* **6**, e1000964.
70. Vanti, M., Gallastegui, E., Respaldiza, I., et al., (2009) Yeast genetic analysis reveals the involvement of chromatin reassembly factors in repressing HIV-1 basal transcription. *PLoS Genet* **5**, e1000339.

Chapter 4

Tools of the Trade: Developing Antibody-Based Detection Capabilities for Recombinant Proteins

Maureen C. Dolan, Giuliana Medrano, Anthony McMickle, and Carole L. Cramer

Abstract

Protein-specific antibodies serve as critical tools for detection, quantification, and characterization of recombinant proteins. Perhaps the most important and widely used antibody-based procedures for recombinant protein applications are Western immunoblotting and enzyme-linked immunosorbent assays (ELISAs). These analyses require well-characterized, sensitive, and high-affinity antibodies that specifically and selectively recognize the recombinant target protein in the native or denatured form. Although the number of commercially available antibodies is quite substantial and rapidly growing, the appropriate antibody tools for many applications currently do not exist. In this chapter, strategies to develop and characterize both polyclonal and monoclonal antibodies directed against a specific protein of interest are discussed. Experimental strategies and methods are presented for producing and selecting the best antibodies and optimizing protocols for Western analyses, ELISAs, and other applications. Once antibody and procedure optimization is completed to ensure specificity, sensitivity, accuracy, and reliability, these immune-based approaches can now serve as powerful and enabling tools in the characterization, detection and diagnostics, structure/function analysis, and quality assessment of recombinant proteins.

Key words: Polyclonal antibodies, Monoclonal antibodies, Western immunoblot, ELISA, Recombinant protein detection

1. Introduction

When considering the synthesis of a foreign protein whether using microbial, animal-, or plant-based expression platforms, there are many aspects of its production process that need to be considered. One important component to determine prior to initiating expression of the protein is the availability of tools for detecting the product throughout the bioproduction process. High-affinity and selective

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_4, © Springer Science+Business Media, LLC 2012

antibodies are powerful tools used in tracking and quantifying a recombinant protein of interest. Robust, well-characterized antibodies and associated ELISA kits are frequently available for many human proteins and a number of other widely used model systems, such as the mouse. However, these important tools are generally lacking or of poor specificity for many recombinant proteins under development for medical, veterinary, bioenergy, and industrial applications. This limitation can be addressed initially by engineering epitope tags such as the widely used histidine, *c-myc*, and FLAG tags into the recombinant protein in order to track the expression of the protein product and facilitate its recovery for initial characterization and study (1). However, the presence of such tags on recombinant products for clinical applications and other commercial strategies is often not desirable. Regulatory constraints linked to the application of many protein products may prohibit presence of a tag and/or the tag may result in unpredictable effects on the protein function. Therefore, in addition to producing purified product for your intended studies, your recombinant protein can be used as an immunogen in generating protein-specific antibodies. These antibodies can ultimately enable quality and quantity assessment of the recombinant protein when integrated into commonly used tests including Western blotting, ELISAs, immunomicroscopy applications, and in vitro bioassays (Fig. 1).

Leveraging plant-based production platforms, our research team has recently expressed a chicken interleukin-12 (ChIL-12) protein with the intention of enhancing vaccine efficacy and/or promoting avian immune health (2). While we outline in a later chapter detailed methods and strategy for identifying the recombinant protein and establishing a product quality assessment process (see Chapter 29), ChIL-12-specific antibodies were needed to effectively execute these objectives. As our laboratory was the first to express recoverable amounts of ChIL-12, protein standards or specific antibodies were neither reported to have been produced nor available commercially. As one embarks on production of a new recombinant protein target, it is worth investigating whether homologous protein standards (e.g., isoform or homolog in a related species) and accompanying antibodies that could be used in detecting your protein of interest are available. In the case of our ChIL-12 protein, an antibody raised against the feline IL-12 (feIL-12) was reported to have moderate cross-reactivity for detecting the chicken protein homolog by Western blot analysis (3). However, we found this antibody lacked the specificity for effective and accurate quantification of our recombinant product either by Western blot densitometry or indirect ELISA. We addressed this limitation in two ways by (a) incorporating a 6×-histidine tag (HIS) that enabled us to use commercially available antibodies to the HIS epitope and (b) using our purified plant-derived ChIL-12 as an immunogen to develop protein-specific antibody tools.

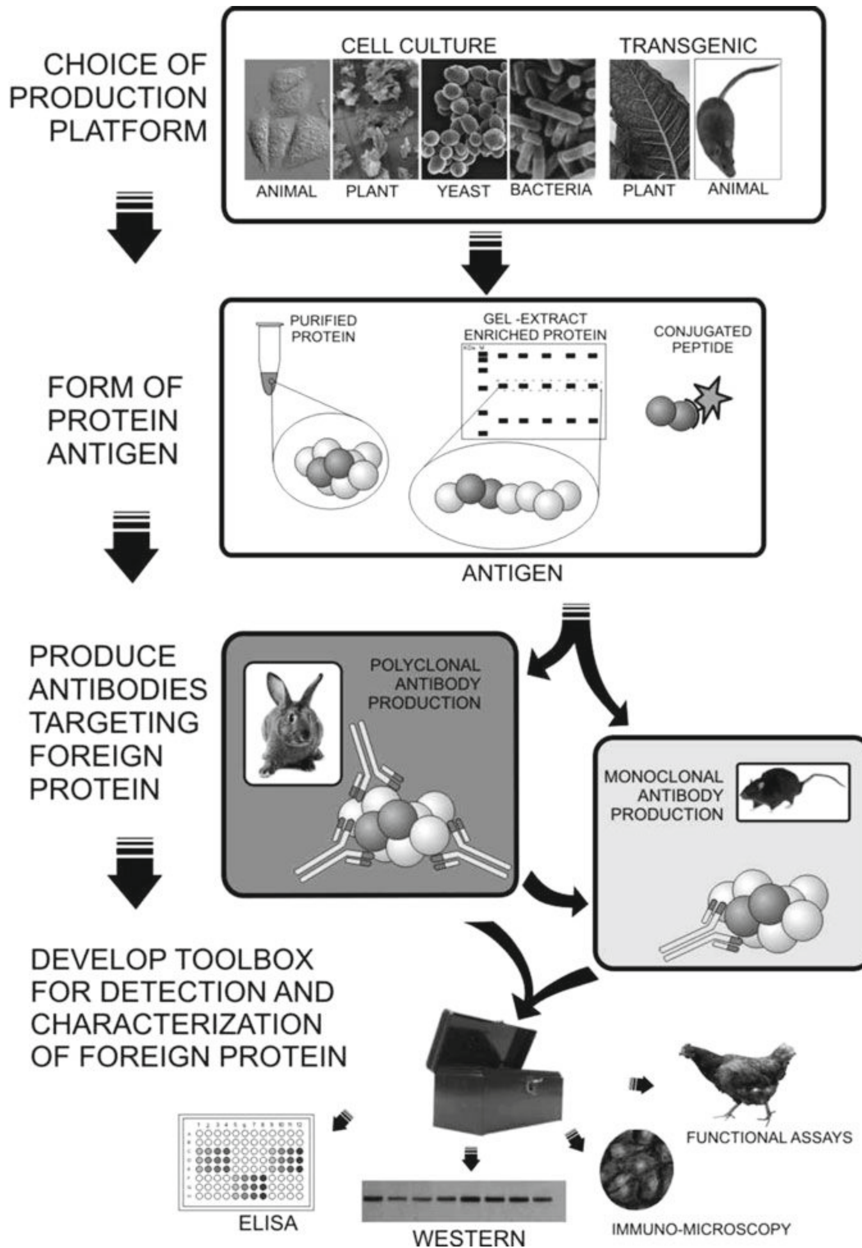


Fig. 1. Strategy for the development of an immunological-based toolbox to support recombinant protein product quality assessment.

In order to develop the appropriate immune tools for ChIL-12, we have taken a systematic approach to first producing a ChIL-12 polyclonal antibody in rabbits and then initiating monoclonal antibody production with a goal of developing paired antibodies for establishing a highly sensitive and quantitative sandwich anti-ChIL-12 ELISA (Fig. 1). Production of these anti-ChIL-12 antibodies

necessitated that we test and optimize their performance on both Western blot analysis and ELISA platforms. In addition, the polyclonal antibody exhibited neutralizing activity on the recombinant ChIL-12 therein providing a valuable tool for conducting bioactivity assays (2).

The following sections outline the strategies and methodologies our group has taken in constructing our "toolbox" for characterizing and detecting plant-expressed recombinant proteins. Specifically, we describe the strategies and methods that we used to:

1. Prepare our recombinant protein as an immunogen
2. Characterize and "clean-up" our newly developed antibodies
3. Select and optimize antibodies and conditions for sensitive and selective Western immunoblotting, and
4. Develop ELISA platforms for detection and quantification of ChIL-12.

This chapter is not intended to provide comprehensive methodologies for antibody production and their integration into assays; we will refer you to appropriate resources throughout this chapter for such details. Rather our goal is to share a general work plan that we have successfully employed for creating antibody tools in detecting and assessing a model recombinant protein, plant-derived ChIL-12. We believe a similar strategy may be applicable when expressing your foreign protein of interest on any one of the commonly used recombinant protein bioproduction platforms discussed in this book.

2. Materials

2.1. Polyacrylamide Gel Electrophoresis and Dye Staining

1. 10% Novex Bis-Tris precast SDS-polyacrylamide gels (Invitrogen).
2. 4× NuPAGE LDS Sample Buffer (Invitrogen).
3. 1 M DTT (dithiothreitol) or 1× commercial reducing agent.
4. 1× SDS-PAGE running buffer: Dilute 20× NuPAGE MOPS/MES SDS running buffer (Invitrogen) 20-fold with dH_2O to obtain the desired dilution.
5. Molecular weight standards (e.g., Precision Plus Protein marker, Bio-Rad; unstained for silver- or Coomassie-stained gels).
6. Mini-format protein electrophoresis unit (e.g., Invitrogen XCell Surelock Electrophoresis Apparatus).
7. Power supply (e.g., Power Pac Basic, Bio-Rad).

8. Gel stain: e.g., 0.1% (w/v) Coomassie Blue R250 (Sigma) in 20% (v/v) methanol, and 10% (v/v) acetic acid or SimlpyBlue SafeStain (Invitrogen Cat# LC6060).
9. Destain solution: 50% (v/v) methanol with 10% (v/v) acetic acid in water (for Coomassie stain) or water (for SimplyBlue Safestain).
10. Deionized water, clean scalpel blade, glass plate, microfuge tubes.

2.2. IgG Enrichment for Polyclonal Antibodies

1. Protein A Sepharose 4B Fast Flow resin (Sigma).
2. 10 mL Poly-prep column (Bio-Rad).
3. Binding buffer (1× PBS): 4 mM KH_2PO_4, 16 mM Na_2HPO_4, and 115 mM NaCl pH 7.4.
4. Elution buffer: 0.1 M Glycine pH 3.0.
5. Neutralization buffer: 1 M Tris–HCl pH 9.5.
6. YM-30 Centricon (Millipore).
7. Glycerol.
8. Spectrophotometer (280 nm; e.g., SmartSpec, Bio-Rad).
9. Column stand, microfuge tubes and tube rack.

2.3. Anti-ChIL-12 Western Immunoblotting Antibody Testing

1. Primary antibody: αChIL-12p70 (rabbit); produced by Cocalico (see Subheading 3.2.1).
2. Secondary antibody: α-rabbit IgG-AP (goat) used at 1:5,000 (Bio-Rad Cat#170-6518).

2.4. Anti-HIS Tag and Anti-ChIL-12 ELISA Checkerboard Tests

1. HIS-tagged protein standard (e.g., HSP90α^{HIS}, SignalChem Cat#H36-50 H).
2. Anti-HIS tag detection antibody [e.g., α-6×HIS(C-term)-HRP (mouse mAb), Invitrogen Cat#R931-25].
3. Antigen-targeting antibodies for sandwich ELISA:
   (a) αFeIL-12/23p40 (goat) (R&D Systems Cat#A-4187).
   (b) α6×-HIS mAb (mouse) (Calbiochem Cat#OB05).
   (c) αChIL-12p70 (rabbit); produced by Cocalico (see Subheading 3.2.1).
4. Universal detection antibodies for sandwich ELISA:
   (d) α-goat IgG-HRP (rabbit) used at 1:4,000 (Southern Biotech Cat#6020-05).
   (e) α-rabbit IgG-HRP (goat) used at 1:2,500 (Pierce ThermoScientific, Cat#31460).
   (f) α-mouse IgG-HRP (goat) used at 1:5,000 (Southern Biotech Cat#1030-05).

3. Methods

3.1. Immunogen Preparation

The structural and/or functional form of the recombinant protein that is used in producing a diagnostic antibody significantly impacts the utility of this new antibody in detecting and characterizing the recombinant protein on Western immunoblot, ELISAs, immune microscopy, and bioassay platforms. As depicted in Fig. 1, the recombinant protein used for immunization of the antibody-producing host animal, referred to as the immunogen, can derive from any number of sources including a purified or enriched fraction containing the protein of interest (typically "native" and/or functional protein) or protein resolved by electrophoresis and recovered as a gel slice or following transfer to a membrane (denatured protein). Alternatively, one can synthesize a peptide corresponding to a potentially immunogenic region (e.g., based on computational predictions) of the recombinant protein of interest and conjugate this to an appropriate "carrier" for immunization. Issues including the quantity and quality of your protein available for antibody production, the animal host used for producing the antibody, and the diagnostic application(s) for which the antibody is intended, all need to be considered when deciding the strategy for antibody production. Production of polyclonal antibodies in rabbit typically requires 300–500 μg of the purified protein of interest (less is required for a peptide immunogen). For monoclonal antibody production in mice, 250 μg of protein is recommended by most service providers. Choosing to develop polyclonal versus monoclonal antibodies (or both), is a strategic decision based on anticipated application(s), amount of antibody required, and cost, discussed further in Subheading 3.2.

3.1.1. Preparation of Soluble "Native" Protein Immunogen

Ideally, if the recombinant protein can be purified in an active form, this is the most desired form of the protein against which to raise antibodies. In practice we have found that this form of the immunogen typically has the highest probability of generating antibodies that will recognize the native protein. The ability to detect protein that is structurally and functionally intact is often important and necessary for their use in ELISAs, immunoprecipitation assays, immunomicroscopy, and bioassessment platforms. The purity of your protein of interest in the immunogen sample will greatly impact the specificity of the resulting antibodies. It should be noted that different proteins can be more or less immunogenic – a minor contaminant that is highly immunogenic can unfavorably skew the resultant antibody population. A sample of the submitted immunogen solution should be retained and used in Western blotting procedures to assess the quality of the antibodies produced. There are ways to enhance the specificity (i.e., to recover sub-populations from a polyclonal antibody population) once you

have your antisera. However, the potential to “chase the wrong band” or to misrepresent the amount of your target protein due to cross-reactivity to endogenous co-purifying contaminant proteins is very real. As described below, characterization of your newly developed antibodies with appropriate controls (e.g., extracts from “mock” or empty vector transfections as negative control; extracts from the gene-origin organism as positive control) are important to developing quality reagents.

In the case of our plant-expressed ChIL-12 protein, we were able to recover sufficient quantities of relatively pure protein to use as antigen in generating both anti-ChIL-12 polyclonal and monoclonal antibodies. The purification of this HIS-tagged recombinant protein is described in detail in Chapter 29. Prior to submission of ChIL-12 protein for antibody production, purity was assessed by silver staining (see Chapter 29, Subheading 3.5.1), the sample was “cleaned up” by buffer exchange to phosphate buffered saline (PBS), and the concentration was estimated using anti-HIS and anti-feIL-12 antibodies. The protein was stored at 4°C and shipped overnight on ice. If your protein is stable following lyophilization, this is also an good option for recombinant protein storage.

In the case of another protein produced in our laboratory using an *Escherichia coli* production platform, the protein of interest was only recovered in the inclusion bodies and not in a soluble or native form. The inclusion bodies containing the recombinant protein were subject to standard refolding protocols containing detergent and chaotropic agents to generate the immunogen. In this case, while this post-production process did enable solubilization of the protein, the recombinant product was rendered inactive likely due to improper refolding of the polypeptide. While this is typically not an ideal immunogen for generating polyclonal antibodies with maximum probability of recognizing the native protein, in this instance the resulting antibody did offer a high-quality reagent that recognized the denatured recombinant protein by Western blotting analysis with antibody titers at 1:20,000 and that exhibited notable sensitivity (as low as 1 ng/lane; data unpublished, M. Reidy and C. Cramer).

3.1.2. Preparation of Immunogen in a Denatured or Non-native Form

If chromatographic purification yields insufficient quantities of your protein of interest, another approach for generating the immunogen source is to size fractionate an enriched protein extract containing your recombinant product by sodium dodecyl polyacrylamide gel electrophoresis (SDS-PAGE). Subsequent excision of the Coomassie-stained gel slice corresponding to the expected protein band size as determined by a protein mass standard will provide recombinant protein that is further enriched from host contaminating proteins. You will need to estimate the amount of the protein in the gel slice by either densitometry of the stained gel or Western blot analysis with an alternate antibody (e.g., anti-6×

HIS antibody if your recombinant protein has been expressed with a histidine tag). The excised gel slice can be processed and used directly for immunization of the host animal. This "presentation" may provide an added benefit of slow release of the immunogen into the host facilitating a good quality antibody response. The downside of this approach is the immunogen is denatured and antibodies generated by this method may not be capable of recognizing your recombinant protein in its native state. Again for some applications (e.g., Western blot analysis) this is generally not an issue. However, if you require the antibody to detect your recombinant protein in its functional and/or structural conformation (e.g., ELISAs, immune microscopy, etc.) this is not the approach of choice. Although not described here, it is also possible to provide your immunogen to an antibody production service on a membrane. In this case, proteins size-separated by SDS-PAGE are transferred to a membrane (e.g., PVDF, nitrocellulose) and visualized by Coomassie staining. The membrane can be submitted directly to the antibody service provided along with a scan of the membrane with the specific band(s) to be used as the antigen clearly delineated. The service provider will excise the band and process the protein of interest according to their established protocols.

The procedure for running a protein sample on an SDS-PAGE in the NuPAGE system has been detailed in a later chapter (see Chapter 29, Subheading 3.2). The following summarizes the process of generating immunogen as a gel slice to be used for subsequent antibody production:

1. Set up the precast gels (e.g., NuPAGE Novex Bis-Tris) in an electrophoresis apparatus as per instructions, prepare 1× SDS-PAGE running buffer and fill chamber with buffer (detailed in Chapter 29, Subheading 3.2.2). All reagents and supplies for this procedure are provided in Subheading 2.1 above.
2. Prepare your protein sample with sample buffer (LDS) and reducing agent as described in Chapter 29 Subheading 3.2.1. Typically, we do not exceed 30 μg total soluble protein per well. Depending on the purity/concentration of your target protein and the amount required as immunogen, more than one lane or a "preparative gel format" may be needed (see Note 1).
3. Load samples into wells of NuPAGE gel. Load an unstained protein marker standard (e.g., Bio-Rad Precision Plus Protein marker) in a lane adjacent to the protein sample lanes (see Note 2) and carry out the electrophoresis as described (see Chapter 29, Subheading 3.2.2).
4. Stain the gel using either a Coomassie stain or other comparable commercially available stain (e.g., SimplyBlue Safe Stain; Invitrogen) and then destain as appropriate to resolve bands corresponding to expected size of the recombinant protein.

5. Transfer the gel to a clean glass plate and wrap the gel/plate sandwich in transparent plastic wrap (e.g., Saran wrap) to hold the gel in position. Using a clean, sharp, scalpel blade, excise the band(s) containing the target protein from the gel and transfer to a plastic tube.
6. Further destain the gel slices. If an organic-based destain is used, it is necessary to perform a final rinse in deionized water to remove any traces of methanol and acids.
7. Store the tube containing gel slices at 4°C for immediate transport on ice to the antibody production facility of choice.

A peptide-based polyclonal antibody may be the best option in the event that (1) sufficient amounts of your recombinant protein cannot be obtained, (2) an antibody to detect native protein is requisite and/or (3) discrimination of your target protein from other related proteins that share considerable homology is needed. Peptide-based antibodies are most commonly derived from C- or N-terminal regions of a target protein and are often selected based on computation predictions of regions having greater potential for antigenicity (4, 5). If these terminal amino acids are predicted or known to present on the surface of the structured polypeptide, the corresponding antibody may in fact be capable of detecting the protein in its native conformation. Most peptide immunogens range from 12 to 16 residues. As a general rule these shorter amino acid sequences generate antibody with selectivity to a single epitope, whereas longer peptides may generate antibodies corresponding to several epitopes on the recombinant protein. To ensure immunogenicity, the peptide needs to be conjugated to a carrier protein such as keyhole limpet hemocyanin (KLH), bovine serum albumin (BSA), ovalbumin or thyroglobulin. There are a number of companies that produce polyclonal antibodies who also provide peptide synthesis and conjugation services (e.g., Invitrogen, Pacific Immunology, Covance, etc.) for an additional fee. This can generally streamline the production process and can be more cost effective than subcontracting these steps to other companies or doing the conjugation chemistry yourself.

As mentioned above, there are recommended amounts of protein immunogen required depending on the animal host used and type of antibody produced. Immunogens may be submitted in a variety of buffers; however, PBS or saline solution (0.9% sodium chloride) are most commonly used (see Note 3). It is often recommended to keep the protein concentration high and buffer volume as low as possible; generally immunogens should be around 1 mg/mL and not lower than 100 μg/mL. In the case of peptide immunogens, most polyclonal antibody production services require 2 mg for production in a rabbit host. Finally, the antigen should be filter sterilized in efforts to reduce inflammation at the injection site of the host animal not only for ensuring good laboratory animal

Table 1
Comparison of key features in considering the production of polyclonal and monoclonal antibodies to detect a recombinant protein of interest

	Polyclonal Ab	Monoclonal Ab
Timeline production/ characterization	≥3 months (e.g., rabbit)	6–12 months
Cost	~$350 to $1,000	$6,000 to $15,000+[a]
Supply	Limited	Unlimited
Relative avidity[b]	Generally higher	Generally lower
Specificity	Multi-epitope[c]	Single epitope
Production host(s)	Varied: rabbit, rodents, goat, donkey, chicken, etc.	Limited: mouse, less commonly rat and rabbit

[a] Widely variable depending on vendor
[b] Avidity represents overall binding intensity between antibodies and a multivalent antigen presenting multiple epitopes
[c] Peptide-based monospecific polyclonal Abs can be directed at a single epitope

practices but to avoid undesired immune responses in the host that could adversely affect the quality and quantity of the antibody desired.

3.2. Producing Antibodies for Detecting Recombinant Protein

The two choices of antibodies most commonly used in R&D for detection of proteins are polyclonal and monoclonal antibodies. In general, monoclonal antibodies detect a single type of antigen binding site (epitope) and are derived from a single B cell clone. Conversely, polyclonal antibodies are generally complex mixtures directed against different epitopes, derived from multiple B cell clones and each antibody type in the mixture has differing affinity for the immunogen. Detailed discussion of these two types of antibodies is beyond the scope of this methods chapter; however, Table 1 provides a comparison of their distinguishing features that can impact the decision process. The choice of antibody production platform to use in generating an antibody against your recombinant protein may be dictated by differences in their production costs, need for epitope specificity, amount of antibody ultimately required, the application for which the antibody will be used and the relationship between the donor of the antigen and the antibody producer. More comprehensive discussion of polyclonal and monoclonal antibodies can be found in a recent review (6) as well as standard protocol manuals (7, 8).

Many universities have core facilities that offer polyclonal and/or monoclonal production services both for their campus clients as well as off-campus patrons. Alternatively, there are numerous commercial vendors that will produce antibodies at a varying range of price points. Here is where it is very important to do your homework!

The first consideration is identifying the producer who will deliver the desired quality antibodies, with an important second consideration that the service is at a cost compatible with the project budget. Based on our experience, we have found that it is important to determine the level of services you desire and have the detailed cost breakout of these services clearly delineated by the service provider producing the antibody. Understanding the number of boosts to be performed, number of animals to be used, if pre-testing of host will be performed, the extent of screening performed by the producer throughout the process, and the quality and quantity of antibody(s) that will be delivered are all important are all determining an accurate cost for producing your antibody.

Due to the lower complexity associated with the overall production process, services for production of polyclonal antibodies are typically simpler to compare and easier to accurately estimate the real cost for the product up front. Conversely, monoclonal antibody production services can vary extensively and differences in advertised costs are generally linked with the methods and schedules used in inoculating the host, the types of technology used in screening and choosing elite antibody candidates and establishing the best antibody pairs, and the amount and number of independent monoclonals guaranteed. If you are considering producing monoclonal antibodies against your recombinant protein, we would strongly encourage contacting the producer prior to production and have them clearly delineate the process, the associated cost of each step and the potential pitfalls in the process that may result in additional fees. Another important aspect that can significantly impact the performance of the monoclonal product(s) generated is associated with the biochemical and structural properties of the target protein. Undoubtedly monoclonal antibody production is a multifaceted process and thus a firm production cost can sometimes be difficult to predict *a priori*.

As the goal in the research setting is often to produce antibodies with high sensitivity (titer) and specificity (affinity), another component to be considered is the adjuvant that will be co-administered with the antigen. An adjuvant aids in stimulating the host immune response and improving the quality and quantity of the antibody produced. In general, there are three commonly used adjuvants for antibody production (9, 10). Freund's adjuvants are water-in-oil emulsions that enable slow-release of the immunogen (up to 6 months) and are associated with eliciting robust immune stimulation for diagnostic antibody production. Complete Freund's adjuvant (CFA) is typically used for the initial inoculation of immunogen. Incomplete Freund's adjuvant (IFCA) must be used in all subsequent booster immunizations if FCA was used in the initial immunization. The use of the Ribi adjuvant (oil-in-water emulsions) has advantages associated with its ability to interact with membranes of immune cells which help to enhance the uptake,

processing and presentation of the target protein. Titermax adjuvant system has gained popularity and more widespread adoption recently as a means of potentially minimizing the inflammatory response that may negatively impact the quality and performance of the antibody produced. In some cases the organization producing your antibody may designate the adjuvant used which is linked with their registered Standard Operating Procedures for producing antibodies in laboratory animals.

A useful website that is regularly updated and a good place to start your search for either basic information on antibodies and antibody production as well as a current list of companies that will produce custom antibodies is http://www.antibodyresource.com/.

3.2.1. Polyclonal Antibodies

In general, polyclonal antibody production is the most common and often best choice for generating antibody to detect your recombinant protein of interest. As outlined in Table 1, the lower relative cost and shorter timeline for generating product are usually the driving factors that favor this antibody production platform in the research setting. However, due to the finite amount of antibody generated from a given host animal, you will want to determine the number of animals to immunize for each target immunogen. Many production facilities recommend immunizing two animals as the animal hosts used are primarily outbred species and therefore each animal can respond quite differently to the same antigen. Following your antibody screening (e.g., Western blotting, ELISA, etc.) if multiple animal hosts produce similar quality antibodies to a given immunogen, the serum may be pooled to create a single lot of polyclonal antibody in meeting the needs of a research laboratory and will last for years if properly stored.

If the antibody is to be more broadly used (e.g., as a commercial research reagent or as a certified QA/QC component of a certified product), one must recognize that the yield per animal is limited and the animal is sacrificed. Immunization of additional animals with the same antigen will require full characterization of the new antibody population, which may demonstrate different affinities. To address these limitations, commercial applications often employ monoclonals or use polyclonal antibodies that have been raised in larger host animals. Polyclonal antibody production in rabbits, clearly the most common host for antibody production, provides an estimated 100 mL of whole serum recovered per animal (8). If your project requires large amounts of an antibody be produced, you may consider a larger mammal (e.g., goat, horse) for production. Another commonly used host for generating polyclonal antibodies is chicken. While polyclonal antibodies produced in mammals are predominantly of the immunoglobulin G subtype (IgG), the circulating Ig counterpart in avian species is IgY. Chicken IgY polyclonals are structurally similar and functionally equivalent to serum IgG antibodies and in some cases offer several advantages

over mammalian-based antibodies (11). Because of the volume of eggs produced by laying chickens, the amount of antibody generated can approach that obtained from a large mammalian host. It should be noted that in our project we intentionally avoided this host system due to its phylogenetic identity to our targeted recombinant ChIL-12 protein. However, we have generated IgY antibodies for other plant-expressed recombinant proteins with success.

For polyclonal antibody production against our plant-expressed recombinant ChIL-12 protein, we chose a rabbit host and used a commercial producer (Cocalico Biologicals). It should be noted that each antibody producer may have slightly different requirements for the immunogen and/or inoculation schedules. For the ChIL-12^{His} expressed transiently in *Nicotiana benthamiana* leaves, we provided purified protein taken through a two-step fast protein liquid chromatography (FPLC) process outlined in detail in Chapter 29. Typically when producing a polyclonal antibody in a rabbit host, approximately 300 μg of protein per animal is requested. In this case, pricing was based on a 90 day production timeline at a cost of $279/rabbit. The immunization schedule of the host rabbits used in producing our ChIL-12 polyclonal antibody was as follows:

- Day 0 – Prebleed/initial inoculation
- Day 14 – Boost
- Day 21 – Boost
- Day 35 – First test bleed-serum tested by Western and ELISA
- Day 49 – Boost
- Day 56 – Second test bleed-serum tested by Western and ELISA
- Exsanguination – Third/final bleed; antibody optimization for establishing standard anti-ChIL-12 Western and ELISA

The initial inoculation of 100 μg of ChIL-12 (soluble in PBS; >90% purity) was administered in Complete Freund's adjuvant (CFA). For the three subsequent boosts 50 μg antigen admixed with Incomplete Freund's adjuvant (ICFA) was used. The route of injection is typically subcutaneous and/or intramuscularly administered at multiple sites (see Note 4).

The antibody producer generally ships the pre-bleed sample for each host at the time the first test bleed serum is sent unless pre-testing serum of multiple hosts prior to initiating the immunization process is requested by the client. The second bleed should exhibit a titer (amount of antibody targeting the immunogen) greater than that exhibited for the first bleed (see Subheading 3.4.2). Initial titer and specificity can be determined by direct ELISA (using your immunogen as the capture reagent) and/or Western

immunoblot (described below in Subheadings 3.3 and 3.4). If the titer is not at the level desired, consultation with the producer may provide an option for performing additional boost(s)/bleed(s) to enhance the antibody pool for detecting the recombinant protein. Upon satisfactory test results from the second bleed sera containing the antibody, the client contacts the producer to confirm the final bleed (exsanguination) of the host animal(s). If multiple hosts for a single immunogen were requested, all serum fractions at each bleed should be tested independently as these are by definition different antibody pools. Overall, the quick turnaround time (about 4 months) for generating a ChIL-12 polyclonal was a significant advantage in this project as it allowed for rapid mobilization of our studies in meeting the requisite research timeline.

In an effort to lower costs, we typically order non-filtered sera and conduct all testing of our polyclonal antibodies in house by Western blot analysis and ELISA. For long-term use (years), aliquots of the polyclonal antibody can either be maintained and stored as whole serum or further purified. In the case of our ChIL-12 polyclonal antibody, sufficient titers and sensitivity were obtained without purification from the rabbit whole sera. We typically store single-use aliquots in a manual defrost −20°C freezer to maximize reproducible performance of the antibody in our assays. Larger aliquots of sera are stored at −70°C. It should be noted that freeze-thaw cycles adversely impact antibody quality and can differentially impact antibody subpopulations potentially changing the behavior of the overall population. Thus storage should address long-term and short-term needs in a way that provides stable storage and minimizes freezing and thawing. Alternative storage options include addition of 30–50% glycerol to limit antibody loss due to denaturing during the thawing process and addition of sodium azide to inhibit any bacterial growth for shorter term maintenance (generally months) of antibody at 4°C. The websites of many polyclonal antibody producers provide relevant information and guidelines for producing antibodies with them. In addition they often provide useful general information on the subject. Some sites we have found helpful are listed at the end of this chapter.

3.2.2. Monoclonal Antibodies

Because of their precise single-epitope selectivity, monoclonal antibodies (mAb) can provide distinct advantages over polyconal antibodies. However, the need, value, and anticipated return on investment for producing monoclonal antibodies to a recombinant protein should be critically evaluated before committing to this process. As highlighted in Table 1, the issues of cost and time are significantly higher than in developing a polyclonal antibody. Therefore, before investing in this major undertaking, clearly identify the questions needed to be addressed regarding the recombinant protein (quantity, protein structure, protein function, etc.) and the "tools" needed in designing the appropriate assay(s) to

answer these questions. It is quite possible that a polyclonal to the recombinant protein, an existing commercial antibody that cross-reacts (e.g., antibody to the protein in a related species) or an antibody targeting a purification tag on the recombinant protein may be sufficient for designing an effective assay.

As part of our project expressing recombinant ChIL-12, we were interested in also developing a diagnostic ChIL-12 ELISA to integrate as part of a quality assessment program for producing recombinant ChIL-12 as a veterinary therapeutic (see Chapter 29 of this book) as well as delivering a high-performance ChIL-12 detection assay of value to the poultry research arena and industry. In meeting this objective, it was important that the antibody sources used in developing this assay were renewable. Due to the heterogeneity of polyclonal antibody pools (animal-to-animal variation of immune response), when one antibody supply is depleted, a new pool of polyclonal antibody must be produced. This new antibody pool is subject to a complete retesting process and the ELISA parameters must be re-optimized to ensure quantitative accuracy of the assay. In contrast, the clonal nature of monoclonal antibodies enables individual B-cells expressing a given antibody to be immortalized and cryopreserved, thus providing continuous access to an antibody reagent having the volume and lot-to-lot consistency that will support a commercially sustainable ELISA or other immune-based assay.

There are many production options for monoclonals; however, a highly skilled and established facility to produce the antibody to your recombinant protein (e.g., Antibody Solutions) can facilitate the process when project timeline constraints are a concern. It is worth mentioning that although at a higher price point, those companies that integrate more advanced technology throughout this process generally deliver higher success rates and a better-quality monoclonal product. Specific vendors may differ in the details of their process for mAb production, but the general steps discussed here are fairly universal for all monoclonal production facilities. While detailing this process is beyond the scope of this chapter, the following summarizes the general workflow we followed for producing monoclonal antibodies to recombinant ChIL-12 protein.

For the production of ChIL-12 monoclonal antibodies, a five step process involving: (1) immunization, (2) hybridoma libraries, (3) monoclonal cultures, (4) pairing analysis and (5) grow/store/purify/conjugate – took approximately 10 months to complete. A total of 2.5 mg of purified protein was required when initiating immunization with four mice. Two mice exhibiting the highest serum ChIL-12 antibody titers (as determined by indirect ChIL-12 ELISAs; see Subheading 3.4) were moved forward in establishing hybridoma libraries to immortalize this antibody response (process that fuses immunized mice spleen cells with an immortal

myeloma cell line). From these libraries, monoclonal cultures originating from single, viable cells selected by flow cytometry cell sorting resulted in 20 independent monoclonal lines being established and cyropreserved. Because our goal in producing monoclonal antibodies for ChIL-12 was to develop a sandwich ELISA, a pairing analysis with the 12 lead monoclonal antibodies was performed that employed the Surface Plasmon Resonance technique. This step, which quantified the affinity and binding equilibrium between different known concentrations of the ChIL-12 antigen and each of the monoclonal antibodies, involved the testing of a total of 72 possible antibody pairs. Following identification of the lead pairs, the monoclonal facility expanded production, purified and cryopreserved the three lead monoclonal cell lines (M6, M8, M19). This process ensured sustainability and provided reagent for testing and assay development. Select monoclonals were also biotinylated for purposes of designing an anti-ChIL-12 sandwich ELISA (see Subheading 3.4).

3.2.3. Purification Options for Antibodies

Several purification methods can be employed in "cleaning up" and/or concentrating a new antibody. For monoclonal antibody production, purification methods are typically integrated during the scale-up phase of the process. As approximately 10% of the serum proteins are immunoglobulins (Igs) and of these the antibodies of interest that react specifically to your recombinant protein typically only account for about 1–5% (reviewed in ref. 6), we commonly perform purification of our polyclonal antibodies in-house. In the case of our ChIL-12 polyclonal antibody, purification was not necessary as the whole sera fraction provided good sensitivity with no background interference in both Western immunoblotting and ELISA. However, in the event that the antibody demonstrates higher background noise in your assay than desired and/or the titer is low (see Note 5), you may chose to use one of a number of available purification approaches to enrich for the antibodies targeted to their protein of interest. In the case of polyclonal antibody production, purification is done after the test sera has been confirmed to have antibody specificity for the intended recombinant protein target (e.g., by Western blot analysis see Subheading 3.3; by ELISA, see Subheading 3.4) and the host animal(s) has been exsanguinated and all sera collected.

There are two basic approaches typically used to enrich for the Igs – salt fractionation and chromatography. In our laboratory we have used both methods for enriching some of the polyclonal antibodies produced. There are a number of references and resources with detailed protocols for both methods (e.g., (12)). The following method focuses on a chromatographic IgG enrichment strategy we have successfully employed that reduced the background

on Western blots compared to that observed when using a whole sera fraction containing rabbit-produced polyclonal antibody to a plant-expressed protein:

1. Use Protein A-based chromatography (e.g., Protein A Sepharose; Sigma) to enrich for the total IgG fraction according to manufacturer's instructions (see Note 6).
2. Equilibrate 1 mL packed resin of Protein A Sepharose (binding capacity = ≥30 mg/mL human IgG) with 5 mL Binding buffer (PBS) by gravity flow in a 10 mL column clamped to a ring stand (see Note 7).
3. Dilute 1.5 mL of sera (based on an estimated IgG concentration of 12–14.5 mg/mL in rabbit sera) with 3 volumes of Binding buffer and load onto the column. Pass the flow-through back through the column several times to ensure binding of IgGs to the resin.
4. Remove unbound material by washing with 3 column volumes of PBS.
5. Prior to elution of the IgG fraction, prepare microfuge collection tubes that contain 30 μL of 1 M Tris–HCl pH 9.5. This allows the elution fractions (0.5 mL) to be immediately neutralized, thus stabilizing the antibodies following acid elution.
6. Elute the bound IgG from the column by adding 10 mL of Elution buffer (100 mM glycine pH 3.0) and collecting 0.5 mL fractions into the collection tubes containing the neutralizing buffer.
7. Assay the fractions for absorbance at 280 nm to identify protein (i.e., antibody) containing fractions; pool fractions containing significant levels of antibody.
8. Repeat this chromatography process for additional fractions of sera; and collect and pool all IgG fractions.
9. Concentrate pooled IgG antibody via YM-30 Centricon (Millipore) following manufacturer's instructions.
10. Add glycerol to a final concentration of 30–50% to stabilize antibody for freezer storage conditions.
11. Aliquot antibodies and store at −20°C (typically working stocks and single-use aliquots) and −70°C (generally larger volumes for longer-term storage). See Subheading 3.2.2 for further details.

Protein G is an alternative resin, which has particular advantages in purifying mouse IgG1 subclass immunoglobulins (the dominant form in IgG in mice) from ascites (e.g., Thermo Pierce NAb Protein G Spin Kit, #89949). Alternatively, Protein A/G columns based on a chimeric protein combining the Protein A and G IgG binding domains (e.g., Thermo Scientific Pierce Protein A/G Resins) provides a more universal option for recovery of multiple IgG subclasses present in polyclonal and monoclonal

antibodies as well as for some of the other Ig subtypes (with the exception of mouse). A variety of binding buffers are available that are designed to effectively bind Ig subtypes and subclasses to the resins as well as optional elution buffers for "gentler" release of antibodies compared to the classic acid buffer release (e.g., see Thermo Pierce for a number of Protein A, G and A/G Binding and Elution buffers). Protein L, which binds to certain immunoglobulin kappa light chains that occur in members of all classes of immunoglobulin (i.e., IgG, IgM, IgA, IgE and IgD), may provide an option for purifying these different classes of antibody. Comparable protocols and kits are also available for IgY purification from eggs (e.g., Pierce Chicken IgY Purification Kits #44918 and 44922) where a major challenge is delipidation to ensure optimal function of IgY antibodies in diagnostic assays.

Often regarded as the "poor man's alternative" to a monoclonal antibody, peptide-generated polyclonal antibodies subsequently processed using a peptide affinity column is a good option. The peptide can also be used to select for antibody detecting a single epitope on the recombinant protein (monospecifc polyclonal antibody) at a fraction of the cost of mAb production and with similar specificity. Likewise, phospho-specific IgG antibodies may be purified using a phospho-peptide column. Also by using a non-phosphorylated peptide, the phospho-specific IgG can be obtained by a subtractive strategy by collecting the flow through.

3.3. Optimizing Antibodies for Use in Western Blot Analysis

The Western blot analysis is commonly the first assay used to "test drive" a new antibody that has been produced to detect a recombinant protein. In this procedure, proteins are denatured by boiling in SDS, size separated by gel electrophoresis, transferred to a membrane, and the protein of interest is specifically detected based on interactions with antibodies. This immune-based test is typically carried out using one or two antibodies. In this assay, the antibody produced to recognize the recombinant protein is referred to as the primary antibody and binds directly to the antigenic epitope(s) on the protein of interest. While the primary antibody can be covalently coupled with an enzyme that reacts with an added substrate in driving the detection of the antigen/antibody complex (e.g., anti-HIS coupled to alkaline phosphatase, see Subheading 3.4.2), a secondary antibody is more often used in Western blot analysis. A secondary antibody whose epitope is selective for the Fc region of the host used in producing the primary antibody is covalently conjugated (in its Fc region) with either biotin or any number of reporter enzymes; the most commonly used are horse radish peroxidase (HRP), alkaline phosphatase (AP), and β-galactosidase. Such conjugated secondary antibodies are widely available from a number of commercial antibody vendors and generally work with a variety of substrates for yielding detectable products. As general rules, secondary antibodies are aliquoted for purposes of maximizing

reagent shelf-life and purchased from a limited number of vendors (see Note 8) to reduce need to re-optimize assay conditions associated with source and/or lot variability (see Note 9). Finally, our laboratory typically uses a chemiluminescence-based system for Western blotting detection as described in Chapter 29, While less sensitive, subheading 3.2.3 of this chapter outlines a colorimetric-based method as an alternative detection option.

When testing a new antibody for Western immunoblotting, running the proper controls during the optimization phase of the assay is crucial. It is useful to establish the specificity of the antibody for the recombinant protein in not only the purified sample but also crude extracts and at various stages during the protein purification process in defining the diagnostic breadth of this immunological tool. Inclusion of these control samples is critical for ensuring that the new antibody indeed has specificity and selectivity for its intended target. If the recombinant protein can be obtained from another source (e.g., expressed in an alternative production system, cells/tissues naturally expressing the protein, synthetic peptide, etc.) this can provide an additional positive control to confirm the selectivity of the new antibody for your target protein. Testing negative controls that represent the corresponding "background" of the sample containing the recombinant protein is vital to this optimization process. For example, in our study, we use samples corresponding to plants transfected with an "empty" vector (i.e., the pBIB-Kan vector lacking the gene that encodes our recombinant protein of interest) and treated in parallel through the extraction and purification process as plants expressing the recombinant ChIL-12. These samples, referred to as "pBK" or "mock," serve as negative controls and ideally should show no cross-reactivity to the newly developed antibodies. Another often overlooked control when optimizing a new Western blot assay is testing the secondary antibody alone (e.g., no incubation with primary antibody) to confirm that it does not cross-react with proteins in the various test samples. The importance of running the proper controls during the optimization of your Western blotting procedures cannot be overstated and is essential to developing an accurate assay that reliably characterizes and defines a recombinant protein product.

3.3.1. Designing a Western Blot Optimization Test for New Antibodies

Among the challenges in working with a new antibody is establishing the appropriate concentrations of not only the new antibody but also the relative concentrations of this antibody to the secondary antibody, both of which are important in confirming the specificity of the new antibody for the intended target protein. While the following section details the procedure we have used in optimizing a ChIL-12 polyclonal antibody, the concepts, controls, and technical

tips used in conducting this matrix test can be applied for developing a well-performing Western blot analysis for any new polyclonal or monoclonal antibody:

1. Load protein samples onto a 12% NuPAGE Bis-Tris gel and resolve according to the protocol detailed later in this book (see Chapter 29, Subheading 3.2.2). Include the recombinant protein (purified and crude extract) and an equivalent volume of the empty vector (pBK) negative control in a repeated series. Multiple gels may need to be run depending on the number of antibody conditions to be tested. It is important to include at least one lane with a protein molecular size standard for establishing/confirming target protein molecular mass on SDS-PAGE (see Note 10).
2. Following electrotransfer of the gel to a nitrocellulose membrane (see Chapter 29, Subheading 3.2.3), remove the membrane, rinse it with water, and place it sample-side up on a clean paper towel (see Note 11).
3. To perform an antibody optimization test, cut the nitrocellulose membrane into strips with each test strip including recombinant protein and pBK negative controls. Light pencil marks can be used to mark on the membrane for cutting membrane test strips (see Note 12).
4. Place each membrane strip in a small container with protein sample side of the membrane facing up (see Note 13). The membrane must lay completely flat in the container to ensure proper coverage by solutions during the incubation and wash processes. Ideally in an effort to limit the amount of antibody reagent required, use a container that minimizes the volume of solution needed while providing coverage of the membrane (see Note 14).
5. "Block" the membrane strips by incubating in 20 mL of 3% BSA in PBS per container for 1 h on a platform shaker at ~100 rpm. While many protocols call for using non-fat dry milk as a blocking agent, we have found much better consistency and improved backgrounds on Western blots using BSA as the blocking agent. In addition, we strongly encourage using the same vendor/product catalog number for your blocking agent of choice to mitigate possible background issues associated with different brands of product (see Note 15).
6. Use 20 mL of solution per container for all primary and secondary antibody incubations and wash steps. Perform these incubations in accordance with the procedure outlined in Chapter 29 Subheading 3.2.3 of this book employing various dilutions of primary and secondary antibody (see discussion below).

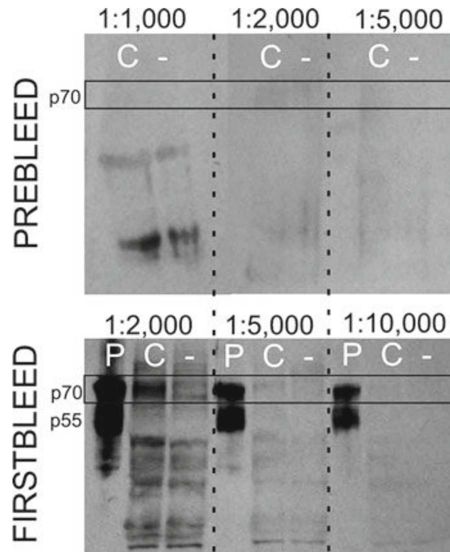


Fig. 2. Optimization of ChIL-12 polyclonal antibody on Western blot analysis. Crude (C) and purified (P) ChIL-12^{HIS} or crude pBIB-Kan "mock" control (–) protein extracts separated by SDS-PAGE run under denaturing conditions. Immunoblotting was performed with serial dilutions of primary antibody in characterizing a new anti-ChIL-12 polyclonal as described in Subheading 3.3. The positions of the ChIL-12 p70 and p55 bands are indicated *on the left*.

7. Following the final wash step, arrange the membrane strips such that each strip is abutted adjacent to each other and positioned between the sheets of a plastic page protector. Add developing solution to all of the aligned membrane strips cover, and incubate to enable visualization of target proteins.

The first experiment in conducting an antibody screening test involves using the pre-bleed serum as the primary antibody to determine the antibody titer at which no signal is observed on the immunoblot under standard assay development conditions. Specifically for our ChIL-12 polyclonal antibody, pre-immune serum was tested at six different dilutions ranging from 1:200 to 1:10,000. In this test the conjugated secondary antibody, anti-rabbit IgG-AP, was used at 1:5,000. As shown in Fig. 2, the pre-immune serum of the host rabbit began to exhibit non-specific binding to proteins in both crude protein extract lanes [ChIL-12 (C) and negative control pBK (–)] at a 1:1,000 dilution that was not detected when antiserum was used at 1:2,000. Following this pre-bleed test, several concentrations of the new first-bleed primary antibody were tested with the lowest dilution limit defined by this pre-bleed serum test, i.e., the dilution level showing no non-specific cross-reactivity. Beginning with this starting dilution, we generally dilute the test-bleed antiserum further by a factor of 2 (depending on the titer of the antiserum) and test a minimum of three different antiserum dilutions. In the case of our ChIL-12

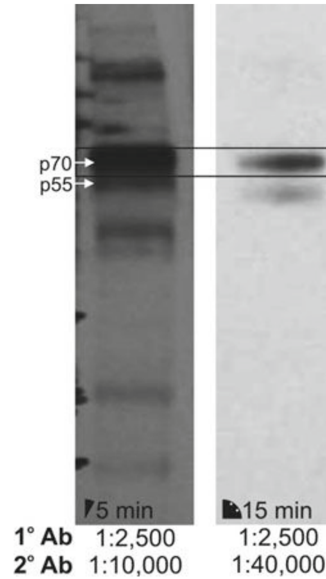


Fig. 3. Optimizing secondary antibody for improved Western blot detection. Purified (P) ChIL-12[HIS] resolved by SDS-PAG and immunoblotted with a mouse monoclonal anti-6×HIS primary antibody was tested at different secondary anti-mouse IgG-AP conjugated antibody to improve signal-noise ratio on Western blot analysis.

antibody, antiserum from the first bleed was tested at 1:2,000, 1:5,000 and 1:10,000 (Fig. 2). While all three dilutions detected the predicted 70 kDa (p70; full length protein) and 55 kDa (p55; cleavage product) bands in purified ChIL-12 lanes, only the lower 1:2,000 dilution of antiserum was capable of detecting this recombinant protein in crude protein extracts. Moreover the pre-immune serum from this same host rabbit showed no cross-reactivity with a band corresponding to either p70 or p55 in both crude and purified ChIL-12 samples further supporting the specificity of this polyclonal antiserum for our ChIL-12 recombinant protein. Subsequent bleeds (e.g., second and final) exhibited slightly improved but similar titers. Additional optimization is also recommended if the antiserum fractions are subsequently pooled and/or purified for establishing the standard assay parameters.

In addition to varying the primary antibody dilutions, the secondary antibody may also need to be optimized. In practice we have found that high background that does not appear to be markedly improved with primary antibody dilution may be associated with non-specific interactions of the secondary antibody. A case in point is shown in Fig. 3, where an anti-HIS tag antibody we use for detecting ChIL-12 exhibited high background at a relatively short exposure time (5 min). By only increasing the dilution of the AP-conjugated anti-mouse IgG secondary antibody fourfold

(1:10,000–1:40,000), a significant reduction in the membrane background was obtained even with longer film exposures. In addition, the ability to increase antibody dilutions (both primary and secondary) is desirable, especially when any of these reagents are limited (e.g., polyclonal antibody).

3.3.2. Some General "Tricks-of-the-Trade" for Optimizing Western Blots

In efforts to optimize the Western blotting conditions for a new antibody, a number of common problems may arise that are independent of adjusting the antibody concentrations. There are several vendor websites listed at the end of this chapter and other references (e.g., (13)) that provide useful troubleshooting tips for Western blotting analysis. Here are just a few "tricks-of-the-trade" our group has employed over the years to reduce or eliminate some of these commonly observed issues that can introduce an uncontrolled variable in determining optimal Western blot conditions for the new antibody targeting your recombinant protein:

1. High background noise:

   In our experience this has been one of the most notable issues with Western blotting using new antibodies whether newly produced or commercially purchased.

   (a) Purify the antibodies from the whole serum or yolk fraction to reduce overall non-specific protein that may attribute to Western blot background (see Subheading 3.2.3).

   (b) Blocking agent may be interacting with the new antibody. Try an alternative blocking agent and make sure the blocking agent used is of high quality. Commonly used blocking agents include non-fat dry milk, casein, BSA, or a commercial blocking solution available from a number of vendors (e.g., Pierce, Roche, KPL) (see Note 16).

   (c) Increase the amount of blocking agent (e.g., BSA) and/or detergent (e.g., Tween 20) used in the blocking step for some antibodies. We typically use 3% BSA in our blocking step; however, BSA concentration in blocking solution can range from 1 to 5% (see Note 15).

   (d) Include BSA in one or both antibody incubation steps to discourage non-specific binding of the antibodies. We frequently include 1% BSA in primary and secondary antibody incubation solutions.

   (e) Increase the relative amount of your target protein through enrichment or concentration; without the presence of a high-affinity target, antibodies as a rule tend to exhibit non-specific binding behaviors.

2. Multiple bands detected

   (a) Reduce the amount of secondary antibody used. We have used this strategy with success on numerous occasions

when optimizing conditions for a new primary antibody (whether a new uncharacterized antibody or one purchased commercially).

(b) Compare secondary antibodies bearing different detection mechanisms, i.e., biotin-based strategies or different enzyme conjugates (alkaline phosphatase versus horse radish peroxidase) for detection.

(c) Further purify antibody subpopulations from your antisera. In some cases, the immunogen may have contained very antigenic contaminants yielding an antibody population that recognizes both the target recombinant protein and the co-purifying "contaminating" endogenous protein(s). Elimination of the non-target antibodies can be facilitated by immunoprecipitation with host protein extracts from production runs lacking the transgene or using affinity columns where these proteins have been conjugated to the resin. Alternatively, we have had some success by spotting non-transgenic protein extracts onto a nitrocellulose membrane and incubating the dried membrane with the primary antibody solution to remove undesired non-specific antibodies (see (14)). The incubation solution now containing the unbound target antibodies is transferred and incubated with the Western blot membrane. It should be noted that many of these procedures can result in significant loss of the antibody.

(d) Multiple bands may represent "legitimate" cross-reactivity, that is, the new antibody may be recognizing multiple glycoforms, isoforms, endogenous homologs, or breakdown products of the targeted recombinant protein. Modification of the blotting procedure cannot fix this but further analyses may provide important insights as to the source of this complexity (see Note 17).

3. Expected band not observed

(a) Extend the development time of the western blot up to overnight to determine protein product detection.

(b) Address non-specific interactions (as discussed in points above) to eliminate competition for the antibody and thus render it capable of binding to its intended target.

(c) Extend the incubation times or increase the incubation temperature to 37°C to promote protein interaction. We do caution against extended incubations especially at room temperature and higher temperatures as this can promote unintended bacterial growth which can advance the appearance of non-specific background on your Western blot. If extending incubation step beyond a few hours consider incubating at 4°C and/or including sodium

azide (generally recommended at 0.02%) in the incubation buffer.

4. Patchy spotting pattern on the Western blot

   This problem can be associated with:

   (a) Bubbles becoming trapped between membrane and gel during the setup of the transblot sandwich cassette. To minimize this problem, carefully roll membrane onto gel surface and remove any potential bubbles using a roller or glass pipette rolling in one direction across the membrane surface.

   (b) Uneven distribution of the chemiluminescence or other developing reagents across the blotted membrane. To minimize this problem use a pipet to regulate the even distribution of the developing reagent across the entire membrane surface.

   (c) At the initiation of the developing period, air bubbles become trapped between the membrane surface with the reagent and the overlaid plastic page protector sheet (or plastic wrap). To avoid introducing bubbles, have the protector page cut to a size slightly larger than your membrane; carefully place it on the membrane starting at one edge and gently “rolling” the plastic cover on such that contact is initiated progressively and any bubbles are guided off of the membrane as the interface is established.

   (d) Incomplete removal of excess developing reagent prior to film exposure or the developing agent becoming trapped in the folds of the plastic sheet or wrap covering the membrane. Use a kimwipe to completely press out the reagent after the development time has elapsed.

   (e) Reused plastic page protector sheets for developing Western blot. Use a new plastic page protector (cut to a size just larger than the membrane) for each blot.

5. Diffuse bands

   This problem may be a feature associated with structure or post-translational modifications of the recombinant protein itself and cannot be avoided. For example, glycan-modified proteins can have a “fuzzy” appearance on Western blot analysis. However, other issues that may result in this problem might be addressed by:

   (a) Decreasing the amount of protein loaded per lane. Don’t exceed suggested well volume and protein concentration given by the gel manufacturer.

   (b) Avoid incubating gel for an extended period in transfer buffer before assembling the unit and engaging the electrotransfer process.

(c) Reducing the distance that your target protein moves by decreasing electrophoresis running time or adjusting acrylamide concentrations to optimize resolution in the size range of your protein.

3.4. Optimizing Antibodies for Use in ELISAs

3.4.1. Choosing Your ELISA Format

The ELISA is a widely used application for analyzing soluble antigens (15). As a general rule there are three common features that all ELISA formats share: (1) one of the reactants (capture component) is bound to the solid surface (usually a 96-well plastic microtiter plate), (2) as other reactants are subsequently added to the system, those molecules with affinity to the capture component become "bound" and the "free" reactants are eliminated with simple washing steps, and (3) the "bound" components are ultimately detected based on a shift from a colorless (or inactive) substrate to a colored (or active) product that indirectly measures the presence and amount of the target analyte (16). In the context of recombinant protein production, ELISAs can provide quantitative, structural or functional information about a target protein depending on the binding specifications of the antibodies used. We have designed many ELISAs in a number of configurations for characterizing several of our plant-expressed proteins (e.g., (2, 17–19)). Some of these more common configurations including direct, indirect, sandwich, and indirect sandwich ELISAs are illustrated in Fig. 4.

If a robust ELISA (e.g., a commercial ELISA kit) is not available for your recombinant protein, consider designing your own assay either using commercially available antibodies that detect your protein or producing your own antibodies (see Subheading 3.2). When designing and optimizing an ELISA, there are a number of features to consider in validating the assay:

1. Linearity of the assay using a serial dilution curve of the protein standard (e.g., purified and validated ChIL-12^{HIS} protein or HSP-90^{HIS} protein standard).
2. Sensitivity of the assay by replicating the lowest detectable standard concentration value relative to the background sample (e.g., pBK plant extract).
3. Specificity of the assay showing no cross-reactivity to negative controls (e.g., pBK crude/parallel purified sample and/or an unrelated recombinant protein expressed in the same bioproduction system).
4. Spike recovery determined by running a serial dilution of standard in the biological matrix (e.g., sample background as in our model pBK crude/parallel purified sample) and obtaining predictable linearity in the concentration curve. If antibodies are specific, such matrix interaction more often results in underestimating the amount of your recombinant protein in the samples.

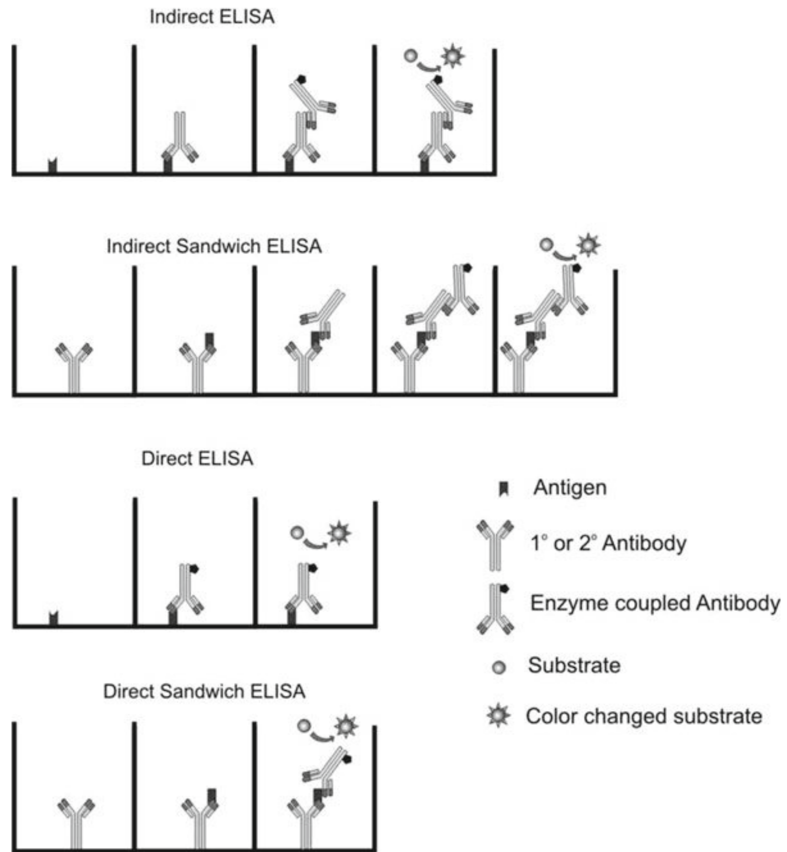


Fig. 4. Commonly used enzyme-linked immunoabsorbant assay (ELISA) formats for characterizing and assessing a recombinant protein of interst.

5. Intra-assay reproducibility for testing repeatability of the absorbance values obtained for the same sample both within a given plate and on different plates; typically we run three experimental replicates for all unknown samples and standards.
6. Inter-assay variation is identified by establishing reproducibility of ELISA standard and sample absorbance values on different plates, run on multiple days and by different technicians.

A measurable parameter in these tests for establishing an acceptable performance rating of the ELISA is the coefficient of variation (CV). The CV is calculated as the standard deviation between samples relative to the mean and is expressed as a percentage of variance to the mean. As detailed discussion of this analysis and other calculations are beyond the scope of this chapter, we refer the reader to more comprehensive ELISA methods development

resources (e.g., (16, 20)). However, in general, for linearity, spike and intra-assay and inter-assay tests, CVs of 80–120% are regarded as acceptable, with values at 100% indicative of optimal performance.

An excellent resource if you need to develop or troubleshoot an ELISA is "*The ELISA Guidebook*" (16), which provides detailed methodology and practical discussion of the topic. In addition, we have listed weblinks at the end of this chapter with useful troubleshooting and protocols published by a number of companies that sell ELISA kits and reagents. In order to provide real-life examples, we have chosen to highlight two ELISAs recently developed for the assessment of our recombinant ChIL-12 protein. Within the context of these examples, we will focus on the overall strategy used and significant aspects in their development that may be broadly applicable to designing an ELISA for your protein of interest.

3.4.2. Developing a Direct ELISA for HIS-Tagged Recombinant Protein Quantification

In expressing a protein for which there is no previously characterized or available standard, the ability to establish an assay that provides quantification for the recombinant protein is critical and a common challenge in the field. For our model recombinant protein ChIL-12, a 6×-histidine tag was engineered at the carboxy terminus that not only facilitated purification (2) but with the commercial availability of anti-HIS tag antibodies, provided a means for determining the quantity of this recombinant expressed protein. While we have used both anti-6×HIS and antiFeIL-12 Western blots in combination with densitometry to estimate the amount of plant-derived ChIL-12, significant limitations including the limited linear range of X-ray film used for detection and the restricted sample lanes on an SDS-PAGE are associated with this gel-based approach. A direct ELISA (see Fig. 4) using a anti-HIS tag antibody can provide a far more sensitive, accurate, simplified, and higher throughput assay for recombinant protein quantification.

To design a direct ELISA using anti-HIS antibodies to quantify your HIS-tagged protein, you must utilize a well-characterized HIS-tagged standard. This is preferably a protein with a similar mass to the target protein and of a known concentration (established and validated in accordance with industry standards), which permits you to effectively establish a relative protein concentration value of a HIS-tagged recombinant protein. We chose the recombinant and commercially available heat shock protein 90 protein expressed with a carboxy-terminal histidine tag (HSP-90^{HIS}) to be used as a relative standard. However, it should be noted that any protein: (1) with a histidine tag matching that of the recombinant protein (e.g., carboxy-terminal 6×-HIS; amino-terminal 10×-HIS; etc.), (2) whose concentration is quantifiable by a standard method and (3) from a sustainable source (e.g., commercially available for

enabling ready access to the reagent) can be used in designing a HIS-tag protein direct ELISA. While the detailed method for this direct ELISA is described in a later chapter (see Chapter 29, Subheading 3.5.3), the following summarizes a checkerboard titration (CBT) strategy used in designing this anti-His tag direct ELISA as a reliable quantification tool for the ChIL-12 protein:

1. Set up a twofold serial dilution of the HSP-90^{HIS} protein (calibrated standard) in PBS (100 μL/well). The dilution scheme initiated in lane 1 with 1 μg/mL of the protein is illustrated in Fig. 5 (see Note 18). Therefore each column has a

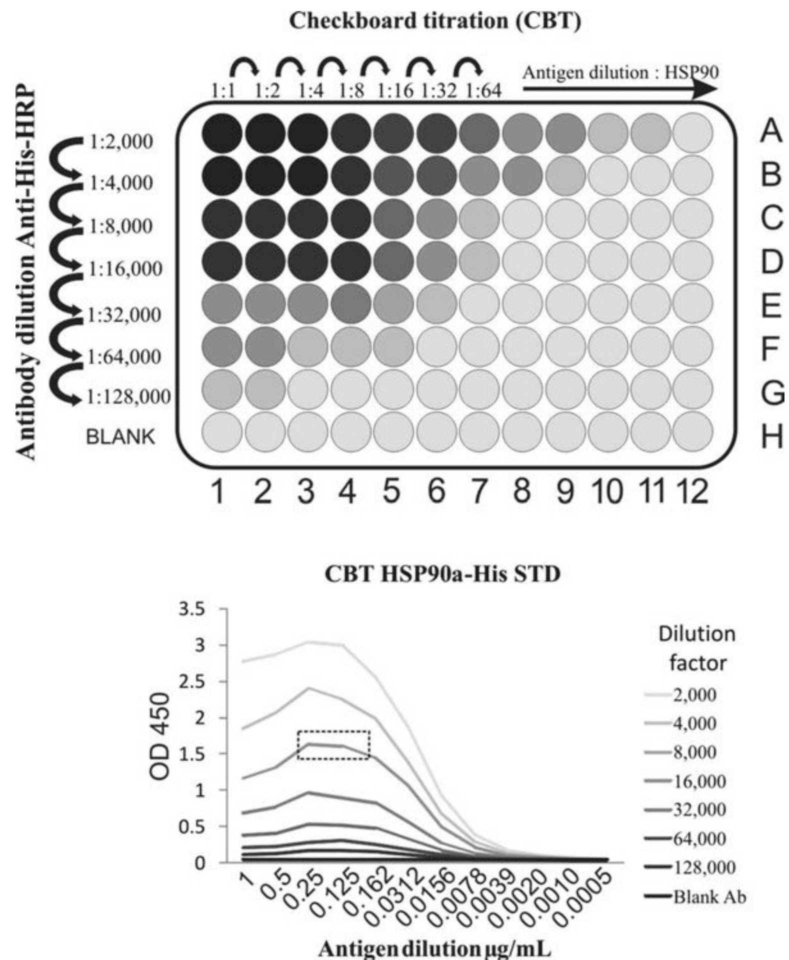


Fig. 5. Strategy using a checkerboard titration test to optimize an anti-HIS tag protein direct ELISA. Serial dilutions of both a HIS-tagged heat shock protein 90 (HSP90) used as a standard HIS-tag protein and a mouse monoclonal anti-6×HIS alkaline phosphatase-conjugated antibody were plated in a checkerboard titration test (*top*) to determine the appropriate antibody dilution (*bottom*) in designing a direct ELISA for quantifying recombinant ChIL-12^{HIS} protein as described in Subheading 3.4.2.

different but predicted concentration of standard protein. Note the last row (H) of each column on the plate is reserved for controls including wells run with no protein as shown.

2. Process ELISA as outlined in Chapter 29 Subheading 3.5.5 of this book to bind protein HSP-90^{HIS} standard protein samples to the wells of the plate, wash to remove excess, unbound protein and block with BSA blocking solution to bind any unoccupied surface of the well.
3. Set up a similar twofold serial dilution range for the HRP-conjugated anti-6×HIS detection antibody beginning at 1:2,000 from row A and on to row G (depicted in Fig. 5); incubate to allow binding (see Note 19).
4. Following a wash step, add chromophore/substrate and allow color to develop.
5. Add stop solution as detailed in Chapter 29 Subheading 3.5.5 to each well and determine the absorbance using a plate spectrophotometer at 450 (see Note 20).

Based on this initial CBT analysis, a best fit curve of the HSP-90^{HIS} serial dilution was determined by plotting standard concentration versus absorbance values and analyzed using Microsoft Excel (or another comparable computer program). As a general rule the most accurate ELISAs have a plateau maximum of ~1.7 O.D and the antigen can be titrated down to background levels as described in further detail in ref. 16. Therefore, based on the antigen titration curves shown in Fig. 5 (dotted line designates targeted plateau level), an antibody dilution of 1:8,000 is predicted to provide the best antibody dilution for the direct ELISA.

With the linear range of this direct ELISA established, serial dilutions of plant-expressed ChIL-12 (e.g., ChIL-12 purified sample) and “mock” plant controls lacking the transgene (pBK) are run in this direct ELISA. For example, a typical direct ELISA analysis might contain the standard curve (minimum of five points in duplicate), a dilution series of your mock control (done in triplicate), and dilution series of the samples to be quantified (done in triplicate). The concentration values for recombinant ChIL-12^{HIS} are interpolated from the HSP-90 standard curve and expressed as a relative concentration to HSP-90^{HIS}. Due to cross-reactivity with competing histidine-containing proteins in plant extracts, we have found this direct ELISA to be most applicable and reproducible with more purified ChIL-12 protein fractions. This direct ELISA along with the one described below are fundamental tools in establishing a quantitative quality assessment program for our recombinant ChIL-12 as outlined in Chapter 29 of this book (see Note 20).

3.4.3. Design Strategy Used in Developing a Sandwich ELISA to Detect ChIL-12

Among the most useful ELISAs for recombinant protein assessment is a two antibody "sandwich" ELISA (Fig. 4). This configuration generally offers improved specificity over a direct ELISA and can determine the absolute amount of the recombinant protein in a sample if a purified protein standard is available. To design a sandwich ELISA, two or more antibodies that target non-overlapping epitopes on the recombinant protein need to be available (see Note 21). If the two antibodies are from the same type of production host (e.g., paired IgG monoclonal antibodies produced in mouse) at least one of the antibodies needs to be linked with a reporter enzyme to enable detection. However, in considering the optimization of an ELISA, this requirement either limits the number of testable antibody configurations (as only one of the antibodies is modified for detection) or becomes a more costly pre-test (as multiple antibodies would need to be conjugated for assessing each in the context of a capture and detection antibody). Conversely, if the antibody pairs being tested were generated in different types of production hosts, an indirect approach (Fig. 4, "indirect sandwich" ELISA) is a commonly used configuration that is especially useful when initially designing a new sandwich ELISA. In this case, all antibodies that target the recombinant protein can be tested in both orientations in the "sandwich" (e.g., binding antigen either in the capture orientation or detection orientation). A third antibody, conjugated to the reporter enzyme that binds the constant region (Fc) of the secondary antibody heavy chain, enables recordable detection. While the use of a third antibody does require an additional incubation and wash steps and extends the assay time, there are many commercially available options for such modified antibodies and they can be used in optimizing additional sandwich ELISAs targeting other proteins. Upon determining the best "matched pair" and their relative orientation in the sandwich configuration, the antibody positioned for detection can subsequently be modified with a reporter enzyme to streamline the assay into a direct sandwich ELISA (see Fig. 4).

One of the initial ChIL-12-specific ELISAs we developed was a sandwich ELISA using our anti-ChIL-12 (αChIL-12) polyclonal antibody. At this point in the project, in addition to our αChIL-12p70 antibody produced in rabbit, we had two other antibodies that could detect ChIL-12, an antibody against the feline IL-12 homolog and an antibody detecting the C-terminal histidine tag of ChIL-12^{HIS}. Both of these antibodies had been shown to detect recombinant ChIL-12^{HIS} by western blot analyses (2). Figure 6 illustrates the overall summary of our initial checkerboard titration test in establishing an αChIL-12 sandwich ELISA. A similar checkerboard titration test used in designing the direct ELISA described above is applicable although a total of four reagents require optimization in an indirect sandwich (compared to two for a direct ELISA). As only two components can be varied by dilution in a

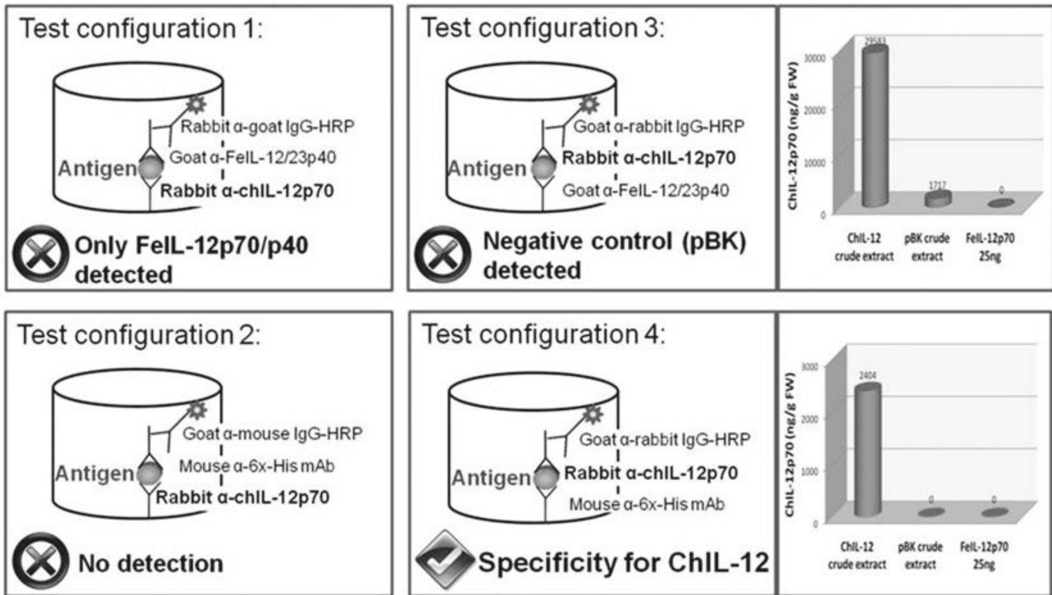


Fig. 6. Designing an anti-ChIL-12 sandwich ELISA. The anti-ChIL-12 polyclonal antibody was challenged in a checkerboard titration test paired with an anti-feline IL-12.23 p40 antibody (αFeIL-12p40) or anti-HIS tag monoclonal antibody (α6×HIS) and detected with the appropriate conjugated tertiary antibody. Four test configurations (TC) were used in establishing the optimal ChIL-12 sandwich ELISA; ELISA results for the two configurations where ChIL-12^{HIS} was detected is shown and further described in Subheading 3.4.3.

given test, initially the dilutions of the tertiary antibodies were held constant and at levels suggested by the vendor (see Subheading 2.4) as well as one of the antibody sets that bind the antigen (the two commercial antibodies, αFeIL-12 and α6×-HIS). In summary, an antibody test matrix with a serial dilution of the new αChIL-12p70 antibody tested first as capture antibody was paired with each of the commercial antibodies (αFeIL-12 and α6×-HIS) tested as a detection antibody each at their respective dilutions. A second antibody matrix tested the reverse orientations with the αChIL-12p70 used as detection and each of the two commercial antibodies used for capture of the antigen. All antibody test matrices were evaluated across a range of antigen (e.g., serial dilution of purified ChIL-12) with the goal of determining the best performing configuration for the sandwich ELISA.

Based on a series of these checkerboard titration tests, it was shown that when the ChIL-12 is captured with the anti-6×-HIS antibody (binds to the C-terminal HIS-tag on the p40 subunit), our αChIL-12p70 polyclonal antisera can detect the protein [see Fig. 6, Test Configuration 4 (TC4)] with good linearity, selectivity, and sensitivity. This configuration appears to be specific for the chicken homolog of IL-12 as neither the FeIL-12 heterodimer or p40 subunit can be detected. In considering the fact that the

FeIL-12 is not detected in TC2 or TC3 as well, this might suggest that the ChIL-12 protein epitopes to which the highest affinity antibodies in the αChIL-12 antisera bind are likely common to high affinity antibodies in the αFeIL-12 antisera (Fig. 6, TC3). Likewise, a sub-population of antibody(s) in the αChIL-12 antisera with the highest affinity for the protein in fact blocks binding to the p40 C-terminal HIS-tag of the recombinant ChIL-12 (Fig. 6, TC2). Furthermore, as TC1 (see Fig. 6) exclusively detects the feline p70 heterodimer and p40 subunit suggest the epitopes the αFeIL-12 antisera recognize on the ChIL-12 protein are common to or hindered by the antibodies in the αChIL-12 antiserum and a subpopulation of antibodies in αFeIL-12 are specific for the feline homolog. Taken together, this example highlights that the inherent heterogeneity of most polyclonal antibodies makes interpretation of such results complex and tedious and elevates the importance of integrating the proper controls into such analyses. This anti-ChIL-12 checkerboard titration test provided the basis for optimizing the first described ChIL-12 ELISA that recognizes a HIS-tagged ChIL-12 protein. This ELISA is an important tool for establishing a quality assessment program for plant-based production of ChIL-12 as discussed in Chapter 29 of this book. Efforts are ongoing to design a quantitative αChIL-12 ELISA that incorporates our ChIL-12 monoclonal antibodies (Subheading 3.2.2) for a more sustainable and universal (e.g., detect native ChIL-12) ELISA with value to the poultry industry and academic R&D sectors.

3.5. Characterization of Antibodies for Additional Applications

High-affinity selective antibodies to your protein-of-interest can be broadly useful to many additional applications that may include assessing function, monitoring protein uptake and trafficking, identifying cellular and subcellular sites of accumulation, or purifying your proteins. While an in depth discussion of these applications is beyond the scope of this chapter, it may be useful to mention some of these approaches and the characteristics of the immuno-tools that may be important for a particular application. More detailed discussions of these applications are available in several reviews and methodology texts (6, 21–23). It should be noted that most of these applications are focused on recognizing the protein in a native conformation which impacts the choice of antigen/immunogen in developing and selecting your antibodies.

3.5.1. Using Antibodies for Functional Analyses

The specificity of some antibodies can provide a “window” into functionality of your protein of interest. Antibodies can be selected for a particular conformation or modification. One example, are antibodies developed to only bind a particular protein in the phosphorylated site thus providing insight into activation of a regulator or signal transduction pathway (e.g., the antibody against phospho-Akt(Ser473) only recognizes the Akt1 protein if it is

phosphorylated at Ser473). Typically antibodies with this level of site-specific selectivity are identified from a monoclonal antibody library. However, it should be noted that particular sub-populations of antibodies can be selected from a polyclonal antibody population. For example, if you are interested in having antibodies that only cross-reacts with the protein component of a glycoprotein, it is possible to enzymatically deglycosylate your recombinant protein and bind it to membranes or resin in order to select the peptide targeted sub-population of antibodies. Alternatively, unrelated glycoproteins can be conjugated to the chromatography resin such that your protein-specific antibodies "flow through" while those binding to the carbohydrate components bind to the resin and are removed.

"Neutralizing" and "activating" antibodies can be very useful for functional analyses (24). These are antibodies that specifically interfere with or promote the protein's function. In the case of neutralizing antibodies this may involve binding or interfering with the active site and thus disrupting enzymatic activity. However, the most common applications are in blocking protein binding to its effector sites. We have used this type of antibody with our plant-derived IL-12 to confirm functionality (18). We demonstrated that our purified IL-12 triggered the induction of interferon-gamma production in primary splenocytes and that this induction was blocked when the IL-12 was pre-incubated with anti-IL-12 neutralizing monoclonal antibody. This confirmed that the interferon-gamma inducing activity observed was directly due to IL-12 and not to other components that may have co-purified from the plant tissues. We have also used neutralizing antibodies to demonstrate that the carbohydrate binding activity of a plant lectin is the critical component mediating mammalian cell uptake of associated payload proteins. Thus, this class of antibodies can provide valuable mechanistic insights into the function of your recombinant protein.

In regards to producing neutralizing antibodies, they can be generated either using a monoclonal or polyclonal production platform. By definition, the "neutralization" aspect is based on interfering with function and thus the antibody is recognizing native epitopes (although it may also cross-react with denatured proteins). Therefore, it is advantageous to immunize with conformationally intact (active) protein when generating your antibody population. In the case of plant-expressed ChIL-12, the polyclonal antibody which was generated from intact soluble immunogen (see Subheading 3.2.2) has been shown to demonstrate neutralizing activity (2). This neutralizing antibody enabled validation that our ChIL-12 protein was in fact responsible for stimulating chicken splenocyte activation and confirmed the functionality of a plant-derived recombinant protein.

In some cases, "conformational antibodies" can be identified. These are antibodies that only recognize proteins when they are in

a particular structural state. An example of a conformational antibody our group has worked with is one available for the mouse IL-12 protein which only recognizes the heterodimeric p70 form of the protein (e.g., rat anti-mouse IL-12p70 clone 48110.111; R&D Systems; used in ref. 18). During the initial characterization and purification of mIL-12 from plants, this antibody provided us quick quantification of the assembled p35 and p40 subunits even in crude plant extracts and excluded detection of any non-function mIL-12 (e.g., p40, p35 or p40 homodimers). To date we have not identified a similar conformational antibody for our plant-expressed ChIL-12.

3.5.2. Microscopy and Immunodetection

Antibodies are also quite valuable for analyzing proteins in the context of cells. Protein-specific antibodies have been very powerful tools for localization of proteins in tissues, cells and organelles using fluorescence or electron microscopy. Typically the antibodies are labeled with either gold particles for electron microscopy or fluors for fluorescent or confocal microscopy and incubated with the target cells or tissues (that may be permeabilized and/or fixed to facilitate antibody access). Again, it should be noted that antibodies that effectively recognize native proteins are needed for such assays. In addition, the specificity of your antibodies is critical for these applications. Because all antibodies in your population will be labeled, if 10% of your antibody cross-reacts non-specifically to a different protein which is highly abundant in your cells or tissue, the potential for generating misleading data and incorrect interpretation is quite likely. For this reason, more extensive clean-up of your antibodies and identifying and exploiting appropriate controls is critical. Similar requirements are also required of the antibodies used in assays that fractionate populations of cells into different groups [e.g., fluorescence activated cell sorting; magnetic activated cell sorting (MACs)] for using select cell subpopulations in follow-on analyses and in assays that quantitatively determine the presence of select proteins on cells by a flow cytometry-based approach.

3.5.3. Affinity Purification of Proteins Based on Antibodies

Once you have identified antibodies specific to your proteins, these can also be developed as protein purification tools. In this application, the antibody is chemically conjugated to commercially available chromatography resins (or magnetic beads) yielding a highly selective affinity purification resin. Several commercial resins and kits (e.g., Pierce's AminoLink Plus Coupling Resin and Kits) are available for such applications. While antibody-based columns can be very useful, there are several caveats and issues that need to be considered. First, the selectivity of an antibody column directly reflects the quality and specificity of the antibody population. This is an important as contaminating proteins will be co-purified along with your target protein if your antibody cross-reacts with additional proteins. Second, antibodies have a tendency to leech off the column

during use. This limits the number of times a column can be reused (often permitting only ten runs) and column lifetime is further impacted by the method of protein elution which can be pH-based (harsh) or peptide-based (gentler). Elution of antibodies from the column can affect the purity of the target protein (antibody fragments may now contaminate your sample), as well as detection and quantification (your secondary antibodies in ELISAs or Westerns may cross-react with the antibody). Finally, if purifying a recombinant human protein, antibody-based purification is typically not supported for clinical applications based on guidance from the regulatory agencies to minimize any animal-sourced reagents in the production process of biologic drugs.

4. Electronic Resources

Antibody and antibody production:

- Great resource for general background about antibodies and troubleshooting antibody production issues and uses. Site also provides updated weblinks to numerous antibody producing facilities http://www.antibodyresource.com/.
- Cocalico Biologicals – used for producing our polyclonal antibody http://www.cocalicobiologicals.com/.
- Antibody Solutions – used for producing our monoclonal antibody http://www.antibody.com/.

Western blot analysis:

- Pierce-ThermoScientific "Optimize antigen and antibody concentrations for Western blots," Technical Resource, #TR0024.0 http://www.thermo.com/eThermo/CMA/PDFs/Articles/articlesFile_7172.pdf.
- Abcam "Western Blotting – A Beginner's Guide" and "Western blot Troubleshooting Tips" http://www.abcam.com/ps/pdf/protocols/WB-beginner.pdf http://www.abcam.com/index.html?pageconfig=resource&rid=11352.

ELISA:

- Pierce-ThermoScientific, "ELISA Technical Guide and Protocols," TechTip#65 http://www.piercenet.com/files/TR0065-ELISA-guide.pdf.
- R&D Systems "ELISA Development Guide," April, 2002 http://www.rndsystems.com/DAM_public/5670.pdf.
- Millpore ELISA Protocols http://www.millipore.com/immunodetection/id3/elisaprotocols.

Abcam "Troubleshooting Tips-ELISA" http://www.abcam.com/index.html?pageconfig=resource&rid=11390.

5. Notes

1. While we have typically run standard multi-welled gels for this procedure, the use of preparative, pre-cast SDS-PAGE gels also commercially available provide a valuable alternative format (one lane for molecular weight standard and one large well for sample). In practice, this reduces the amount of polyacylamide accompanying your antigen sample.
2. Unstained protein standards provide a better estimation of size than pre-stained standards. Stained standards are generally ideal for confirming the electrophoresis run and estimating the efficiency of transfer onto a membrane.
3. Other buffers including 6–8 M Urea, SDS or other detergents, distilled water, HEPES, or Tris that provide improved solubility and/or stability of your protein of interest may be used; consult the facility that will be producing your antibody to confirm that your buffer is compliant with their protocols.
4. Cocalico Biologicals accepts antigens supplied in a variety of forms including liquid, lyophilized, in polyacrylamide gels, as a precipitate or contained on nitrocellulose membranes. Check with the company producing your antibody for antigen forms accepted for immunizations.
5. In this context we define "low titer" as <1:1,000 dilution of the antisera from the second and/or final bleed containing the polyclonal antibody yet is capable of detecting the intended target on a Western immunoblot.
6. Several vendors sell Protein A purification kits that can simplify the IgG enrichment of your antibodies from serum (e.g., Sigma PURE1A; ThermoPierce #44667). In addition depending on the antibody sources, select subclasses or subtypes of antibodies can also be obtained by similar purification.
7. We typically purify only a small portion of a polyclonal antibody supply by Protein A chromatography at a given time. While the method detailed uses 1 mL of resin, this procedure is commonly scaled proportionately to 5 mL of Protein A sepharose.
8. Secondary enzyme reporter-conjugated antibodies we use includes: Bio-Rad goat anti-mouse IgG-AP (#1170-6520); Bio-Rad goat anti-rabbit IgG-AP (#170-6518); Abcam goat anti-rabbit IgG-HRP (#ab6721-1).

9. For plant-expressed recombinant proteins, we have found that AP-conjugated secondary antibodies provide better selectivity, particularly in crude plant extracts, due to the presence of plant endogenous peroxidases that interact with the substrate and may result in the appearance of additional, non-specific protein bands being detected.
10. A positive control such as the feline IL-12 protein standard (a species variant of ChIL-12) in this case, can be used to establish if the new antibody is recognizing epitopes common to both homologs.
11. If extra lanes are present on your SDS-PAGE gel, either skipping a lane between the sample series or running a dyed molecular weight marker between each set of samples will provide a convenient guide for cutting the membrane into strips.
12. It is best to minimize handling of the blotted membrane surface at this step which can contribute to increased background on the developed Western blot. When cutting the membrane into strips we often place it on a layer of two kimwipes or a paper towel, grasp the membrane with forceps and cut the membrane using either a clean scalpel blade or sharpened scissors.
13. An alternative to accelerate the Western immunoblot process and quickly test and optimize the use of new antibodies is to use the Millipore SnapID unit. This vacuum-based Western blot detection system allows up to six independent antibody conditions to be tested at a single time in about 35 min. To conduct a matrix test for new primary and secondary antibodies, the unit will need to be configured with two, triple well blot holders. Some important considerations in using this system include the standard buffer condition recommendations suggested by the manufacturer may deviate from your standard detection method; be aware that altering these condition when switching detection systems may impact the outcome of your Western blot. It is important to refer to the User Instructions in calculating the appropriate dilution of the antibody when using the multi-well blot holders so that antibody concentrations can be properly translated when running a standard Western analysis. Some technical considerations include: (1) when adding antibody use a transfer pipet to ensure complete coverage of the membrane surface; (2) an alternative to three times – 10 mL washes is a single 30 mL wash; (3) to pre-equilibrate the membrane for detection by vacuuming a small volume of detection buffer prior to removing membrane and proceeding with development step; (4) blot holders should be marked and not used more than ten times (or less depending on the antibodies).

14. If single lane membrane strips are being used they can be placed into the troughs of a mini strip eight channel tray (e.g., Bio-Rad immunological screening tray); incubations are conducted on a platform shaker at 50 rpm with 2 mL of solution.
15. There are multiple types of BSA in a standard chemical catalog that varies in purity, cost, extraction method, presence/absence of other albumin fractions. We in fact use a higher quality BSA for both Western blotting and ELISAs; Sigma BSA Cat #A7906; ≥98% (agarose gel electrophoresis) to reduce previous background noise issues.
16. While many protocols use non-fat dry milk as blocking agent, we have found much better consistency and improved Western blot background with BSA. In addition, we strongly encourage using the same vendor/product catalog number for your blocking agent of choice to mitigate possible background issues associated with different brands of product.
17. To determine if the additional bands comprise alternate glycoforms, deglycosylation experiments can be performed (see Chapter 29 of this book for more details). Cross-reactivity to endogenous homologs can be detected by comparing extracts from the production host with and without the target recombinant protein (see above). Lower molecular mass bands may reflect degradation products. This interpretation can be further supported if time course analyses show a temporal reduction of full-length product correlated with an increase in the smaller products. This can be confirmed by either N-terminal sequencing or mass spectrometry (see Chapter 29, Subheadings 3.4.1 and 3.4.2).
18. When establishing this direct ELISA assay, it is advisable to the run standard in triplicate; upon establishing this assay the HSP-90^{HIS} standard curve can be run in duplicate.
19. In addition to peroxidase-based reporter enzymes (e.g., HRP), alkaline phosphatase (AP) conjugated detection antibodies, while less sensitive, can also be used.
20. Substrates used with the reporter horse radish peroxidase (HRP) enzyme includes 2,2′-azo-bis(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS), o-phenylenediamine (OPD) and 3,3′5,5′-tetramethylbenzidine base (TMB), which yield green, orange and blue colors, respectively.
21. The amount of effort taken to optimize a sandwich ELISA warrants determining at the onset that the antibodies chosen and used in this ELISA assay are of sufficient supply and availability.

Acknowledgments

This work was supported by grants to C Cramer and M Dolan from the Arkansas Biosciences Institute at Arkansas State University and from USDA SBIR Phase I and II Grants awarded to BioStrategies LC (D Radin/C Cramer). A special thanks to Nora Rubio, Jorge Ayala, and Ross Fergus for technical support and expert troubleshooting skills on the projects supporting the development of the above technologies and "tools of the trade" for the ChIL-12 and RTB projects in our laboratory. Several of the troubleshooting and technical recommendations were contributed from immunological tool development supporting the seafood diagnostic program at Applied Food Technologies, Inc. (funding through several USDA SBIRs/PD: L Applewhite).

References

1. Terpe, K. (2003) Overview of tag protein fusions:from molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.* **60**, 523–533.
2. Medrano, G., Dolan, M.C., Stephens, N.T., et al., (2010) Efficient plant-based production of chicken interleukin-12 yields a strong immunostimulatory cytokine. J. *Interferon Cytokine Res.* **30**, 143–54.
3. Degen, W.G., van Daal, N., van Zuilekom, H.I., et al., (2004) Identification and molecular cloning of functional chicken IL-12. *J. Immunol.* **172**, 4371–4380.
4. Sheibani, N., Frazier, W.A. (1998) Direct use of synthetic peptides for antiserum production. *Biotechniques* **25**, 28–32.
5. Grant, G.A. (2002) Synthetic peptides for production of antibodies that recognize intact proteins. *Curr. Protoc. Immunol.* 9.2–9.2.18.
6. Burry, R.W. (2009) Immunocytochemistry: A Practical Guide for Biomedical Research. Springer-Verlag. New York.
7. Yokoyama, W.M. (2008) Production of Monoclonal Antibody Supernatant and Ascites Fluid. *Curr. Protoc. Mol. Biol.* **83**, 11.10.1–11.10.10.
8. Cooper, H.M., Paterson, Y. (2008) Production of Polyclonal Antisera. *Curr. Protoc. Immunol.* **82**, 2.4.1–2.4.10.
9. Landi, M. (ed.) (1995) Adjuvants and Antibody Production. *ILAR J.* **37**, 92–152.
10. Stills, J.F. Jr. (2005) Adjuvants and antibody production: Dispelling the myths associated with Freund's complete and other adjuvants. *ILAR J.* **46**, 280–293.
11. Gassmann, M., Thommes, P., Weiser, T., et al., (1990) Efficient production of chicken egg yolk antibodies against a conserved mammalian protein. *FASEB J.* **4**, 2528–2532.
12. Harlow, E., Lane, D. eds. (1988) Antibodies: A Laboratory Manual. Cold Springs Harbor Press, Cold Springs Harbor, NY.
13. Alegria-Schaffer, A., Lodge, A., Vattem, K. (2009) Performing and optimizing Western blots with an emphasis on chemiluminescent detection. *Methods Enzymol.* **463**, 573–599.
14. Nature Methods (2005) Immunoblot affinity purification. *Nat. Methods* **2**, 797–798.
15. Hornbeck, P. (1991) Enzyme-linked immunosorbent assays. In *Curr. Protoc. Immunol.* (Coligan, J.E., ed.), pp. 2.1.2–2.1.22. John Wiley & Sons, Hoboken, N.J.
16. Crowther, J.R. (ed). (2001) ELISA Guidebook. Humana Press. Totowa, New Jersey.
17. Reed, D.G., Nopo-Olazabal, L.H., Funk, V., et al., (2005) Expression of functional hexahistidine-tagged ricin B in tobacco. *Plant Cell Rep* **24**, 15–24.
18. Liu, J., Dolan, M.C., Reidy, M., et al., (2008) Expression of bioactive single-chain murine IL-12 in transgenic plants. *J. Interferon Cytokine Res* **28**, 381–392.
19. Medrano, G., Reidy, M., Liu, J., et al., (2009) Rapid system for evaluating bioproduction capacity of complex pharmaceutical proteins in plants. In: Faye L, and Gomord V (eds). *Methods Mol Biol.* **483**, 51–67. Humana Press, Totowa, NJ.
20. Plested, J.S., Coull, P.A., Gidney, M.A. (2003) ELISA. *Methods Mol Med.* **71**, 243–261.

21. Howard, G.C., Kaser, M.R. eds. (2007) Making and using antibodies: a practical handbook, CRC Press. New York, NY.
22. Janeway, C.A., Jr., Travers, P., Walport, M., et al., eds. (2007) Immunobiology 7th edition, Garland Science Publishing. London, England.
23. Howard, G.C., Bethell, D.R. eds. (2001) Basic methods in antibody production and characterization, CRC Press, New York, NY.
24. Lipman, N.S., Jackson, L.R., Trudel, L.J., et al., (2005) Monoclonal versus polyclonal antibodies: Distinguishing characteristics, applications, and information. *ILAR J* **46**, 258–268.

Part II

Prokaryotes

Chapter 5

Heat-Shock Protein Fusion Vectors for Improved Expression of Soluble Recombinant Proteins in *Escherichia coli*

Christos A. Kyratsous and Christos A. Panagiotidis

Abstract

Molecular chaperones or heat-shock proteins (HSPs) are protein machines that interact with unfolded or partially folded polypeptides and assist them in attaining their proper conformation. The folding reaction relies on a complex array of scaffolding effects and ATP-driven conformational changes that mediate the temporary unfolding and subsequent refolding of protein substrates. DnaK and GroEL are the two major *Escherichia coli* chaperones. They belong to the HSP70 and HSP60 families of proteins, respectively, and play a major role in protein folding. Here, we describe a set of bacterial expression vectors that permits the fusion of a protein of interest to DnaK or GroEL and its subsequent quantitative expression in a soluble, easily purifiable form. We also provide a set of compatible co-chaperone expression constructs that permit the simultaneous co-expression of the DnaK and GroEL physiological partners to further increase protein solubility. The system was successfully tested using the murine prion protein (PrP). Although PrP is normally insoluble when expressed in *E. coli*, we show that utilizing our vectors it can be produced in a soluble form as a DnaK or GroEL fusion. This system is useful for the production of a large array of proteins that fail to fold properly when expressed in *E. coli*.

Key words: Chaperones, DnaK, GroEL, Fusion proteins, Protein folding

1. Introduction

The ever-growing need for vast quantities of specific proteins in biotechnological or pharmaceutical applications cannot be met solely through their isolation from natural sources. Modern biotechnology has developed a large number of recombinant protein expression systems that allow high-level protein expression in a wide variety of host organisms. Of these, the undisputed champion remains the enterobacterium *E. coli*. The popularity of this lowly bacterium as

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_5, © Springer Science+Business Media, LLC 2012

a cell factory for the production of recombinant proteins has much to do with its ease of culture, its rapid growth under relatively low-cost conditions, the extensive knowledge of its genetics, and the abundance of available protein expression systems optimized for this organism (8–10, 14, 17).

The overproduction of recombinant proteins in *E. coli* is not always a straightforward process. Protein accumulation can be severely reduced by bacterial proteases (14). Additionally, the differential codon utilization of mammals and bacteria can result in the translational stalling of many mammalian genes expressed in *E. coli* and this may subsequently lead to premature transcriptional termination. However, the major factor that hinders the value of this bacterium as an expression host is the inability of many overproduced heterologous proteins to fold properly (14), often leading either to their accumulation within compact intracellular structures termed inclusion bodies or to their degradation by bacterial proteolytic systems (1, 4, 23). This inability to fold properly could stem from the fact that the synthesis of massive amounts of recombinant proteins overwhelms the cell's protein-folding systems (11). Without the aid of such systems recombinant proteins, more often than not, fail to fold properly (6, 20). Such misfolded polypeptides are rather prone to aggregation due to interactions of their exposed hydrophobic surfaces with other hydrophobic components within the very crowded intracellular environment (1, 4, 23).

To circumvent this problem of limiting amounts of protein-folding chaperone proteins (6), a number of plasmid systems have been developed to allow co-expression of various combinations of chaperones together with the target proteins (3, 16, 18, 19, 21, 22, 25, 26). While such systems have improved the solubility for a significant number of previously insoluble bacterially expressed proteins, their success in protein folding is far from universal. Thus, a need for even more efficient refolding systems for proteins overproduced in *E. coli* remains.

Here, we describe the use of a set of plasmid vectors that allow for the expression of recombinant proteins as cleavable N-terminal fusions with the two major chaperones (HSPs) of *E. coli*, i.e., DnaK (Hsp70) and GroEL (Hsp60). We have shown (13) that gene fusions between a target-protein and an HSP can improve target-protein folding, possibly due to an increase in the HSP local concentration within the proximity of the folding area. Since the activity of these chaperones requires the presence of co-chaperone proteins, i.e., DnaJ and GrpE for DnaK and GroES for GroEL, we also describe the effects of co-expression of these co-chaperones on the solubility of the recombinant HSP-fusions.

2. Materials

2.1. Bacterial Strains and Media

All strains are grown at 37°C, either on 2XYT agar plates or with aeration in 2XYT broth, supplemented with 100 μg/mL ampicillin, 50 μg/mL kanamycin, and/or 25 μg/mL chloramphenicol, as necessary.

1. *E. coli* strains: *Top10* (F^- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *deo*R *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G λ^-) and *BL21(DE3)* (F^- *omp*T *hsd*S_B ($r_B^- m_B^-$) *gal dcm* (DE3)).
2. Media: 2XYT Broth is prepared by dissolving 16 g tryptone, 10 g yeast extract, and 5 g of NaCl in 900 mL deionized H_2O, adjusting the pH to 7.2 with 5N NaOH, bringing the volume to 1 L with distilled H_2O and sterilizing by autoclaving at 121°C for 20 min. Solid 2XYT/agar for plates is made by adding 16 g of agar in 1 L of 2XYT prior to the autoclaving step.
3. 1,000× Ampicillin stock: 100 mg/mL Ampicillin sodium salt in H_2O, filter-sterilized with 0.22 μM pore-size syringe filtration device, and stored at −20°C.
4. 1,000× Kanamycin stock: 50 mg/mL Kanamycin in H_2O, filter-sterilized with 0.22 μM pore-size syringe filtration device, and stored at −20°C.
5. 1,000× Chloramphenicol stock: 25 mg/mL Chloramphenicol in ethanol; filter-sterilized with 0.22 μM pore-size syringe filtration device, and stored at −20°C.
6. Isopropylthiogalactoside (IPTG) is dissolved in distilled H_2O at 1 M, filter-sterilized with 0.22 μM pore-size syringe filtration device, and stored at −20°C.

2.2. Plasmids, Plasmid Construction, and Cloning

1. Tris–EDTA (TE) buffer: 10 mM Tris–HCl, pH 8, and, 1 mM EDTA.
2. 5× TBE buffer: Dissolve 54 g Tris base and 27.5 g boric acid in 900 mL H_2O, add 20 mL 0.5 M EDTA, pH 8.0 and bring the final volume of the solution to 1 L.
3. 0.5× TBE buffer: Dilute 5× TBE tenfold with H_2O.
4. Ethidium bromide (EtBr) solution: Prepared as a 10-mg/mL stock solution in H_2O and kept at room temperature (RT), protected from the light. Extreme care must be taken in handling and disposing of EtBr, as it is mutagenic.
5. SEVAG: Prepared by mixing 24 volumes chloroform with 1 volume isoamyl alcohol. SEVAG is stored at RT in a glass bottle.
6. Phenol–SEVAG: Commercially available equilibrated phenol stabilized with 0.1% (w/v) 8-hydroxyquinoline.

7. TA-cloning derivative of pZErO 2.1: Prepare 20 μL aliquots at a solution at a final concentration of 50 ng/μL and store at −80°C (see Note 1).
8. pET-21a(+) (Novagen) is used for the expression of control recombinant proteins carrying a carboxy-terminal hexahistidine (6×His) tag.
9. The chaperone-fusion vectors pXCK-K and pXCK-EL (13) allow for the expression of recombinant proteins carrying an HSP fused at their amino-terminus (DnaK for pXCK-K and GroEL for pXCK-EL) and a 6×His tag fused at their carboxy-terminus, as shown in Fig. 1.
10. The co-chaperone expression vectors pXCK-E/J and pXCK-ES (13) express the DnaK and GroEL co-chaperones DnaJ/GrpE and GroES, respectively, under the control of the T7 promoter, as shown in Fig. 2.
11. DNA modifying enzymes: *Not*I, *Nhe*I, *Hind*III, and *Eco*RV restriction enzymes, Taq polymerase and T4 DNA ligase (along with their appropriate assay buffers) are commercially available.
12. Low-melt agarose gels (2% w/v): To prepare a single agarose mini-gel (8 × 10 cm) in a horizontal gel apparatus, 1 g of low-melt agarose is added to 50 mL 0.5× TBE containing 0.3 μg/mL EtBr and heated sufficiently in a microwave oven, with intermittent swirling, to dissolve the agarose. The hot agarose is cooled to about 60°C before the solution is poured into the sealed, clean gel mold, a comb is inserted and it is left to harden (at least 30 min at 4°C). The gel is then removed from the mold, placed in the electrophoresis tank and a sufficient volume 0.5× TBE/EtBr buffer is poured into the tank to cover the gel. The samples are loaded and the gel is run at a constant voltage of 60 V to separate the DNA bands of interest.

2.3. Cell Lysis, Protein Purification, and Protein Solubility Assays

1. Phosphate-buffered saline (PBS): 1 mM KH_2PO_4, 10 mM Na_2HPO_4, 137 mM NaCl, 2.7 mM KCl, pH 7.4. PBS is sterilized by autoclaving and stored at RT.
2. Cell lysis solution: PBS supplemented with 10 mM β-mercaptoethanol, 1 mM EDTA, 0.2 mg/mL lysozyme, and Complete™ (Roche) protease inhibitor cocktail. Prepare immediately before use and keep at 4°C.

Fig. 1. (continued) 6×His tag and the thrombin cleavage site, (**c**) Schematic diagram of the expressed HSP-fusion proteins including the locations of the thrombin cleavage site and the C-terminal 6×His tag, (**d**) Expression of chaperones and chaperone-PrP fusions. Extracts from control cells [pET-21a (+)] or from cells expressing PrP-6×His (pX-PrP), DnaK-6×His or GroEL-6×His (pXCK-K and pXCK-EL) and the DnaK-PrP-6×His or GroEL-PrP-6×His fusions (pX-DnaK-PrP and pX-GroEL-PrP) are shown. Equal extract amounts were analyzed on 10% SDS-PAGE gels, which were either stained with Coomassie Brilliant Blue R250 (total protein staining), or electrotransferred to a nitrocellulose membrane and probed with an antipolyhistidine monoclonal antibody (Western blot analysis). The positions of the relevant proteins are indicated by *arrows.*

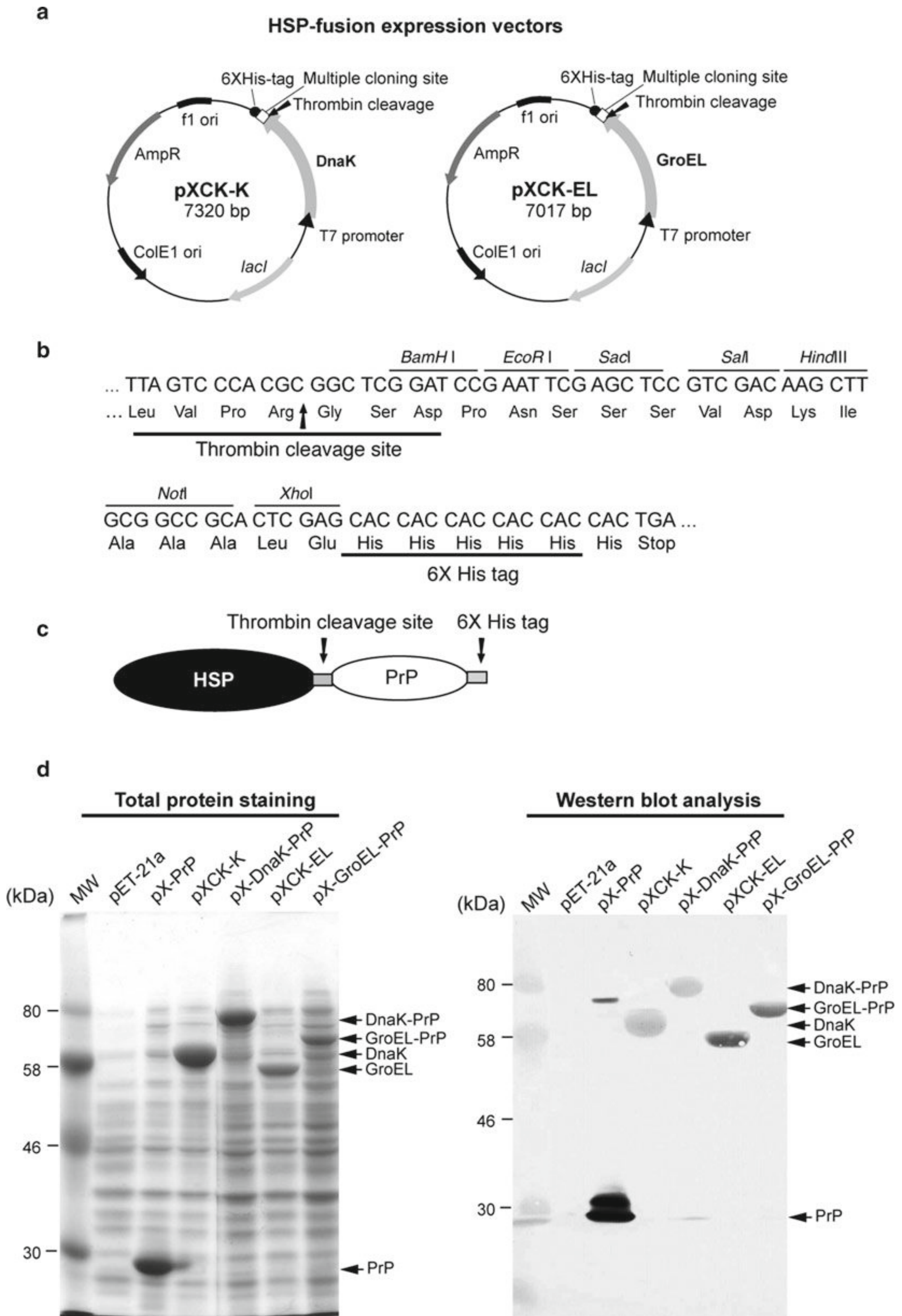


Fig. 1. Cloning and expression of chaperone fusions. (**a**) Maps of the DnaK and GroEL fusion vectors, (**b**) Sequence of the multiple cloning site of the HSP-fusion vectors, showing the translational reading frame and indicating the positions for the C-terminal

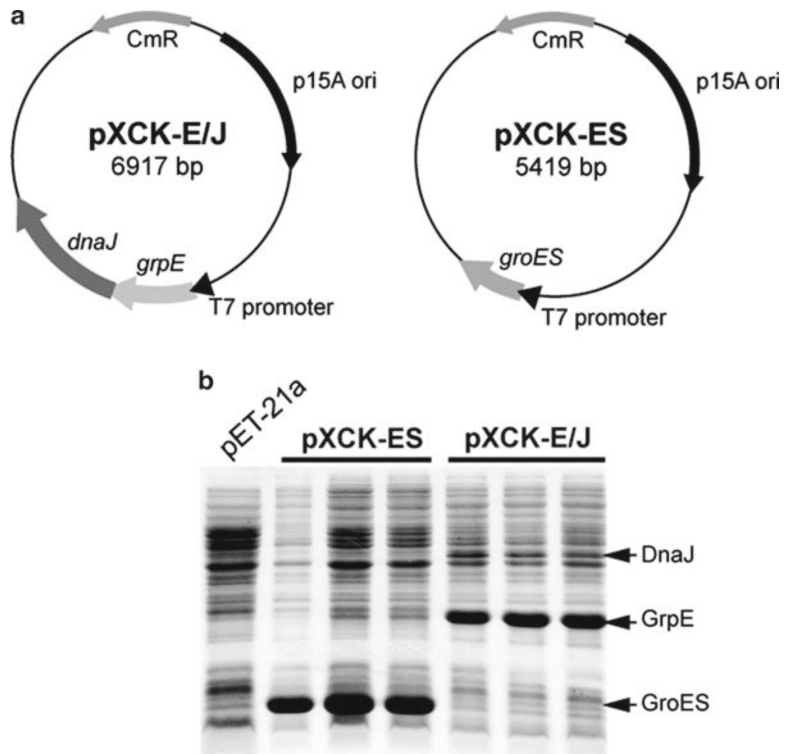


Fig. 2. The co-chaperone expression vectors. (**a**) Maps of the vectors expressing either GrpE and DnaJ (pXCK-E/J) or GroES (pXCK-ES), (**b**) Expression of co-chaperones. Extracts from cells carrying an empty vector (pET-21a(+)) or the co-chaperone expression vectors (pXCK-E/J and pXCK-ES) in triplicate were analyzed on a 15% SDS-PAGE gel and total proteins were visualized by Coomassie Brilliant Blue R250 staining. The *arrows* indicate the positions of the relevant proteins.

3. EDTA-free cell lysis solution: PBS supplemented with 10 mM β-mercaptoethanol, 5 mM imidazole, 0.2 mg/mL lysozyme, and Complete™ EDTA-free (Roche) protease inhibitor cocktail. This solution is prepared immediately before use and kept at 4°C.
4. 20% (w/v) Triton X-100: 20 g Triton X-100 is added to 75 mL distilled H_2O and the mixture is stirred until completely mixed. The volume is adjusted to 100 mL with distilled H_2O and the solution is stored at 4°C.
5. Ni-NTA agarose (e.g. Qiagen).
6. Wash buffer: 20 mM Tris–HCl pH 8.0, 150 mM NaCl, 20 mM imidazole, 1 mM PMSF, and 2 mM β-mercaptoethanol.
7. Elution buffer: 20 mM Tris–HCl pH 8.0, 150 mM NaCl, 300 mM imidazole, 1 mM PMSF, and 2 mM β-mercaptoethanol.
8. Dialysis buffer: 25 mM Hepes–NaOH pH 7.2, 100 mM NaCl, 0.5 mM dithiothreitol (DTT), 0.5 mM PMSF, and 0.05% (w/v) Tween-20.

9. Thrombin dilution buffer: 50 mM Tris–HCl pH 8.0, 150 mM NaCl, and 2.5 mM $CaCl_2$.
10. 10× Thrombin cleavage buffer: 500 mM Tris–HCl pH 8.0, 1.5 M NaCl, and 25 mM $CaCl_2$.

2.4. SDS-Polyacrylamide Gel Electrophoresis

1. Sodium dodecyl sulfate (SDS) solution: SDS is prepared as a 10% (w/v) stock solution and is stored at RT.
2. 5× SDS-PAGE sample buffer: 250 mM Tris–HCl pH 6.8, 500 mM DTT, 10% (w/v) SDS, 0.15% (w/v) bromophenol blue, and 50% (v/v) glycerol. Sample buffer is stored frozen at −20°C in 1 mL aliquots.
3. 5× Stacking gel buffer: 0.625 M Tris–HCl pH 6.8. Store at RT.
4. 5× Separating gel buffer: 1.875 M Tris–HCl pH 8.9. Store at RT.
5. 10× Electrode buffer: 250 mM Tris base, 1.92 M glycine, and 1% (w/v) SDS (no pH adjustment is required). Store at RT.
6. Ammonium persulfate: Prepare a 30% (w/v) solution. Ammonium persulfate solutions may be stored at 4°C for no more than 2 weeks.
7. Acrylamide solution: Acrylamide is dissolved in H_2O to give a 30% (w/v) solution. Filtrated through 0.45 μm pore-size membrane filters and stored in the dark at 4°C. Handled with care, since acrylamide is neurotoxic.
8. Bis-acrylamide is dissolved in H_2O to yield a 2% (w/v) solution. Filtrated through 0.45 μm pore-size membrane filter and stored in the dark at 4°C. Handled with care, since bis-acrylamide are neurotoxic.
9. *N*,*N*,*N*′,*N*′-Tetramethylethylenediamine (TEMED): Stored at 4°C. TEMED must be handled with care, since it is highly flammable, corrosive, and toxic.
10. Molecular weight markers: Prestained Protein Marker, Broad Range (7–175 kDa) may be purchased from New England Biolabs and stored in small aliquots (30 μL) at −20°C (see Note 2).
11. Fixing solution: 40% (v/v) Methanol and 10% (v/v) glacial acetic acid in H_2O.
12. Destain solution: 10% (v/v) Methanol and 5% (v/v) glacial acetic acid in H_2O.
13. Gel staining solution: 0.1% (w/v) Coomassie Brilliant Blue R250 in fixing solution.

2.5. Western-Blot Analysis of Recombinant Proteins Carrying Polyhistidine Tags

1. Transfer buffer: 25 mM Tris base, 190 mM glycine, and 20% (v/v) methanol (no pH adjustment is required). Transfer buffer is stored at RT.
2. Protran® BA85 nitrocellulose membranes and GB002 blotting paper may be purchased from Schleicher & Schuell BioScience.
3. PBS with Tween-20 (PBST): PBS containing 0.1% (w/v) Tween-20 (see Note 3).

4. Blocking buffer: 5% (w/v) nonfat dry milk in PBST.
5. Antibody dilution buffer: Both the primary and the secondary antibodies are diluted in PBST supplemented with 0.5% (w/v) nonfat dry milk.
6. Primary antibody: Mouse monoclonal antipolyhistidine antibody (clone HIS-1) is purchased from Sigma-Aldrich (see Note 4).
7. Secondary antibody: Goat anti-mouse IgG polyclonal antibodies conjugated to horseradish peroxidase may be purchased from Sigma-Aldrich.
8. Enhanced chemiluminescence (ECL) reagents may be purchased from GE Healthcare (Amersham™ ECL Plus) and BioMax XAR films are from Kodak (see Note 5).

3. Methods

The inability of a large number of mammalian proteins to properly fold when overproduced in *E. coli* results in either their proteolytic degradation or the accumulation of their aggregates in bacterial inclusion bodies (1, 4, 23). We chose one such protein, the mouse prion protein (PrP) that is implicated in spongiform encephalopathy, a mammalian protein folding disease (2, 7, 12, 15, 24), as a model for assessing the ability of the HSP fusion system to direct the proper folding of recombinant proteins expressed in *E. coli*.

The HSP-fusion expression vectors that we have developed (13) permit the expression of the target protein as a fusion with an HSP fused to its N terminus and a 6×His tag fused to its C terminus, as shown in Fig. 1. The HSP part of this hybrid aids the protein in attaining its properly folded state, whereas the 6×His tag is used for protein purification by immobilized metal affinity chromatography (IMAC). These HSP-fusion proteins are expressed under the control of the T7 promoter; thus, regulated expression will take place only in *E. coli* strains that synthesize the T7 RNA polymerase, e.g., BL21[DE3] (see Note 6).

Since the protein chaperones operate as multiprotein molecular machines (5), we have also constructed plasmids that permit expression of the appropriate co-chaperones within the same cells that express the HSP-fusions. Plasmid pXCK-E/J expresses the DnaK co-chaperones GrpE and DnaJ, whereas pXCK-ES expresses GroES the co-chaperone of GroEL. The fact that these co-chaperone expression vectors carry the p15A origin of replication makes them compatible with the HSP-fusion expression vectors bearing the ColE1 origin of replication. Additionally, the co-chaperones are also expressed under the control of the phage T7 promoter. Thus, the synchronized induction of both HSP-fusions and co-chaperones should allow for maximum effect on the folding of the recombinant protein.

3.1. Cloning into HSP-Fusion Protein Expression Vectors

1. The gene encoding the mouse PrP was amplified from pVPmPrP14 (a gift from Dr. Sklaviadis, Aristotle University, Thessaloniki, Greece) using primer pairs that allow for the murine PrP gene's in-frame cloning either in the bacterial expression vector pET-21a(+), i.e., mPrP_Nhe_F: 5′-GGGCTAGCAAAAAGCGGCCAAAGCCT-3′ and mPrPNot_R: 5′-GCGCGGCCGCGGATCTTCTCCCGTCGTA-3′, or in the HSP-fusion expression vectors pXCK-K and pXCK-EL, i.e., mPrPHind_F: 5′-GGCAAGCTTAAAAAGCGGCCAAAGCCT-3′ and mPrPNot_R: 5′-GGCGCGGCCGCGGATCTTCTCCCGTCGTA-3′ (see Note 7). The PCR reactions are performed in a final volume of 50 μL, with each reaction mix containing: 10 ng target DNA (plasmid VPmPrP14), 5 μL 10× Promega Thermo Buffer – $MgCl_2$-free, 4 μL 25 mM $MgCl_2$ (final concentration 2 mM), 5 μL 2.5 mM dNTP mixture (0.25 mM final concentration for each dNTP), 5 μL of each 20 μM primer (2 μM final concentration for each primer), and 1 μL of Taq polymerase (2.5 U/μL). Following an initial denaturation period of 4 min at 94°C, the PCR reactions proceed for 30 cycles of 1 min at 94°C, 1 min at 62°C, and 1 min at 72°C, followed by 5 min at 72°C.
2. The PCR products are cloned, using TA-cloning, in T-tailed plasmid pZErO 2.1 (see Note 8). The ligation reaction, which is performed in a final volume of 20 μL, contains 1 μL T-tailed pZero2.1 (50 ng/μL), 3 μL PCR products (reaction mixture), 2 μL 10× T4 DNA ligase buffer, 13 μL H_2O, and 1 μL T4 DNA ligase (400 U/μL). The ligation reactions are incubated at RT for 2 h and then at 4°C overnight prior to being used for the transformation of chemically competent *E. coli* Top10. Specifically, 3 μL of ligation mixture are gently mixed with 200 μL competent *E. coli* Top10 and the mixture is incubated on ice for 30 min prior to being placed into a 42°C water bath for 90 s. Following heat-shock, the mixtures are placed on ice for 2 min before 800 μL of 2XYT medium are added. The cells are then allowed to recover by incubation at 37°C for 45 min with agitation. Subsequently, the cells are plated on 2XYT/agar plates containing kanamycin and incubated overnight at 37°C (see Note 9).
3. Preparation of chemically competent *E. coli*: Bacteria from a frozen stock are struck out on 2XYT/agar plates and incubated overnight at 37°C. The next day, a large colony is picked using an autoclaved toothpick and it is placed into a sterile 500 mL Erlenmeyer flask containing 50 mL 2XYT. The culture is grown at 37°C on a rotating platform under vigorous shaking (200 rpm) until its absorbance at 600 nm (A_{600}) reaches about 0.2 (approximately 3.5 h). The culture is then poured into a prechilled sterile 50 mL conical tube and placed on ice for 10 min before centrifuging the cells at 5,000 × *g* for 10 min in a refrigerated low-speed centrifuge. The supernatant is poured

off and the cell pellet is thoroughly resuspended in 10 mL ice-cold, sterile 0.1 M $CaCl_2$ by repeated pipetting under sterile conditions. The resuspended cells are incubated on ice for 30 min before centrifuging again, as above. After decanting the supernatant, the cells are thoroughly resuspended in 2 mL ice-cold, sterile 0.1 M $CaCl_2$ and kept on ice until use (maximum 2 days).

4. Colonies carrying recombinant plasmids are identified by the presence of appropriate size inserts following digestion of plasmid DNA minipreps with pertinent restriction enzymes. In our test of the system, the recombinant plasmid containing the PCR product generated with the mPrP_Nhe_F/mPrPNot_R primer set was named pZeTA-PrP, whereas the plasmid carrying the PCR product generated with the mPrPHind_F/mPrPNot_R primer set was named pZ-PrP (see Note 7).
5. The PrP sequences are released from pZeTA-PrP by *Nhe*I/*Not*I digestion and from pZ-PrP by *Not*I/*Hind*III digestion and the restriction enzyme digestion products are analyzed on 2% (w/v) low-melt agarose/TBE/EtBr gels. Using visualization with UV-light, the bands containing the PrP sequences are located, carefully excised from the gels using clean razor blades and each is placed in a new, sterile microfuge tube (see Note 10).
6. The isolated PrP inserts are used in ligation reactions, without any further purification, after melting the isolated agarose slices containing them at 65°C for 10 min. PrP fragments released from pZeTA-PrP by *Nhe*I/*Not*I digestion are cloned into *Nhe*I/*Not*I-digested pET-21a (+) to construct pX-PrP, whereas pX-DnaK–PrP and pX-GroEL–PrP are constructed by releasing PrP from pZ-PrP with *Hind*III/*Not*I digestion and cloning into *Hind*III/*Not*I-digested pXCK-K and pXCK-EL, respectively. Specifically, 3 μL of insert in melted low-melt agarose are mixed with 1 μL of appropriately digested expression vector (pET-21a (+), pXCK-K, and pXCK-EL) (50 ng/μL), 2 μL 10× T4 DNA ligase buffer, 13 μL H_2O, and 1 μL T4 DNA ligase (400 U/μL). The ligation reactions are incubated for 2 h at RT prior to being used for the transformation of chemically competent *E. coli* Top10.
7. The transformed Top10 bacteria are plated on 2XYT agar supplemented with ampicillin (100 μg/mL) and incubated overnight at 37°C. Plasmid DNA minipreps are prepared after picking individual colonies and growing overnight cultures at 37°C in 3 mL 2XYT containing ampicillin (100 μg/mL) (see Note 11).
8. The recombinant plasmids are verified by restriction enzyme digestion of the plasmid DNA miniprep preparations. The pX-PrP minipreps are digested with *Nhe*I/*Not*I, whereas the pX-DnaK–PrP and pX-GroEL–PrP minipreps are digested with *Hind*III/*Not*I, to verify the presence of mPrP gene inserts.

3.2. Verification of HSP–PrP Fusion Protein Expression by SDS-PAGE and Western Blot Analysis

1. Chemically competent *E. coli* BL21[DE3] (see Note 6) are transformed with the recombinant protein expression constructs, i.e., pX-PrP, pX-DnaK–PrP, and pX-GroEL–PrP, and the bacteria are grown overnight at 37°C on 2XYT/agar plates supplemented with ampicillin.
2. Individual transformed colonies are used to start overnight cultures in 3 mL 2XYT medium with ampicillin at 37°C.
3. The next day these cultures are diluted tenfold (0.1 mL culture in 0.9 mL of fresh 2XYT media with ampicillin) and incubated for 1.5 h at 37°C under vigorous agitation (200 rpm).
4. Induction of protein expression is initiated by adding IPTG to a final concentration of 1 mM. After a further 1.5 h incubation at 37°C the bacteria are harvested by a 2-min centrifugation at top speed in a microfuge.
5. The supernates are decanted, the tubes are blotted dry on paper towels, and the cell pellets (from 1 mL culture each) are resuspended in 300 μL 1.5× SDS-PAGE sample buffer. The samples are boiled for 6 min prior to being loaded (10 μL/lane) on a 10% SDS-PAGE gel with a 4% stacking-gel.
6. 1.5 mm-thick Mini gel sandwiches for SDS-PAGE are prepared using the appropriate commercially available gel-casting apparatus and components, e.g., glass plates, alumina plates, spacers, combs, etc.
7. The separating gel polymerization mixture (for 50 mL) consists of 16.7 mL 30% acrylamide solution, 8.6 mL 2% bis-acrylamide solution, 10 mL 5× separating gel buffer, 0.5 mL 10% SDS, 14.1 mL H_2O, 120 μL 30% ammonium persulfate, and 20 μL TEMED.
8. The components are mixed thoroughly and sufficient solution is poured into each gel sandwich to reach approximately 2.5 cm below the top edge of the smaller glass (or alumina) plate, leaving space for the stacking gel. Each gel is overlaid with 0.5 mL isopropanol and the apparatus is covered with plastic wrap to prevent isopropanol evaporation. The gel is allowed to polymerize at RT (40–60 min).
9. The isopropanol is poured off from each sandwich, the gel surfaces are rinsed twice with distilled H_2O, and the sandwiches are inverted onto paper towel briefly to drain excess H_2O.
10. The stacking gel polymerization mixture for two mini gels (10 mL) consists of 1.33 mL 30% acrylamide solution, 0.67 mL 2% bis-acrylamide solution, 2 mL 5× stacking gel buffer, 100 μL 10% SDS, 5.8 mL H_2O, 100 μL 30% ammonium persulfate, and 10 μL TEMED. Once the components are mixed, 0.5 mL of the solution are very quickly pipetted into each of the two sandwiches to rinse the gel surfaces. The stacking gels are

poured immediately, the combs inserted, being careful to avoid trapping bubbles, and the gels are allowed to polymerize (30–45 min at RT).

11. Once polymerization of the stacking gel is complete, the comb is removed and the wells of the gel are washed twice with 1× electrode buffer using a 1-mL pipetter.
12. The gel sandwiches are attached to the gel running module, which is then placed into the electrophoresis tank. The upper and lower chambers of the apparatus are filled with electrode buffer and 10 μL aliquots of each sample are loaded into individual wells. A lane containing 10 μL prestained molecular weight markers is included in each gel.
13. Once the proteins of the bacterial cell lysates have been separated by SDS-PAGE, the gels are either stained for total protein or they are transferred electrophoretically to nitrocellulose membranes with the aid of a mini protein-transfer unit (e.g., Hoefer TE22).
14. For transfer, a piece of nitrocellulose membrane that is slightly larger (0.5 cm larger in width and length) than the gel is cut for each gel. The membrane is first placed onto the surface of distilled H_2O in a large clean plastic weigh-boat to hydrate it. Once wet, the membrane is rocked gently in the boat to submerge it and the rocking is continued for 5 min. The water is then substituted for transfer buffer and the rocking is continued for another 10 min.
15. Gels for transfer are removed from the gel sandwiches, placed in weigh-boats containing transfer buffer (at least 30 mL/minigel) and left to incubate for 10–20 min with gentle agitation.
16. The gel transfer cassettes are assembled as follows: Two pieces of precut GB002 blotting paper are wetted with transfer buffer and laid on top of the cassette sponge. The equilibrated gel is placed on top of the wet blotting papers, then the nitrocellulose membrane is laid on top of the gel and two more pieces of prewetted blotting paper are placed on top of the membrane. The surface of this stack of paper, gel, and membrane is rolled gently with a glass-rod to remove any bubbles trapped between the gel and the nitrocellulose membrane. Finally, the second sponge is placed on the top of the stack and the cassette is closed and latched.
17. The cassettes are placed into the transfer electrophoresis unit so that the gel side of it is facing the cathode (black lead) and the membrane side is oriented toward the anode (red lead). A stir-bar is deposited into the tank, the lid is added, the tank is then placed onto a magnetic stir-plate, and the leads are attached to a power supply. Transfer of the proteins is achieved by running the electrophoresis at a constant voltage of 36 V for 3 h at RT.

18. Once the transfer is finished, the membranes are removed from the cassettes and incubated in PBST for 15 min before being placed in blocking solution, either for 2 h at RT or overnight at 4°C, with gentle agitation.
19. To detect the presence of the C-terminal 6×His tag in the recombinant proteins, the membranes are removed from the blocking solution and placed in 10 mL antibody dilution buffer containing a 1:1,000 (v/v) dilution of the monoclonal antipolyhistidine antibody. The membranes are incubated with this primary antibody for 1.5–2 h at RT.
20. The membranes are then washed three times, 5 min each, with PBST before 10 mL of HRP-conjugated secondary antibody [1:5,000 (v/v) dilution in antibody dilution buffer] are added. The membranes are incubated with the secondary antibody for 1.5–2 h at RT.
21. The membranes are washed twice with PBST and twice with PBS (5 min per wash) before they are incubated for 1 min with ECL reagents.
22. Excess ECL reagent is removed by blotting one corner or side of a membrane on paper towel and the blot is then placed between the leaves of a plastic page-protector, which in turn is placed into an X-ray film cassette.
23. In a darkroom, an appropriately sized piece of film is laid over the blot, the cassette is closed, and the film is exposed to the blot long enough for recombinant protein detection (see Note 12).
24. Staining of the gels with Coomassie Brilliant Blue R-250 for detection of total protein content is achieved by placing them in gel staining solution for 2–6 h, with gentle rocking.
25. The gels are then destained by gently rocking in fixing solution, with multiple changes, followed by storage in destain solution to prevent excessive destaining of the protein bands during long-term storage.

3.3. Solubility Assays

1. Using a sterile toothpick, one bacterial colony, carrying the proper expression construct, is picked from a freshly prepared 2X/YT agar plate and is placed into a sterile culture tube containing 3 mL of 2XYT medium with the appropriate antibiotics(s). The culture is incubated overnight at 37°C with vigorous agitation.
2. Fresh media (40 mL) is inoculated with 0.4 mL from each overnight culture and these subcultures are incubated at 37°C under vigorous agitation. When the A_{600} of the subcultures reaches approximately 0.5, IPTG is added to a final concentration of 0.5 mM, and the cultures are incubated for an additional 4 h.

3. The bacteria are harvested by centrifugation at 7,000×*g* for 6 min at 4°C and the bacterial pellets are washed twice with ice-cold PBS.
4. Each washed cell pellet is resuspended in 3 mL cell lysis solution and stirred at 4°C for 30 min. 150 μL of 20% (w/v) Triton X-100 is added and stirring is continued for an additional 5 min period.
5. Using a probe-type sonicator set at 60% power, each cell suspension is sonicated three times, 30 s each at 4°C, with 1–2 min cooling intervals between the pulses. Aliquots of 300 μL from each suspension are placed into clean microfuge tubes, 150 μL 5× SDS-PAGE sample buffer is added to each and the tubes are boiled for 5 min. This portion represents the total protein content of the cell and is identified as *Total Lysate*.
6. Another 0.5 mL from each lysate are centrifuged at top-speed (15,000×*g*) for 30 min at 4°C in a microfuge. 300 μL from each clarified supernate are placed in clean microfuge tubes, mixed with 150 μL 5× SDS-PAGE sample buffer each, and boiled for 5 min. This portion represents the soluble protein content of the cell and is identified as *Soluble Lysate*.
7. The pellets from the 30 min centrifugation step are each suspended in 750 μL 1.5× SDS-PAGE sample buffer and boiled for 5 min. These samples represent the insoluble protein content of the cell that is present in the bacterial inclusion bodies and are identified as *Inclusion Bodies.*
8. Examination of the distribution of the recombinant proteins in the soluble lysate and the inclusion bodies fractions is achieved by analysis of equal volumes (10 μL) of the above-prepared extracts by SDS-PAGE followed either by total protein staining or Western blot analysis with antipolyhistidine antibodies. The results presented in Fig. 3 indicate that while PrP accumulated mainly in the inclusion bodies, the majority of the DnaK–PrP and GroEL–PrP fusion is found in the soluble fraction.

3.4. Co-expression of the DnaK and GroEL Co-chaperones to Increase the Solubility of the HSP-PrP Fusions

1. *E. coli* BL21[DE3] carrying either the unfused PrP expression plasmid pX-PrP or the appropriate HSP-fusion expression constructs, i.e., pX-DnaK–PrP and pX-GroEL–PrP, are made chemically competent (see Subheading 3.1 step 3) and then transformed with the co-chaperone expression constructs pXCK-E/J or pXCK-ES.
2. The transformed bacteria are selected by overnight growth at 37°C on 2XYT/agar plates supplemented with both ampicillin and chloramphenicol.
3. Individual transformed colonies are used to start overnight cultures in 3 mL 2XYT medium with ampicillin and chloramphenicol at 37°C and induction of protein expression and solubility assays are performed as described in Subheading 3.3.

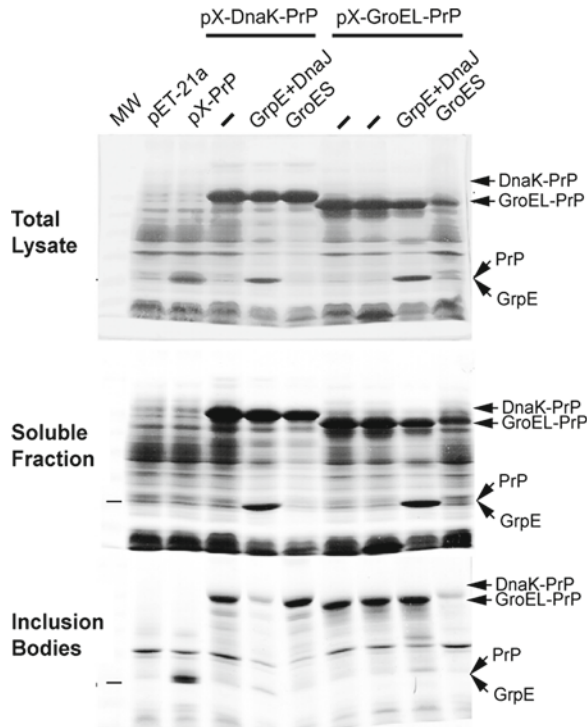


Fig. 3. Comparison of the partitioning of unfused or chaperone-fused PrP to the soluble fraction or inclusion bodies. Total lysates were prepared from control cells [pET-21a (+)], from cells expressing PrP-6×His (pX-PrP), DnaK–PrP-6×His or GroEL–PrP-6×His fusions (pX-DnaK–PrP and pX-GroEL–PrP), either in the absence (–) or the presence of GrpE and DnaJ (GrpE + DnaJ) or GroES. Aliquots of these lysates (total lysates) were put aside and the remainder of each was centrifuged for 30 min at 15,000 × *g* to derive the soluble protein fraction (soluble fraction) and the insoluble protein fraction (inclusion bodies). All three fractions were analyzed directly on 11% SDS-PAGE gels and total proteins were visualized by Coomassie Brilliant Blue R250 staining. The *arrows* indicate the positions of the relevant proteins.

4. The results in Fig. 3 indicate that the parallel expression of the co-chaperone increases the solubility of the fusion protein of their cognate partner, i.e., DnaJ/GrpE expression increases the DnaK–PrP solubility, and GroES increases the GroEL–PrP solubility. In contrast, the presence of co-chaperone(s) has no effect on the solubility of the unfused PrP or the noncognate HSP-protein fusion.

3.5. Purification of the 6×His-Tagged HSP-Fusion Proteins

1. An overnight culture is initiated by placing a single bacterial colony, from a freshly prepared plate, into 30 mL 2XYT medium with the proper antibiotics(s) and it is incubated overnight at 37°C with constant vigorous agitation on a rotating platform (200 rpm).
2. Overnight cultures (30 mL) are used to inoculate 500 mL sterile 2XYT/antibiotic media in 2 L Erlenmeyer flasks. The cultures

are grown at 37°C under vigorous agitation (200 rpm) and induction of recombinant protein expression is initiated by adding IPTG (0.5 mM final concentration) when the A_{600} reaches 0.5. The cultures are induced for 4 h at 36.5°C before collecting the cells by centrifugation at 7,000 × *g* for 6 min at 4°C. The cell pellets are washed twice with ice-cold PBS before being stored frozen at −80°C.

3. The frozen cell pellet is thawed in EDTA-free cell lysis solution (15 mL/g wet cell pellet), stirred for 30 min at 4°C, and then Triton X-100 is added to achieve 1% (w/v) concentration. Stirring is continued for another 5 min before the suspension is sonicated three times for 30 s each at 4°C with 1 min cooling intervals using a probe-type sonicator at 60% power setting (see Note 13).
4. The soluble protein extract is obtained by centrifuging the cell lysate at 12,000 × *g* for 30 min at 4°C and taking the clear supernatant.
5. The soluble protein extract is mixed with an appropriate amount of a 50% (v/v) suspension of the Ni-NTA agarose pre-equilibrated in EDTA-free cell lysis solution/1% Triton X-100 (~3 mL 50% suspension/g of wet cell pellet). The suspension is incubated for 30 min at 4°C, under constant mixing (see Note 14).
6. The agarose gel slurry is poured into the column, letting the agarose settle for 5 min and then starting the flow (slowly, drop-wise) to pack the material into the column. The protein extract is recirculated through the column matrix two to three times to ensure maximum binding of the 6×His-tagged protein.
7. The column is washed with at least 3 column volumes of EDTA-free cell lysis solution, followed by 10 column volumes of the same buffer supplemented with 10 mM $MgCl_2$ and 10 mM ATP (see Note 15).
8. The column is washed with at least 10 column volumes of the *Wash Buffer* to remove any proteins that bind the Ni-NTA with low affinity.
9. The 6×His-tagged HSP-fusion recombinant proteins are eluted from the affinity column by passing at least 10 volumes *Elution Buffer* through the resin. 1 mL protein fractions are collected, the protein content is determined (see Note 16), and peak fractions are pooled.
10. The pooled protein solution is dialyzed overnight at 4°C against 500 volumes *Dialysis Buffer* and the dialyzed protein is aliquoted and kept frozen at −80°C.

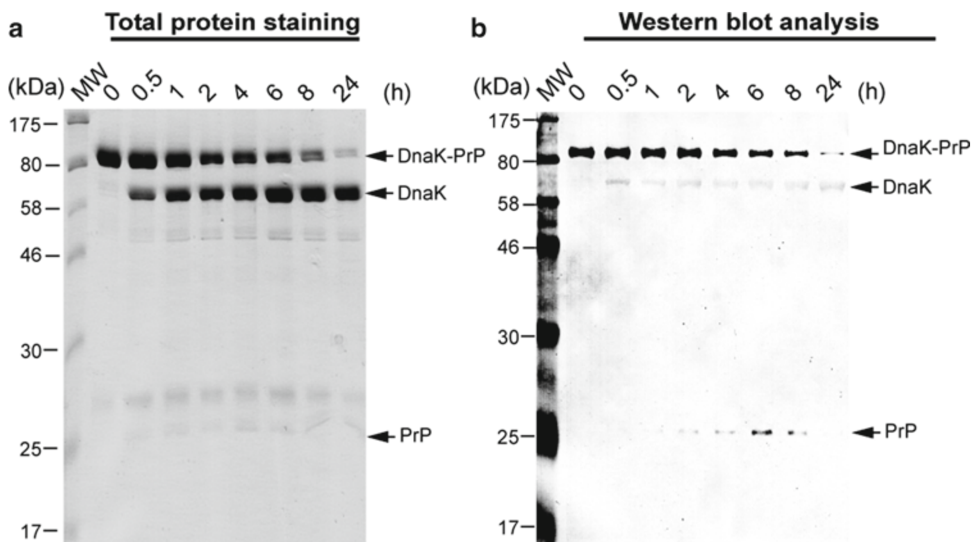


Fig. 4. Thrombin cleavage of the DnaK-PrP fusion. Purified DnaK–PrP-6XHis was incubated with thrombin (6 U/mg fusion protein) and aliquots were withdrawn at the indicated time points. The digestion products were analyzed on 12% SDS-PAGE gels, which were either (**a**) stained with Coomassie Brilliant Blue R250 (total protein staining), or (**b**) transferred electrophoretically to a nitrocellulose membrane and probed with an antipolyhistidine monoclonal antibody (Western blot analysis). The positions of the relevant proteins are indicated by *arrows.*

3.6. Thrombin-Cleavage of PrP from Its HSP Partner

1. To recover the PrP, the purified HSP–PrP fusion protein is proteolytically cleaved with thrombin, a cleavage site for which had been incorporated at the C terminus of the HSP protein as shown in Fig. 1.
2. To identify the optimal conditions for the specific and complete cleavage of the fusion protein, a time-course experiment is performed. Specifically, 20 μL purified DnaK–PrP (2.8 mg/mL) are mixed with 10 μL 10× thrombin cleavage buffer and 3 μL diluted thrombin (0.1 U/μL, final thrombin concentration 6 U/mg fusion protein) and the final reaction volume is adjusted to 100 μL with H_2O.
3. The reaction is carried out at RT and 20 μL aliquots are withdrawn at the indicated times, mixed with 5 μL 5× SDS-PAGE sample buffer, and boiled for 5 min.
4. As shown in Fig. 4, the products of the thrombin digestions are analyzed by SDS-PAGE on 12% gels, which are either stained for total proteins with Coomassie Brilliant Blue R250 (a) or analyzed by Western blotting using an antipolyhistidine monoclonal antibody (b).
5. The results presented in Fig. 4 indicate that the DnaK–PrP fusion can be cleaved readily with thrombin (see Note 17) to yield PrP protein. This cleaved PrP remains soluble even after the removal of its HSP-fusion partner (see Note 18).

4. Notes

1. A T-tailed derivative of the plasmid vector pZErO 2.1 (Invitrogen) is used for TA-cloning of the PCR products. This vector is prepared as follows: 38 μL *Eco*RV-digested pZErO 2.1 (7 μg), 5 μL 10× Promega Thermo buffer ($MgCl_2$-free), 1 μL 25 mM $MgCl_2$, 2 μL 25 mM dTTP, and 1 μL Taq polymerase (2.5 U/μL) are mixed and the 50-μL reaction is incubated at 72°C for 2 h. The reaction products are extracted sequentially with phenol–SEVAG and SEVAG prior to precipitation of the DNA, which is achieved by adding 2.5 volumes of absolute ethanol to the aqueous phase. Following two washes with 70% ethanol, the DNA pellet is air dried for 10 min before being dissolved in TE to a final vector concentration of 50 ng/μL. Each frozen aliquot of the T-tailed pZErO 2.1 TA-cloning vector may not be subjected to more than three freeze–thaw cycles, since its cloning efficiency is severely reduced by such treatment.
2. The prestained protein marker aliquot should be placed in a 42°C water bath for approximately 3 min before loading on the SDS-PAGE. For best results, avoid freezing and thawing the aliquots more than three times.
3. PBST is best used fresh, although it can be stored at 4°C for a few days.
4. Monoclonal and polyclonal antipolyhistidine antibodies can be obtained from several suppliers. Be careful to use the proper secondary antibody when you use a rabbit polyclonal antipolyhistidine antibody, e.g., a goat anti-rabbit IgG antibody conjugated to HRP.
5. ECL reagents and X-ray films optimized for ECL-detection can be purchased from several other vendors.
6. *E. coli* BL21[DE3] is a lysogen carrying a lamda prophage bearing the T7 polymerase gene under the control of the *lac* promoter. Thus, the expression of T7 polymerase can be initiated by the addition of IPTG, thereby providing the polymerase needed to transcribe the recombinant genes. This strain's deficiency in both *lon* and *ompT* proteases further increases its value as a recombinant protein expression host.
7. The primers carry sequences containing specific restriction enzyme sites to aid the in-frame incorporation of the target gene sequences into the appropriate expression vector. Note that different restriction sites are used when the PrP gene is to

be produced as an HSP-fusion and when it is to be expressed without such an N-terminal fusion tag in the control vector pET-21a(+).

8. There is no absolute requirement for TA-cloning of the PCR amplimers, since they can be cloned directly into expression vectors following restriction enzyme digestion. However, since direct cloning can be a low-yield process due to incomplete digestion of the restriction sites placed near the ends of the DNA amplimers, we recommend this intermediate step to increase efficiency.
9. When no insert has been introduced, plasmid pZErO 2.1 expresses the lethal *ccdB* gene and kills the cells. Thus, the vast majority of the colonies obtained when using the pZErO 2.1-based TA-cloning system are positive for PCR amplimers.
10. Use minimal exposure of DNA to UV radiation to avoid DNA damage.
11. A number of miniprep plasmid DNA isolation kits are available from several suppliers. The pertinent manufacturer's instructions should be followed.
12. The films are developed with the appropriate developing and fixing reagents, either manually or in an automated film developing unit.
13. The presence of EDTA in the protein solution, or in any of the buffers, will remove the nickel ions from the Ni-NTA-agarose column thereby either preventing the protein binding to the resin or effecting its quantitative elution from the column.
14. Do not use magnetic stirrer bars for mixing, because they may crush the agarose beads.
15. The column wash buffer containing $MgCl_2$ and ATP is designed to dissociate any protein complexes with HSPs (e.g., the complexes of DnaK either with GrpE or polypeptide substrates dissociate in the presence of Mg/ATP) and thus increase the purity of the isolated HSP-fusion proteins.
16. Commercially available protein determination kits are used according to the manufacturer's instructions.
17. Unlike DnaK–PrP, the GroEL–PrP fusion fails to be cleaved by thrombin. This could either be due to a lack of accessibility of the thrombin recognition site, which may be "buried" within the oligomeric (14-mer) GroEL structure, or it may be indicative of improper folding of the GroEL–PrP fusion.
18. The PrP protein that is cleaved from the DnaK–PrP fusion is soluble and can be purified away from the DnaK with IMAC under native conditions.

Acknowledgments

The authors would like to thank Prof. Saul J. Silverstein (Columbia University, N.Y.C., USA.) for his continuing interest and support, Prof. Theodoros Sklaviadis (Aristotle University, Thessaloniki, Greece) for plasmid pVPmPrP14, and Dr. Cynthia H. Panagiotidis for critically reading the manuscript and providing valuable comments.

References

1. Baneyx, F. and Mujacic. M. (2004) Recombinant protein folding and misfolding in Escherichia coli. *Nat Biotechnol* **22**, 1399–408.
2. Cohen, F. E. (1999) Protein misfolding and prion diseases. *J Mol Biol* **293**, 313–320.
3. de Marco, A., Deuerling, E., Mogk, A., Tomoyasu, T. and Bukau, B. (2007) Chaperone-based procedure to increase yields of soluble recombinant proteins produced in *E. coli. BMC Biotechnol* 7, 32.
4. Fahnert, B., Lilie, H. and Neubauer, P. (2004) Inclusion bodies: formation and utilisation. *Adv Biochem Eng Biotechnol* **89**, 93–142.
5. Fink, A. L. (1999) Chaperone-mediated protein folding. *Physiol Rev* **79**, 425–449.
6. Georgiou, G. and Valax, P. (1996) Expression of correctly folded proteins in *Escherichia coli. Curr Opin Biotechnol* 7, 190–197.
7. Gerber, R., Tahiri-Alaoui, A., Hore, P. J. and James, W. (2007) Oligomerization of the human prion protein proceeds via a molten globule intermediate. *J Biol Chem* **282**, 6300–6307.
8. Gold, L. (1990) Expression of heterologous proteins in *Escherichia coli. Methods Enzymol* **185**, 11–4.
9. Hockney, R. C. (1994) Recent developments in heterologous protein production in *Escherichia coli. Trends Biotechnol* **12**, 456–463.
10. Hodgson, J. (1993) Expression systems: a user's guide. Emphasis has shifted from the vector construct to the host organism. *Biotechnology (N Y)* **11**, 887–893.
11. Hoffmann, F. and Rinas, U. (2004) Roles of heat-shock chaperones in the production of recombinant proteins in *Escherichia coli. Adv Biochem Eng Biotechnol* **89**, 143–161.
12. Kupfer, L., Hinrichs, W. and Groschup, M. H. (2009) Prion protein misfolding. *Curr Mol Med* **9**, 826–835.
13. Kyratsous, C. A., Silverstein, S. J., DeLong, C. R. and Panagiotidis, C. A. (2009) Chaperone-fusion expression plasmid vectors for improved solubility of recombinant proteins in *Escherichia coli. Gene* **440**, 9–15.
14. Makrides, S. C. (1996) Strategies for achieving high-level expression of genes in *Escherichia coli. Microbiol Rev* **60**, 512–38.
15. Moore, R. A., Taubner, L. M. and Priola, S. A. (2009) Prion protein misfolding and disease. *Curr Opin Struct Biol* **19**, 14–22.
16. Nishihara, K., Kanemori, M., Kitagawa, M., Yanagi, H. and Yura, T. (1998) Chaperone coexpression plasmids: differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in *Escherichia coli. Appl Environ Microbiol* **64**, 1694–1699.
17. Olins, P. O., and Lee, S. C. (1993) Recent advances in heterologous gene expression in *Escherichia coli. Curr Opin Biotechnol* **4**, 520–525.
18. Perez-Perez, J., Martinez-Caja, C., Barbero, J. L. and Gutierrez, J. (1995) DnaK/DnaJ supplementation improves the periplasmic production of human granulocyte-colony stimulating factor in *Escherichia coli. Biochem Biophys Res Commun* **210,** 524–529.
19. Schlapschy, M., Grimm, S. and Skerra, A. (2006) A system for concomitant overexpression of four periplasmic folding catalysts to improve secretory protein production in *Escherichia coli. Protein Eng Des Sel* **19**, 385–390.
20. Schlieker, C., Bukau, B. and Mogk, A. (2002) Prevention and reversion of protein aggregation by molecular chaperones in the *E. coli* cytosol: Implications for their applicability in biotechnology. *J Biotechnol* **96**, 13–21.
21. Stevens, J. M., Rao Saroja, N., Jaouen, M., Belghazi, M., Schmitter, J. M., Mansuy, D., Artaud, I. and Sari, M. A. (2003) Chaperone-assisted expression, purification, and characterization of recombinant nitrile hydratase NI1 from *Comamonas testosteron*i. *Protein Expr Purif* **29**, 70–76.

22. Vonrhein, C., Schmidt, U., Ziegler, G. A., Schweiger, S., Hanukoglu, I. and Schulz, G. E. (1999) Chaperone-assisted expression of authentic bovine adrenodoxin reductase in *Escherichia coli*. *FEBS Lett* **443**, 167–169.
23. Wetzel, R. (1994) Mutations and off-pathway aggregation of proteins. *Trends Biotechnol* **12**, 193–198.
24. Winklhofer, K. F., Heller, U., Reintjes, A. and Tatzelt, J. (2003) Inhibition of complex glycosylation increases the formation of PrPsc. *Traffic* **4**, 313–322.
25. Xu, Y., Weng, C. L., Narayanan, N., Hsieh, M. Y., Anderson, W. A., Scharer, J. M., Moo-Young, M. and Chou, C. P. (2005) Chaperone-mediated folding and maturation of the penicillin acylase precursor in the cytoplasm of *Escherichia coli*. *Appl Environ Microbiol* **71**, 6247–6253.
26. Yanase, H., Moriya, K., Mukai, N., Kawata, Y., Okamoto, K. and Kato, N. (2002) Effects of GroESL coexpression on the folding of nicotinoprotein formaldehyde dismutase from *Pseudomonas putida* F61. *Biosci Biotechnol Biochem* **66**, 85–91.

Chapter 6

The Use of a Flagellar Export Signal for the Secretion of Recombinant Proteins in *Salmonella*

Ferenc Vonderviszt, Ráchel Sajó, József Dobó, and Péter Závodszky

Abstract

The flagellum-specific export system is a specialized type III export machinery, which exports external flagellar proteins through the central channel of the flagellar filament. A number of evidence indicates that short segments within the disordered N-terminal region of flagellar axial proteins are recognized by the flagellum-specific export apparatus. Recently, we have demonstrated that the 26–47 segment of *Salmonella typhimurium* flagellin is capable of mediating flagellar export. N-terminal flagellin segments containing the export signal combined with a hexahistidine tag can be attached to heterologous proteins (preferentially in the size range of 9–40 kDa) facilitating their secreted expression and easy purification from the medium. Certain over-expressed proteins that are easily degraded within the cells are found intact in the medium implying a potential application of this expression system for proteins of high proteolytic susceptibility.

Key words: Flagellum-specific export, Type III secretion, Export signal, Recombinant protein, *Salmonella typhimurium*, Secreted over-expression, Flagellin

1. Introduction

Flagella are the locomotion organelles of bacteria (1). A membrane embedded molecular motor rotates a long helical filament that works as a propeller driving the bacterium through the liquid environment. The filamentous portion of the flagellum extends from the cytoplasm to the cell exterior and involves several substructures: the rod, the hook, the hook–filament junction, the long helical filament, and a cap at the filament tip (2). The helical filament may comprise as many as 30,000 flagellin subunits, and can grow up to about 15 μm. External flagellar proteins, lying beyond the cytoplasmic membrane, are synthesized in the cell and exported sequentially by the flagellum-specific export apparatus from the cytoplasm to the

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_6, © Springer Science+Business Media, LLC 2012

site of assembly at the distal end of the growing filament (3, 4). They are assumed to move through the narrow, 2–3 nm wide, central channel of the flagellum (5) as partially unfolded monomers.

The flagellar protein export system is located at the cytoplasmic face of the basal part of the flagellum to distinguish flagellar proteins from other cytoplasmic proteins and to facilitate their transportation. Recent studies have demonstrated that the proton motive force across the cytoplasmic membrane is responsible for driving the export process (6, 7). The flagellar protein export system belongs to the family of the type III secretion systems (8), which also include those for secretion of virulence factors by a wide variety of pathogenic bacteria (9). A growing number of evidence indicates that the export signal is located in the disordered N-terminal region of the secreted proteins (10–12). We have demonstrated that the 26–47 disordered segment of *Salmonella* flagellin contains the recognition signal for the flagellar export machinery (13).

The flagellar export system has the potential to be used as an effective tool for the secretion of recombinant proteins. Our studies have revealed that the recognition signal can facilitate translocation of a wide variety of attached foreign proteins through the flagellum-specific export pathway into the outer medium (13–15). It was observed that specific characteristics of given proteins may cause significant differences in export efficiency. The major reason hindering secretion can be inclusion body formation within the cells. When only the soluble fraction of the expressed proteins is considered the export efficiency is fairly high: the average efficiency was above 50% overall (15).

A *lac* promoter-based expression vector was created to facilitate efficient over-expression, secretion, and purification. Our constructs were designed to contain a hexahistidine tag at the N terminus to enable easy detection, and purification in one step on a Ni^{2+} affinity matrix. Another feature of the constructs is that they contain an enterokinase (EK) cleavage site right after the export signal. The EK site enables selective removal of the flagellar and other sequence elements preceding the expressed protein. As an example, secreted expression and purification of 26–47-MBP (maltose binding protein and 26–47 segment of flagellin as a fusion protein) is characterized in detail in this chapter. The flagellin deficient *Salmonella typhimurium* strain, SJW2536, is used as a host strain in our secretion experiments (16). As a result, the export apparatus is not overloaded with the most abundant flagellar protein, flagellin. Since SJW2536 assembles only the basal body and the hook part of the flagellum, the expressed fusion proteins do not need to travel through a very long filament.

A major advantage of secreted protein expression is that cell disruption is no longer required before purification, and the starting material is less contaminated with host proteins, lipopolysaccharides, and nucleic acids. It gives a chance that certain proteins which are

degraded or form intracellular aggregates can be expressed in intact and soluble form. Solubility, however, does not necessarily mean that the protein of interest is correctly folded. Flagellar proteins are translocated in a largely unfolded conformation through the narrow export channel by the flagellar export machinery. It is plausible to assume that fusion proteins containing the export signal are also transported to the medium in an unfolded form, where they may, or may not adopt their native conformation. Anyway, the lack of competing aggregation is supposed to highly facilitate the folding process. This expression system has a potential for unfolded or unstructured proteins that are easily degraded intracellularly but may escape proteolysis if secreted efficiently into the medium.

2. Materials

2.1. DNA Work

1. Synthetic gene (GenScript).
2. pGFP plasmid (Clontech).
3. High-fidelity DNA polymerase and buffer.
4. T4 DNA ligase.
5. Competent *E. coli* cells (e.g., TOP10, XL1-Blue).
6. Gel extraction kit (e.g., Invitrogen, PureLink Quick Gel Extraction kit).
7. Plasmid miniprep kit (e.g., Promega, Wizard Plus SV Minipreps Purification System).
8. Restriction endonucleases (New England Biolabs, or Fermentas).
9. Agarose.
10. TAE buffer: 50 mM Tris–acetate pH 8.0, 20 mM Na–acetate, and 2 mM EDTA.
11. TE buffer: 10 mM Tris–HCl pH 7.5 and 0.1 mM EDTA.

2.2. Cell Culture and Protein Expression

1. *S. typhimurium* ΔfliC strain, SJW2536.
2. Luria-Bertani medium (LB): 10 g Bacto-tryptone, 5 g bacto-yeast extract, 10 g NaCl, add H_2O to a final volume of 1 L, adjust pH to 7.2 with NaOH, and sterilize by autoclaving.
3. LB-agar for plates: LB medium supplemented with 15 g agar per liter. After autoclaving cool to about 50°C while stirring, then pour the plates.
4. Ampicillin stock solution: Prepare a 100-mg/mL stock in water, and sterilize by filtration through a 0.22-μm sterile syringe filter.
5. LB–ampicillin, LB–ampicillin–agar: As LB and LB-agar but containing 100 μg/mL ampicillin, cool the LB-agar to 50°C before adding ampicillin and pouring the plates.

6. 15% (v/v) Glycerol solution: Sterilize by autoclaving.
7. 20% (w/v) Glucose solution: Sterilize by filtration through a 0.22-μm sterile syringe filter.
8. Metal stock solution: 8 mL concentrated HCl, 5 g $FeCl_2 \cdot 4H_2O$, 184 mg $CaCl_2 \cdot 2H_2O$, 64 mg H_3BO_3, 40 mg $MnCl_2 \cdot 4H_2O$, 18 mg $CoCl_2 \cdot 2H_2O$, 340 mg $ZnCl_2$, 605 mg $Na_2MoO_4 \cdot 2H_2O$ dissolved in H_2O in a final volume of 100 mL; store at room temperature (RT).
9. Thiamine solution: Dissolve 22 mg thiamine HCl in H_2O in a final volume of 100 mL, and sterilize by filtration through a 0.22-μm sterile syringe filter. Thiamine is light sensitive, therefore, store this solution at 4°C in a dark bottle.
10. Minimal medium: Dissolve 12 g $Na_2HPO_4 \cdot 7H_2O$, 3 g KH_2PO_4, 0.5 g NaCl, 1 g NH_4Cl in H_2O in a final volume of 986 mL, add 10 mL of 20% glucose, 1 mL of thiamine solution, 40 μL of metal stock solution, 1 mL of 1 M $MgCl_2$, 15 μL of 1 M $CaCl_2$, and 1 mL 100 mg/mL ampicillin into the solution. Sterilize the glucose, thiamine and metal stock solution by filtration through a 0.22-μm sterile syringe filter, and the other solutions by autoclaving.
11. Minimal medium supplemented with 5% LB: Add 50 mL of sterile LB medium to 950 mL of minimal medium.
12. Petri dishes.
13. Sterile syringe filter, 0.22 μm.
14. Electroporation system (BioRad).
15. 2-mm electroporation cuvettes (BioRad).

2.3. Protein Purification and Processing

1. 80% (w/v) Trichloroacetic acid (TCA): Dissolve 40 g TCA in H_2O in a final volume of 50 mL, and store at 4°C.
2. Standard SDS-PAGE sample buffer: 125 mM Tris–HCl pH 6.8, 15% glycerol, 4.5% sodium dodecyl sulfate (SDS), 0.03% bromophenol blue, and 5% β-mercaptoethanol.
3. Nonstandard SDS-PAGE sample buffer: 250 mM Tris–HCl pH 8.0, 15% glycerol, 4.5% SDS, 0.03% bromophenol blue, 5% β-mercaptoethanol (for TCA precipitated proteins).
4. Dialysis tubing.
5. Ni-NTA Superflow resin (Qiagen).
6. Empty 20-mL Econo-Pac gravity-flow chromatography column (BioRad).
7. Loading buffer: 400 mM NaCl, 10 mM imidazole, 50 mM Na-phosphate pH 7.5, and 0.01% NaN_3.
8. Elution buffer: 400 mM NaCl, 400 mM imidazole, 50 mM Na-phosphate pH 7.5, and 0.01% NaN_3.

9. Enterokinase (EK), recombinant, light chain (Sigma).
10. EK buffer: 50 mM Tris–HCl pH 7.5, 0.2 mM $CaCl_2$, and 0.1% Tween-20.
11. Amylose resin: Store at 4°C.
12. Empty 2-mL Poly-Prep gravity-flow chromatography column (BioRad).
13. Amylose wash buffer: 20 mM Tris–HCl pH 7.4, 200 mM NaCl, and 1 mM EDTA.
14. Amylose elution buffer: Same as amylose wash buffer but containing 10 mM maltose.
15. Soybean trypsin inhibitor agarose: Store at 4°C.
16. 10× Stripping buffer: 1 M NaCl, and 1 M formic acid pH 3.0.
17. 10× Binding buffer: 5 M NaCl, and 500 mM potassium phosphate pH 8.0.
18. Gel documentation system (BioRad).

3. Methods

3.1. Construction of the Expression Vector, pVJGFPa

1. Order a synthetic gene (usually supplied in a plasmid) containing a His_6-tag, the 26–47 flagellin segment of *S. typhimurium* and an EK cleavage site (GenScript). (Fig. 1. boxed sequence between the *Hind*III and *Kpn*I sites).
2. Digest about 2 μg of plasmid containing the synthetic cassette and 1 μg of a pGFP vector with 10 U of *Hind*III and *Kpn*I restriction endonucleases each in a final volume of 20–50 μL according to the manufacturer's instructions.
3. Run the samples on 1% agarose gel in TAE buffer, excise the appropriate bands, and extract the DNA using a gel extraction kit according to the manufacturer's instructions.
4. Ligate the extracted cassette and pGFP plasmid using T4 DNA ligase: Mix 20–100 ng of the vector and the insert in 1:3 molar ratio and add sterile H_2O in a final volume of 17 μL. Heat the solution to 42°C for 2 min, cool it down on ice, add 2 μL 10× T4 ligase buffer and 1 μl T4 ligase. The final volume of the solution should be 20 μL. Incubate the solution at 16°C for 2 h.
5. Transform the ligation products into *E. coli* competent cells. Thaw a 100-μL aliquot of *E. coli* competent cells on ice and add the total volume of ligation product. Incubate it on ice for 30 min, then put the mixture to 42°C for 90 s. Cool it on ice for 2 min, then add 1 mL of sterile LB.
6. Incubate the transformed cells at 37°C for 1 h.

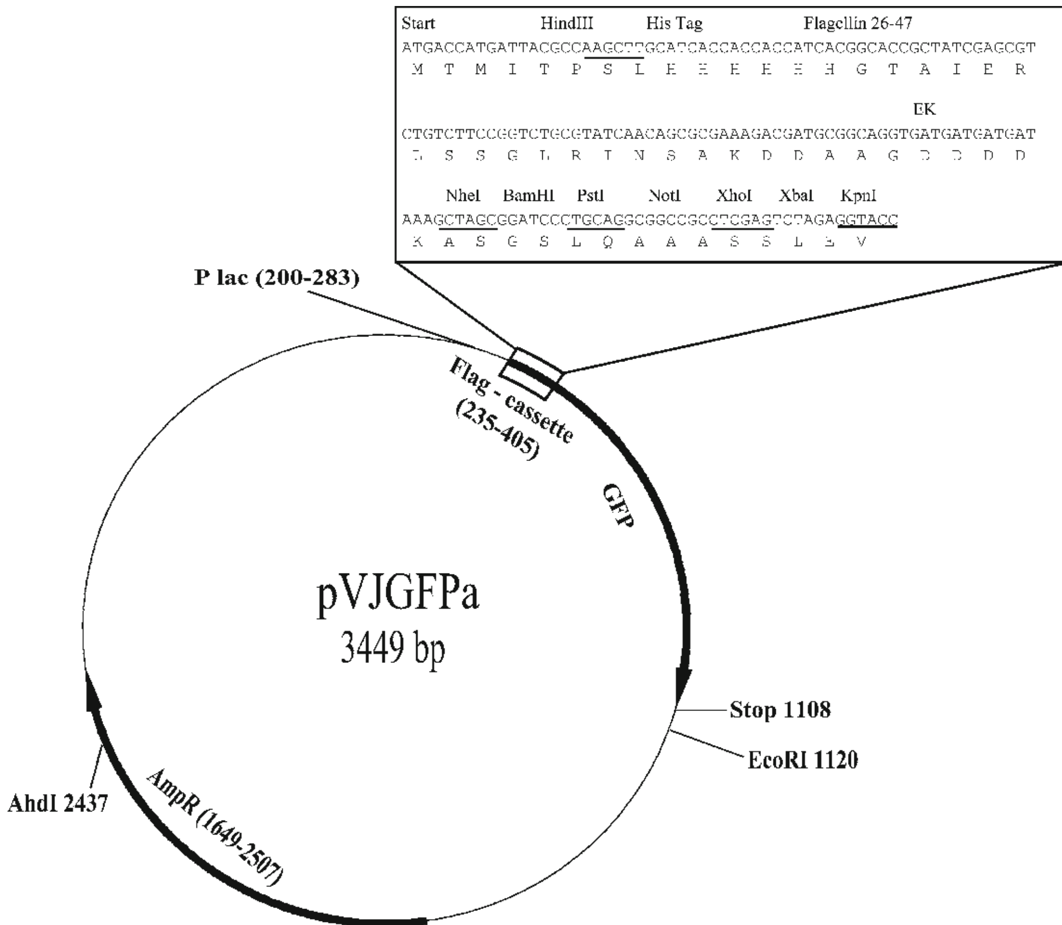


Fig. 1. Map of the pVJGFPa vector. A synthetic cassette, containing a His_6-tag, the flagellar export signal, and an enterokinase cleavage site, is inserted into a pGFP plasmid between the *Hind*III and *Kpn*I sites resulting in the pVJGFPa plasmid. The sequence of the cassette is shown in the *box.*

7. Streak the cells onto LB-agar plates containing 100 μg/mL ampicillin.
8. Pick up a single colony from the plate and add 3 mL sterile LB containing 100 μg/μL ampicillin. Grow it overnight at 37°C.
9. From a 2-mL overnight culture isolate the ligated plasmid with a plasmid miniprep kit.
10. Digest the plasmid with *Hind*III and *Kpn*I as above in a final volume of 10 μL and analyze the products by agarose gel electrophoresis to ensure it contains the synthetic cassette (see Note 1). The resulting plasmid (Fig. 1) is called pVJGFPa (15).

3.2. Subcloning of the Gene of Interest into the Expression Vector

The gene of the protein of interest can be cloned into any of the restriction sites (from *Nhe*I to *Kpn*I) listed on Fig. 1. Use a restriction site on the 5′ end to ensure that your gene is in frame with the flagellar signal. The 3′ end should contain a stop codon before the other

restriction site. If your gene does not contain the appropriate restriction sites, amplify it by PCR with primers (see Note 2) that will add the required cleavage sites to both ends (see Note 3).

1. Digest both the vector (pVJGFPa) and the amplified insert (both approx. 1 μg) with the chosen enzymes (10 U each) according to the manufacturer's instructions. We recommend using two enzymes.
2. Run the samples on 1% agarose gel and extract the appropriate bands using a gel extraction kit.
3. Ligate and transform as above (steps 4 and **5** of Subheading 3.1).
4. Incubate the transformed cells at 37°C for 1 h.
5. Streak 50 μL fractions (see Note 4) of the cells onto LB-agar plates containing 100 μg/mL ampicillin and incubate them overnight at 37°C.
6. Pick up a single colony from the plate and add 3 mL of sterile LB containing 100 μg/μL ampicillin. Grow the cells overnight at 37°C.
7. From a 2 mL overnight culture isolate the ligated plasmid with a plasmid miniprep kit.
8. Digest the plasmid with the same enzymes that were used for cloning and analyze the products by agarose gel electrophoresis to ensure it contains your gene.

3.3. Electroporation into Salmonella

1. Grow *S. typhimurium* SJW2536 cells in 20 mL sterile LB medium at 37°C until OD_{600} reaches approximately 0.6.
2. Harvest the cells by centrifugation at 3,000 × *g* for 10 min at 4°C and resuspend them in 10 mL of ice-cold sterile 15% glycerol solution.
3. Centrifuge again and remove the supernatant. Repeat the washing step with 10 mL of ice-cold sterile 15% glycerol.
4. Finally, resuspend the cells in 1 mL of ice-cold sterile 15% glycerol by careful up and down pipetting. These cells are electrocompetent (see Note 5).
5. Perform all the processes on ice. To a 100-μL aliquot of the cells add 50–200 ng of plasmid DNA in a maximum volume of 10 μL. The DNA should preferably be dissolved in sterile ultra pure H_2O rather than in TE buffer.
6. Electroporate the cells using an exponential voltage decay protocol (2.4 kV, 25 μF, 400 Ω, 2-mm cuvette, the time constant should be 9–13 ms).
7. Add 1 mL of LB. After 1 h incubation at 37°C, streak appropriate fractions of the cells (e.g., 50 μL) onto LB-agar plates containing 100 μg/mL ampicillin.
8. Incubate the plates overnight at 37°C.

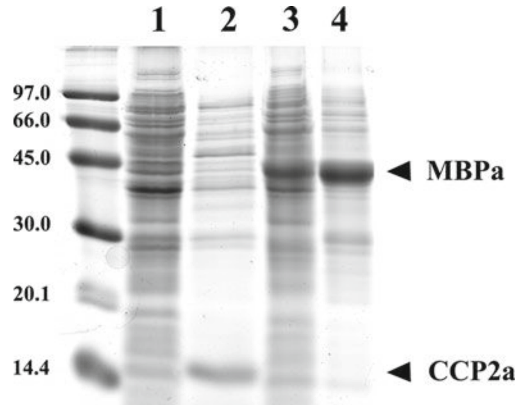


Fig. 2. Expression level of two model proteins (CCP2, MBP) within the cells and in the cell culture medium. Each lane of the SDS polyacrylamide gel represents cells (pellet) from 100 μL, or medium (supernatant) from 250 μL of cell culture. Before analysis cells were centrifuged and resuspended in 15 μL of water and equal volume of standard SDS-PAGE sample buffer, while the cell culture media were precipitated with 80% TCA, and resuspended in 30 μL of nonstandard SDS-PAGE sample buffer. Lane 1: disrupted SJW2536 cells containing pVJCCP2a. Lane 2: concentrated medium of SJW2536 harboring pVJCCP2a. Lane 3: disrupted SJW2536 cells containing pVJMBPa. Lane 4: concentrated medium of SJW2536 harboring pVJMBPa. *Small triangles* point to the expressed fusion proteins.

3.4. Protein Expression

Protein expression and purification is demonstrated by the example of two model proteins: MBP (17) and CCP2 (complement control protein 2 domain of C1r) (18) in Figs. 2 and 3.

3.4.1. Cell Growth and Harvest

1. Pick up a single colony from the plate containing the electroporated cells and add 3 mL of sterile LB containing 100 μg/μL ampicillin. Grow the cells overnight at 37°C.
2. Inoculate sterile LB medium or minimal medium supplemented with 5% LB with 1% (v/v) overnight SJW2536 culture (e.g., add 200 μL of overnight culture to 20 mL medium), grow the cells at 37°C at 220 rpm until OD_{600} reaches approximately 1 (see Notes 6 and 7).
3. Centrifuge the culture at 3,000 × *g* for 10 min at 4°C. Separate the supernatant (medium) from the cells. Keep the supernatant.

3.4.2. Experimental Investigation of Export Efficiency

1. After harvesting the cells by centrifugation at 3,000 × *g* for 10 min as described in Subheading 3.4.1, and precipitate the medium by adding 80% TCA. The final concentration of TCA should be about 16%.
2. Incubate the solution for 1 h at 4°C.
3. Collect the precipitated proteins by centrifugation at 3,000 × *g* for 10 min. Carefully remove all TCA from the tube. Spin down again the tube containing the pellet (centrifuge at 3,000 × *g* for 2 min). Remove and discard the remaining TCA.

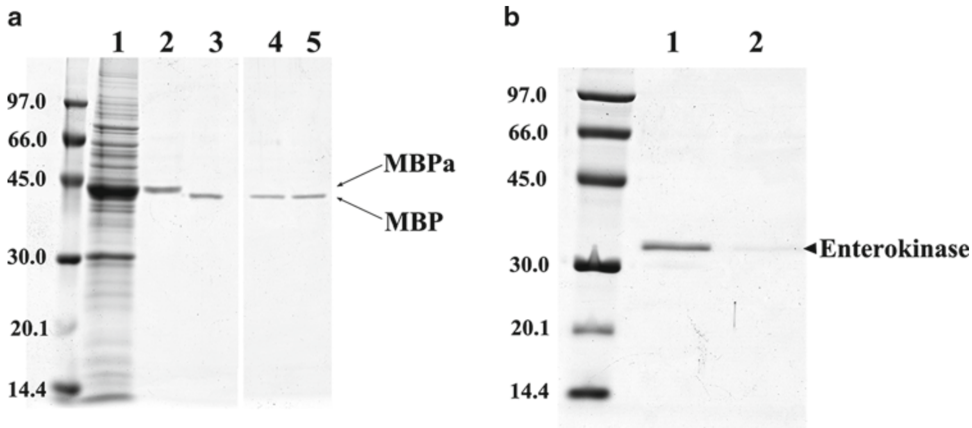


Fig. 3. Purification by Ni affinity chromatography and removal of the signal sequence. The MBPa fusion protein was purified from the medium by Ni^{2+}-affinity chromatography and cleaved by enterokinase (EK). If correctly folded the EK cleaved MBPa binds to amylose–agarose. The EK was removed by centrifugation with soybean trypsin inhibitor agarose. (**a**) Lane 1: concentrated medium (equivalent to 2 mL original medium) of SJW2536 harboring pVJMBPa. Lane 2: MBPa purified on Ni-NTA Superflow column. Lane 3: MBPa treated with EK (at MBPa:EK = 20:1 w/w ratio, 40 h). Lanes 4 and 5: fractions containing the EK-treated MBP eluted from amylose agarose beads with a maltose containing buffer. (**b**) Demonstration of EK removal. Lane 1: binding buffer containing EK. Lane 2: EK was removed by soybean trypsin inhibitor agarose. In one step approximately 90–95% of the EK was removed. The reaction does not contain MBP, because such large amount of EK that is visible on an SDS polyacrylamide gel would degrade it.

4. Dissolve the pellet in nonstandard SDS-PAGE sample buffer. The final volume of the dissolved proteins should be about 1/50 of the original medium (e.g., add 80 μL sample buffer to a pellet precipitated from 4 mL medium) (see Note 8). Heat the samples at 95°C for 5 min before analyzing them by SDS-PAGE.
5. Resuspend the harvested cells in 1/5 volume of sterile water then add equal volume of standard SDS-PAGE sample buffer (e.g., add 2 mL of water and 2 mL of standard SDS-PAGE sample buffer to cells harvested from 10 mL of medium). Lyse the cells by two cycles of heating at 95°C for 5 min and vortexing for 20 s. This ensures that the genomic DNA breaks.
6. Load 20–20 μL of the prepared cell lysate and medium samples on a 12% or 15% gel depending on the size of the protein of interest and analyze them by SDS-PAGE using a standard Tris–glycine buffer system (Fig. 2) (19) (see Note 9).

3.4.3. Calculating the Export Efficiency

1. Estimate the protein concentration within the cells and in the cell culture supernatant by applying densitometric analysis of the Coomassie-stained SDS polyacrylamide gels (Fig. 2) using an appropriate software or by simple visual estimation comparing the intensity to a known amount of protein (see Note 10).
2. Use the bands of the molecular weight marker as a reference to determine the amount of the protein in each lane (see Note 11).

3. Calculate the export efficiency from the level of the fusion protein in the cells and in the medium as follows (see Note 7):

$$\frac{\text{amount of protein in the medium}}{\text{amount of protein in the medium} + \text{amount of protein in the cells}} \times 100\%$$

3.5. Protein Purification by Ni-Affinity Chromatography

It is recommended to grow the cells for protein purification in minimal medium instead of LB because amino acids in LB bind to and remove Ni ions from the Ni-NTA resin. Addition of 5% LB to the minimal medium facilitates cell growth, but it does not significantly hinder purification by Ni-affinity chromatography. If cells are grown in pure LB, extensive dialysis is needed against large volumes of loading buffer before Ni-NTA purification.

1. Grow transformed SJW2536 cells as above in 1 L minimal medium supplemented with 5% LB.
2. Separate the cells by centrifugation at 5,000×*g* for 15 min at 4°C.
3. Pour a 5-mL (bed volume) Ni-NTA Superflow column using an Econo-Pac gravity-flow chromatography column.
4. Equilibrate the 5-mL Ni-NTA Superflow column with 20 mL of loading buffer.
5. Add NaCl, NaH_2PO_4 and imidazole to the sample in a final concentration of 400, 50, and 10 mM respectively.
6. Load 200 mL of the sample to the column.
7. Wash it twice with 10–10 mL of loading buffer to eliminate the nonbinding components.
8. Finally elute the protein of interest with 3×5 mL of elution buffer.
9. Analyze the eluted samples by SDS-PAGE, then combine the clear fractions containing the secreted protein (Fig. 3a).

3.6. Removal of the Signal Sequence by Enterokinase Cleavage

1. Dialyze the combined fractions against >1,000-fold volume of EK buffer (e.g., 400 μL of the fractions against 2 L buffer).
2. Add recombinant EK at about fusion protein: EK=20:1 to 40:1 w/w ratio (see Note 12) to the solution of 0.1–0.15 mg/mL protein, and incubate for 8–40 h at RT.
3. Remove EK by mixing with soybean trypsin inhibitor agarose resin followed by centrifugation. Use 30 μL of 50% slurry per 1 unit (0.05 mg) of enzyme.
4. Resuspend and wash the resin twice with 4× bed volume of 1× stripping buffer then with the same volume of 1× binding buffer.
5. After pelleting the resin by centrifuging for 1 min at 800×*g*, remove and discard the supernatant. Resuspend the resin to make a 50% slurry in binding buffer. Add 1/10 volume of 10× binding buffer to the EK digestion mixture.

6. Add the prewashed resin to the digestion mixture.
7. Keep the resin suspended in the reaction mixture by slowly rocking the tube or by gently pipetting up and down for 15 min.
8. Centrifuge the resin for 2 min at 5,000 × *g* and save the supernatant. The supernatant contains your protein.
9. Check the purity by SDS-PAGE analysis (Fig. 3b). If the EK is not totally removed from the reaction mixture, use larger amount of resin or add fresh prewashed resin to the protein solution containing EK, and repeat the rocking/pipetting step.
10. Apply the reaction mixture to a Ni-affinity column as above (Subheading 3.5, steps 3–6). The cleaved His_6-tags and the nondigested proteins bind to the Ni-resin. The purified cleaved protein will be in the flow-through fraction after loading the sample to the column.

3.7. Functional Assay

As previously mentioned, the secreted protein is not necessarily correctly folded, therefore, it is important to prove by appropriate structural or functional studies that the EK-digested protein is functional, and gains its correct three-dimensional structure. As an example, here, we present the functional assay for MBP, which, if is correctly folded, binds to an amylose resin.

1. Pour a 1-mL (bed volume) amylose resin column using a Poly-Prep gravity-flow chromatography column.
2. To check amylose binding, load 5 mL of the EK digested protein to the amylose resin column.
3. Wash the column twice to eliminate the nonbinding components with 2-mL of amylose wash buffer.
4. Elute the protein with 4 mL amylose eluting buffer. Collect 1 mL fractions.
5. Demonstrate by SDS-PAGE analysis, that the eluted fractions contain clean MBP, indicating that it is correctly folded (Fig. 3a).

4. Notes

1. More detailed protocols of basic DNA techniques can be found in (20), Sections 8.37–8.41.
2. The primers should contain at least four nucleotides at the end (e.g., GCGC sequence) then the recognition sequence of the chosen endonuclease, at last a 15 base pair long sequence matching the gene of interest. The melting point of the two primers should be about the same temperature.
3. Choose restriction enzymes that do not cut inside the gene of interest.

4. If 50 μL fractions do not result in appropriate number of colonies, streak larger volumes on the plates.
5. The cells in 15% glycerol can be stored at −80°C for years. According to our observations these cells are viable but not electro-competent. It is recommended to use freshly made electro-competent cells each time. If you do not need to store the cells for longer time, it is not necessary to add glycerol to the cells, they are electro-competent in water as well.
6. IPTG induction is not required, because *lac* repressor is not present in *Salmonella* cells.
7. If your protein is produced in the form of inclusion bodies it may cause low export efficiency. Lowering the growth temperature to 30–16°C might help.
8. If the protein solution dissolved in nonstandard SDS-PAGE sample buffer becomes yellow, add a minimal amount of Tris–base to the solution until it gains the original blue color of the sample buffer.
9. More detailed protocols of basic protein techniques can be found in (20), Sections A8.40–A8.50.
10. Correct the band intensities with the intensity of the host protein band with the same apparent molecular weight in the neighboring lane.
11. The detection limit in Coomassie-stained gels is about 1 mg/L fusion protein in the cells and 0.2 mg/L in the medium.
12. The optimal ratio of the required EK varies from protein to protein.

Acknowledgments

Support of the Hungarian Scientific Research Fund (OTKA) grant NK77978, the Hungarian National Office for Research and Technology (NKTH – KPI) grant GVOP-3.1.1-0386/3.0, and the National Development Agency grants TAMOP-4-2-2/B-10/1-2010-0025 and KMOP-1.1.2-07/1-2008-0003 are acknowledged.

References

1. Macnab, R. M. (1995) Flagella and motility. In *Eschericia coli* and *Salmonella*: Cellular and Molecular Biology (Neidhardt, F. C., Ingraham, J., Low, K. B., Magasnik, B., Schaechter, M. & Umbarger, M. E., eds), 2nd ed., pp. 123–145, American Society for Microbiology Publications, Washington, DC.
2. Namba, K. and Vonderviszt, F. (1997) Molecular architecture of bacterial flagellum. *Quart. Rev. Biophys.* **30**, 1–65.
3. Minamino, T., Imada, K., and Namba, K. (2008) Mechanisms of type III protein export for bacterial flagellar assembly. *Mol. Biosyst.* **4**, 1105–1115.

4. Macnab, R. M. (2004) Type III flagellar protein export and flagellar assembly. *Biochim. Biophys. Acta*. **1694**, 207–217.
5. Yonekura, K., Maki-Yonekura, S., and Namba, K. (2003) Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. *Nature* **424**, 643–650.
6. Minamino, T. and K. Namba. (2008) Distinct roles of the FliI ATPase and proton motive force in bacterial flagellar protein export. *Nature* **451**, 485–488.
7. Paul, K., Erhardt, M., Hirano, T., Blair, D. F., and Hughes, K. T. (2008) Energy source of flagellar type III secretion. *Nature* **451**, 489–492.
8. Blocker, A., Komoriya, K., and Aizawa, S-I. (2003) Type III secretion systems and bacterial flagella: insights into their function from structural similarities. *Proc. Natl. Acad. Sci. USA*. **100**, 3027–3030.
9. Journet, L., Hughes, K. T., and Cornelis, G. R. (2006) Type III secretion: a secretory pathway serving both motility and virulence. *Mol. Membr. Biol.* **22**, 41–50.
10. Kuwajima, G., Kawagishi, I., Homma, M., Asaka, J-I., Kondo, E., and Macnab, R. M. (1989) Export of an N-terminal fragment of *Escherichia coli* flagellin by a flagellum-specific pathway. *Proc. Natl. Acad. Sci. USA* **86**, 4953–4957.
11. Kornacker, M.G. and Newton, A. (1994) Information essential for cell-cycle-dependent secretion of the 591-residue Caulobacter hook protein is confined to a 21-amino-acid sequence near the N-terminus. *Mol. Microbiol.* **14**, 73–85.
12. Weber-Sparenberg, C., Pöplau, P., Brookman, H., Rochón, M., Möckel, C., Nietschke, M., and Jung, H. (2006) Characterization of the type III export signal of the flagellar hook scaffolding protein FlgD of *Escherichia coli*. *Arch. Microbiol.* **186**, 307–316.
13. Végh, B. M., Gál, P., Dobó, J., Závodszky, P., and Vonderviszt, F. (2006) Localization of the flagellum-specific secretion signal in *Salmonella* flagellin. *Biochem. Biophys. Res. Commun*. **345**, 93–98.
14. Gál, P., Végh, B. M., Závodszky, P., and Vonderviszt, F. (2006) Export signals. *Nat. Biotechnol.* **24**, 900–901.
15. Dobó, J., Varga, J., Sajó, R., Végh, B. M., Gál, P., Závodszky, P., and Vonderviszt, F. (2010) Application of a short, disordered N-terminal flagellin segment, a fully functional flagellar type III export signal, to expression of secreted proteins. *Appl. Env. Microbiol.* **76**, 891–899.
16. Ohnishi, K., Ohto, Y., Aizawa, S-I., Macnab, R. M., and Iino, T. (1994) FlgD is a scaffolding protein needed for flagellar hook assembly in *Salmonella typhimurium*. *J. Bacteriol.* **176,** 2272–2281.
17. Boos, W., and Shuman, H. (1998) Maltose/maltodextrin system of Escherichia coli: transport, metabolism, and regulation. *Microbiol Mol Biol Rev*. **62**, 204–229.
18. Láng, A., Szilágyi, K., Major, B., Gál, P., Závodszky, P., and Perczel, A. (2010) Intermodule cooperativity in the structure and dynamics of consecutive complement control modules in human C1r. *FEBS J.* **277**, 3986–3998.
19. Laemmli U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
20. Sambrook, J., and Russel, D.W. (2001) Molecular cloning. A laboratory manual. 3rd Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Chapter 7

Optimization of Purification Protocols Based on the Step-by-Step Monitoring of the Protein Aggregates in Soluble Fractions

Ario de Marco

Abstract

Soluble protein fractions are often considered containing exclusively monodispersed and correctly folded molecules. This is not the case, being soluble aggregates of different complexity widely represented in such fractions. The use of fusing target protein domains to highly soluble carriers may strongly contribute to soluble aggregate accumulation. Therefore, reliable analytical methods must be used to evaluate the biophysical characteristics of soluble proteins. On the other hand, conventional methodologies are often technically demanding and time consuming. In this method paper, a protocol is presented that enables the rapid evaluation of the protein monodispersity from the initial step aimed at screening several conditions in parallel to the setup of the complete protocol for large-scale purification. The analysis is performed by means of simple lab equipment and starting from small sample volumes.

Key words: Aggregation index, Light scattering, Magnetic beads, Size-exclusion chromatography, Soluble aggregates

1. Introduction

For a long period, solubility has been accepted as a sufficient parameter for considering a protein correctly folded. However, the repeated reports of the presence of aggregates in soluble fractions containing recombinant fusion proteins expressed in bacteria clearly urged for a more accurate analysis and monodispersity evaluation (1–3). The complexity of the soluble aggregate structures was demonstrated as well as the progressive development of larger aggregates starting from tiny aggregate seeds (3–7). This last feature implies that the conditions leading to aggregation must be

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_7, © Springer Science+Business Media, LLC 2012

prevented from the initial purification steps for avoiding a cumulative aggregation process. Ideally, the presence of soluble aggregates should be monitored in a reliable way at each protocol step to identify the conditions that are critical for preserving or preventing the protein native folding. Given the limiting conditions of time and sample volumes that are usual during the phases of a protocol setup, such analyses should be performed in a short period and without loss of material.

In the laboratory praxis, the development of a purification protocol usually consists of a first phase, in which several conditions (strains, constructs, growth conditions, osmolyte addition, foldase coexpression, etc.) are compared simultaneously in small scale, and of an advanced phase, in which the single steps of the scale-up protocol (lysis conditions, materials, chromatographic steps, buffers, etc.) are progressively tested. Recently, we demonstrated that it was feasible to analyze in 30 min and by size-exclusion chromatography (SEC) 20 μL of sample eluted by a screening-scale amount of magnetic beads and that larger volumes (140 μL) from preparative samples could be evaluated in a few minutes by the fluorimetric aggregation index (AI) assay (8, 9). The rapid analyses and consequent evaluation allows the fast choice among options and the step-by-step design of an optimized purification flowchart. It is important to underline that the biophysical analyses incorporated in this protocol are easily performed with the standard equipment available in a biology lab.

2. Materials

2.1. Small-Scale Cell Culture

1. Culture medium (Lauria Bertani – LB or Terrific Broth – TB).
2. I M IPTG dissolved in H_2O.
3. 40% glucose.
4. BL21 (DE3) bacteria transformed with the suitable constructs.

2.2. Small-Scale Protein Affinity Purification

1. 1 mg/mL DNase I.
2. 100 mg/mL lysozyme.
3. Ni-NTA magnetic beads.
4. Lysis buffer: 50 mM Tris–HCl, pH 8.0, 500 mM NaCl, and 5 mM $MgCl_2$.
5. Washing buffer: 50 mM Tris–HCl, pH 8.0, 500 mM NaCl, 15 mM imidazole, and 0.02% Triton.
6. PBS.
7. Water bath sonicator.

2.3. Small-Scale SEC

1. Superdex 75 5/150 GL (GE Healthcare).
2. ÄKTA-FPLC (GE Healthcare).
3. SEC buffer: 25 mM Tris–HCl, pH 7.8, 150 mM NaCl or PBS.

2.4. Optimization of the Protein Production Protocol

1. Spectrofluorometer.
2. Desalting cartridge (for instance, a HiTrap desalting column – GE Healthcare).
3. Dialysis device (see Note 1).

3. Methods

3.1. Small-Scale Cell Culture

1. Inoculate a 10-mL tube containing 2 mL of LB medium plus 1% glucose with the BL21(DE3) cells transformed with the target and grow overnight at 30°C in an inclined rack (30°) inside a shaker (180 rpm). Repeat the procedure for each construct.
2. The day after, add 100 μL of the precultures from each of the different constructs to a 50-mL Erlenmeyer flask filled with 15 mL of LB (or TB).
3. Let the bacteria grow at 37°C in an orbital shaker (210 rpm) until the OD_{600} reaches 0.4, then switch the temperature to 20°C, and after 30 min (the OD_{600} of the culture reaches approximately the value of 0.6) induce the protein expression with 0.2 mM IPTG.
4. Let the cultures grow for 18 h at 20°C and then harvest the pellet by centrifuging (15 min × 11,000 × *g* at room temperature).
5. Remove the medium and store the pellets at −20°C.

3.2. Small-Scale Protein Affinity Purification (See Notes 2 and 3)

1. For each sample to purify, add 35 μL of magnetic bead slurry to a 2-mL Eppendorf tube (see Note 4).
2. Set the tubes into a magnetic rack to separate the beads from the solution and carefully remove it.
3. Transfer the tube into a standard rack and wash the beads by resuspending them in 400 μL of PBS.
4. Set the tube to the magnetic rack and remove the buffer.
5. Transfer the tube in a standard rack, and recover the beads with 50 μL of lysis buffer.
6. Resuspend each of the bacteria pellets in 500 μL of lysis buffer (see Note 5).
7. Sonicate for 5 min in a water bath at room temperature.

8. Add lysozyme to a final concentration of 1 mg/mL and DNase I to a final concentration of 50 μg/mL and incubate for 20 min at room temperature by continuous rocking (see Note 6).
9. Separate the supernatant fractions by centrifugation (5 min at 16,100 × *g* at room temperature) and add them to the tubes with the pretreated beads.
10. Incubate the tubes for 30 min under constant rotation (see Note 7).
11. Separate the beads from the supernatant using the magnetic rack and discard the supernatant.
12. Remove the magnet, resuspend the beads with 400 μL of washing buffer, and incubate for 30 min under constant rotation.
13. Repeat steps 11 and 12.
14. Separate the beads from the supernatant by using the magnetic rack and carefully remove any remaining buffer.
15. Remove the magnet and elute the proteins by incubating the beads for 10 min in the presence of 40 μL of 50 mM Tris–HCl, pH 8.0, 250 mM NaCl, and 250 mM imidazole.
16. Separate the beads by means of the magnet and recover the elution fraction.
17. Load 20 μL of the affinity-purified protein fractions on the preequilibrated SEC column and use the remaining material for running a denaturing polyacrylamide gel.

3.3. Small-Scale SEC

1. Preequilibrate the SEC column in the appropriate buffer.
2. Load a mini-loop of 15 μL of volume.
3. Run the gel filtration at a flow rate of 0.2 mL/min and collect the absorbance signal at 280 nm (see Notes 8 and 9).
4. If available, collect the data from the multiangle light scattering (see Note 10).
5. The analysis of the peak distribution and intensity allows the identification of the expression conditions that assures the highest yield as shown in Fig. 1.

3.4. Optimization of the Protein Production Protocol

Once the optimal conditions for protein expression are identified by means of the small screening protocol, large amounts of bacteria can be prepared for scaling up the affinity purification step. It is possible to evaluate the monodispersity of the protein fractions starting from the material collected by elution of the affinity column until the final step yielding the sample having the requested purity and buffer conditions. Here is the method we use for calculating the aggregation index (see Note 11).

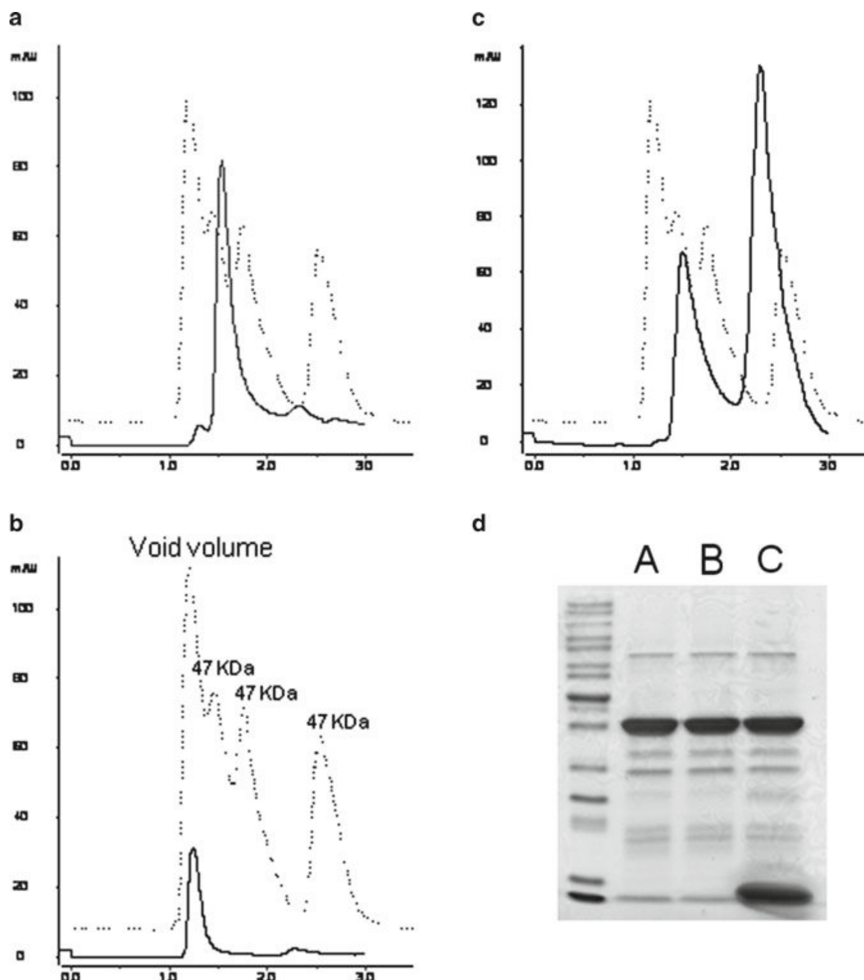


Fig. 1. Examples of gel filtration profiles of TAT–Cre samples with variable degree of aggregation and contamination. (**a**) Monodispersed protein. The two extra minimal peaks may correspond to polymers and a contaminant of low molecular mass, respectively. (**b**) Single peak eluted in the void volume, corresponding to large aggregates. (**c**) Peaks corresponding to the monodispersed protein and to a low-molecular-mass contaminant. (**d**) SDS-PAGE of the samples separated by gel filtration.

1. Set up the spectrofluorometer. Excitation at 280 nm and emission at 340 nm, scan rate 5, and emission recovery between 260 and 400 nm.
2. Insert the cuvette (140 μL of volume).
3. Sensitivity: 600 V can be used as a default, but lower values should be chosen for saturating signals. It allows recovering precise data for the final calculations.
4. Run and record the values at 280 and 340 nm.
5. Calculate the ratio between the values at 280 and 340 nm (see Note 12).

The critical steps during the purification protocol are strictly protein dependent. It is good practice to prepare aliquots

at the end of each step and to test them in parallel using variable conditions during the successive procedure steps. Different strategies and relative drawbacks are discussed taking the purification protocol of the TAT–Cre (Cre recombinase fused to a TAT peptide) as an example. In this case, alternative protocols yielded similar protein amounts, but in some cases the enzymatic activity was completely lost. We have realized that the activity loss correlated with the amount of soluble aggregation evaluated by AI analysis.

6. Cell lysis and affinity purification.
7. AI measurement: If the value is below 0.2, the sample can be divided in two aliquots and the buffer exchange step can begin. In case the AI value is higher, modify one by one the experimental parameters (lysis components, purification temperature, affinity resins, and conditions) and check the AI of the resulting affinity-purified fraction. In the case of TAT–Cre protein, the crucial factor was the chelating agent due to the high enzyme sensitivity to free metal ions.
8. One protein aliquot is buffer exchanged by means of a desalting cartridge and another one by using a two-step dialysis system. Compare the AI at the end of the process and consider the option with the lower AI. TAT–Cre was very sensitive to the desalting cartridge, but its monodispersity was preserved during dialysis.
9. Compare aliquots recovered from the buffer exchange step after a further purification step intended to eliminate endotoxin contaminations. According to the final use, the preference can be given to methods highly efficient in removing the contaminants or preserving the functional activity (see Note 13).
10. It is good norm to compare the storage buffer as well. Chemical chaperones, such as trehalose, betaine, or glycerol, can be added alone or in combination and at different concentrations before protein storage (see Note 14). After thawing, monodispersity can be evaluated by AI and the values compared with those of the samples before storage. Similarly, their stabilizing effect can be evaluated also at the conditions used for the experimental applications of proteins.

4. Notes

1. Dialysis membranes are available with pores of different diameter and must be chosen according to the mass of the target protein. It should be kept in mind that proteins with masses slightly larger than the theoretical exclusion size can slowly pass the barrier. Therefore, especially when only ions and other

small molecules must be exchanged and dialysis is not used as a cleaning step for diluting away low-mass peptides, it is safer using membranes with pores significantly smaller than theoretically acceptable. It is recommended to immerge the dialysis bag under slow stirring for 2 h at 4°C in a large backer containing the final buffer (dilution factor: at least 50 times) and then to replace the contaminated buffer with fresh one before incubating overnight under the same conditions.

2. This protocol has been conceived for proteins expressed in bacteria; however, it can be applied to eukaryotic systems as well. Times and volumes change according to the culture needs and, probably, different parameters are tested. For instance, the number of insect cell strains suitable for baculovirus expression is limited, but several combinations of constructs can be compared to identify the optimal conditions for protein complex production and copurification (10).
3. This section is an update version of a previously described protocol (11). Some variations have been introduced to increase volume and protein concentrations for obtaining samples suitable for the successive steps.
4. It is also possible to use any nonmagnetic bead product commercialized for immobilized metal-affinity chromatography (IMAC) purification. However, magnetic beads are more convenient since they allow more accurate washing steps with the consequent improved removal of contaminants (8).
5. The genes for the Lon and Omp proteases have been deleted in the BL21(DE3) strain. As a consequence, the use of protease inhibitors in the lysis buffer is not usually necessary.
6. At the end of this step, no viscous material indicating the presence of indigested nucleic acids should be detectable. Otherwise, it interferes with the next washing steps. Lysis conditions can be optimized by tuning DNase and $MgCl_2$ concentrations or using longer incubation time.
7. Several proteins can be manipulated at room temperature. However, 4°C can be necessary in the case of temperature-sensitive polypeptides.
8. We have noticed that large aggregates can apparently disappear during SEC. The most probable explanation is that aggregated protein is trapped by the filter above the column. According to the characteristics of the samples, SEC columns that separate in different mass ranges should be selected.
9. The peaks observed during SEC are not suitable for accurate protein quantification due to the sequence-specific absorbance that each single sequence shows at 280 nm. However, SEC can be used to evaluate relative peak variations (for instance, aggregated

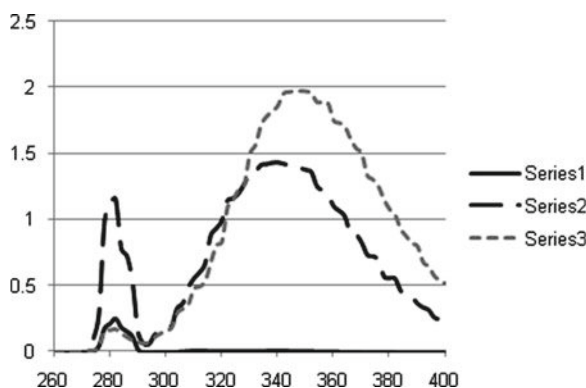


Fig. 2. Spectrofluorimetric profiles of TAT–Cre samples with different monodispersity. Three paradigmatic protein profiles are compared in which the ratio between light scattering contribution and emission at 340 nm varies from >40 (series 1) to 0.1 (series 3).

and monomeric) of a single protein. In Fig. 1, the sample *A* corresponds to monomeric and monodispersed protein. The sample *B* contains the same amount of protein according to the SDS-PAGE (*D*), but it is clearly composed by soluble aggregates that are in part blocked before entering the column and in part elute in the void volume. Sample *C* represents monodispersed target protein in the presence of high amount of low-molecular-weight contaminant.

10. SEC is already an informative analytical method. Furthermore, it offers the advantage of being easily coupled to an online multiangle light-scattering instrument. By such a way, SEC information is integrated with data about absolute molar mass and size of the proteins that allow inferring indications concerning conformational changes due to partial unfolding or misfolding (12).

11. The evaluation of the protein monodispersity can be performed by means of different approaches, such as gel filtration or turbidity analysis. However, the AI (2, 3, 9) is fast, reliable, does not waste material, and requires simple equipment. For these reasons, it is preferable for rapid comparisons of samples. In Fig. 2, three paradigmatic examples are reported. The fluorescence profile of sample 1 shows a unique signal at 280 nm (light scattering), but no light emission at 340 nm (contribution of accessible aromatic rings) indicating a completely aggregated sample. Sample 3, in contrast, is characterized by elevated emission at 340 nm in comparison to the light scattering contribution and is typical of monodispersed proteins. Sample 2 presents an intermediate fluorescence profile that suggests a nucleation phase of soluble aggregates.

12. The AI does not provide absolute values. Therefore, its maximal utility is obtained when samples originated from the same material and that underwent different treatments are analyzed in parallel to identify what conditions preserve the original structural conditions and which induce aggregation. However, a good approximation of the absolute monodispersity state can be obtained also by a direct analysis. In our experience, AI values lower than 0.2 correlate with stable and monodispersed protein fractions, as confirmed by gel filtration analyses.

13. In the case of TAT–Cre protein, the most efficient method for endotoxin removal was based on two-phase partition (13). However, it resulted in a dramatic increase of the AI index and completely inhibited the enzymatic activity. Therefore, alternative methods were preferred, although less efficient for endotoxin elimination. The sensitivity to single treatments is protein specific and, in case of TAT–Cre, the reliable functional test for confirming its activity was based on Cre-recombinase-induced cell morphological modifications. Since the readout of such an assay is available after 4–6 days, the advantage of having a method, such as the AI that allows the evaluation of the protein quality in a few minutes, is evident and absolutely necessary to take rapid decisions for planning the successive steps of the purification protocol.

14. For an exhaustive description of compounds known for their stabilizing effects on macromolecules, a good review is Roberts' work (14).

References

1. Sachdev, D., Chirgwin, J.M. (1999) Properties of soluble fusions between mammalian aspartic proteases and bacterial maltose-binding protein. *Biochem. J.* **338**, 77–81.
2. Nominé, Y., Ristriani, T., Laurent, C., et al., (2001) A strategy for optimizing the monodispersity of fusion proteins: application to purification of recombinant HPV E6 oncoprotein. *Prot. Engineer.* **14**, 297–305.
3. Schrödel, A., de Marco, A. (2005) Characterization of the aggregates formed during recombinant protein expression in bacteria. *BMC Biochem.* **6**, 10.
4. Martinez-Alonso, M., Gonzalez-Montalban, N., Garcia-Fruitos, E., et al., (2008) The functional quality of soluble recombinant polypetides produced in *Escherichia coli* is defined by a wide conformational spectrum. *Appl. Environ. Microbiol.* **74**, 7431–7433.
5. Martinez-Alonso, M., Garcia-Fruitos, E., Villaverde, A. (2008) Yield, solubility and conformational quality of soluble proteins are not simultaneously favored in recombinant *Escherichia coli*. *Biotechnol Bioeng.* **101**, 1353–1358.
6. Nataliello, A., Santarella, R., Doglia, S.M., et al., (2008) Physical and chemical perturbations induce the formation of protein aggregates with different structural features. *Protein Expr. Purif.* **58**, 356–361.
7. Nataliello, A., Liu, J., Ami, D., et al., (2008) The osmolyte betaine promotes protein misfolding and disruption of protein aggregates. *Proteins* **75**, 509–517.
8. Sala, E., de Marco, A. (2010) Screening optimized protein purification protocols by coupling small-scale expression and mini-size exclusion chromatography. *Prot. Expr. Purif.* **74**, 231–235.
9. Capasso, P., Aliprandi, M., Ossolengo, G., et al., (2009) Monodispersity of recombinant Cre recombinase correlates with its effectiveness *in vivo*. *BMC Biotechnol.* **9**, 80.

10. Berger, I., Fitzgerald, D.J., Richmond, T.J. (2004) Baculovirus expression system for heterologous multiprotein complexes. *Nat. Biotechnol.* **22**, 1583–1587.
11. de Marco, A. (2007) Protocol for preparing proteins with improved solubility by co-expression with molecular chaperones in *Escherichia coli. Nature Protocols* **2**, 2632–2639.
12. Wen, J., Arakawa, T., Philo, J.S. (2004) Size-exclusion chromatography with on-line light-scattering, absorbance, and refractive index detectors for studying proteins and their interactions. *Anal. Biochem.* **240**, 155–166.
13. Liu, S., Tobias, R., McClure, S., et al., (1997) Removal of endotoxin from recombinant protein preparations. *Clinical Biochem.* **30**, 455–463.
14. Roberts, M.F. (2005) Organic compatible solutes of halotolerant and halophilic microorganisms. *Saline Sys.* **1**, 5.

Chapter 8

Heterologous Protein Expression by *Lactococcus lactis*

Julio Villatoro-Hernández, Oscar P. Kuipers, Odila Saucedo-Cárdenas, and Roberto Montes-de-Oca-Luna

Abstract

This chapter describes the use of *Lactococcus lactis* as a safe and efficient cell factory to produce heterologous proteins of medical interest. The relevance of the use of this lactic acid bacterium (LAB) is that it is a noncolonizing, nonpathogenic microorganism that can be delivered in vivo at a mucosal level. The use of strains of *L. lactis* in clinical trials in humans to alleviate inflammatory bowel diseases has opened up the possibility of using this same LAB to target other diseases.

Several crucial aspects are addressed in this chapter, such as the expression of heterologous protein, subcellular compartment into which the heterologous protein is located, and description of a standardized protocol to process samples in cell and cell-free fractions to detect the targeted protein expressed by *L. lactis*.

Key words: Antigen, Expression, Cellular location, Heterologous proteins, *Lactococcus lactis*

1. Introduction

The last decade has seen the metamorphosis of *Lactococcus lactis*, a lactic acid bacterium (LAB) that has moved from being solely considered as a microorganism for food fermentations to a useful resource to develop biopharmaceuticals. This includes the administration in vivo of genetically modified *L. lactis* and the production of molecules of medical interest by this LAB (1).

There are several reasons that make *L. lactis* an excellent candidate for the development of vaccines, e.g., it is noninvasive and noncolonizing, it can be safely administered to humans, and the genome of *L. lactis* is currently fully sequenced and it can be easily manipulated (2).

The development of several vectors specifically designed to express heterologous proteins in *L. lactis* has given us the possibility

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_8, © Springer Science+Business Media, LLC 2012

of producing a wide variety of proteins of different origins in *L. lactis*, either prokaryotic or eukaryotic. The expression of the target protein in *L. lactis* may be regulated by inducible or strong constitutive promoters. Among the most used promoters in *L. lactis* is the nisin-inducible promoter Pnis, which is the key component of the nisin-inducible controlled expression (NICE) system (3). These plasmids are designed to insert the DNA-coding sequence of the protein of interest in a relatively easy manner.

In this chapter, we describe the most recurring vectors used to express proteins in *L. lactis*, the different cellular compartments in which the heterologous protein can be located, and methods for the detection of these proteins.

1.1. Criteria to Be Considered Before Grabbing the Pipette

In the past 5 years, *L. lactis* has transcended from being considered a food-fermenting microorganism to a biopharmaceutical product and a safe vaccine. This achievement has also provided biotechnological tools to manipulate *L. lactis* to express a variety of molecules in different cellular compartments. Table 1 summarizes the criteria to be thoroughly considered to design an efficient construct.

1.2. Heterologous Gene: Plasmid Expression or Chromosomal Integration?

Depending on the requirements of the researcher, the heterologous gene coding for the target protein can be either expressed in a plasmid or engineered and integrated into the *L. lactis* chromosome. There are pros and cons for each choice. Using plasmid DNA is relatively easy because of a number of factors, e.g., small size, replication in *E. coli*, multiple cloning sites, and easy manipulation. These factors expedite the construct in general. Plasmids may be lost if the host strain is not kept properly and the levels of expression may vary depending on the plasmid copy number. Integrating the DNA into the *L. lactis* chromosome could be advantageous because it can provide a controlled number of genes present in the chromosome (4). Replacing the gene of interest by an essential gene of *L. lactis* would make this strain sensitive to the lack of a substrate. This strain is regarded as biologically contained and, for clinical

Table 1
Criteria to be considered in heterologous protein expression in L. lactis

Criteria	Considerations
Localization of heterologous gene	Extrachromosomal (plasmid) or integrated to chromosome (fixed number of copies)
System of expression	Constitutive or inducible (promoter)
Protein location in *L. lactis*	Cytoplasmic, extracellular, or associated to cell wall
Biological contingency of the genetically manipulated *L. lactis*	Deletion of essential genes, auxotroph strains

trials in humans, could guarantee the controlled proliferation of the strain once it is administered to humans. The major drawback of this system is the lengthy construction time.

1.3. To Be Inducible or Constitutive, that Is the Question

There are several research groups that differ about this basic decision; allow the protein expression in *L. lactis* under an inducible or a constitutive manner. This decision should be addressed in the requirements of the heterologous protein itself. The nature of the protein, e.g., toxicity, folding, or solubility, may affect its expression in *L. lactis*. It is possible that the constitutive expression of a protein could affect its metabolism, possibly making *L. lactis* generate mutations or inactivating the expression of the coding sequence; for this, a strong regulatory promoter is preferred. For the patented strain of *L. lactis* able to produce IL-10 used in clinical trials in humans to treat inflammatory diseases of the digestive tract, the constitutive expression was preferred, which appears to be unaffected overall, with the advantage of being a drug-resistant, gene-free system (4).

1.4. Cellular Location of the Target Protein: Where Should It Go?

The use of *L. lactis* as a microbial cell factory has transcended into a potential biopharmaceutical to cause immune responses and be used as a potential vaccine (5). For this reason, it is important to consider the cellular compartment of the expressing cell as the most adequate to address the target protein. Depending on the nature of the target protein, this can be retained intracellularly, secreted extracellularly, or associated with the cell wall.

For cytokines, the IL-10, IL-2, IP-10, and lymphotactin, secretion was the preferred choice (4, 6–8). Because of the nature of the cytokine, it is important that it is released to the extracellular environment and allowed to exert its activity. To achieve this, the protein is usually fused to a signal peptide (preferentially from *L. lactis*, i.e., the signal peptide from *usp*45). This signal peptide addressed the protein to the membrane and it has a motif that is cut from the protein. The result is secretion of the mature protein without the signal peptide retaining its biological activity (9) (Fig. 1).

For antigens, there are several debates among researchers. A strain of *L. lactis* able to intracellularly express the PspA surface protein from *Streptococcus pneumoniae* has been reported (10). This strain was used to immunize mice intranasally. After some weeks, the mice were challenged with the causal agent of pneumonia (*S. pneumoniae*) resulting in their protection against this pathogen. It was proven that the immunization with a strain of *L. lactis* able to intracellularly produce and retain the antigen PspA of *S. pneumoniae* resulted in protection against the causal agent of pneumonia, *S. pneumoniae* (10).

Another study analyzed the antigenicity of three different strains of *L. lactis* able to secrete the same antigen in different cellular compartments: (1) intracellular, (2) extracellular, and (3) attached

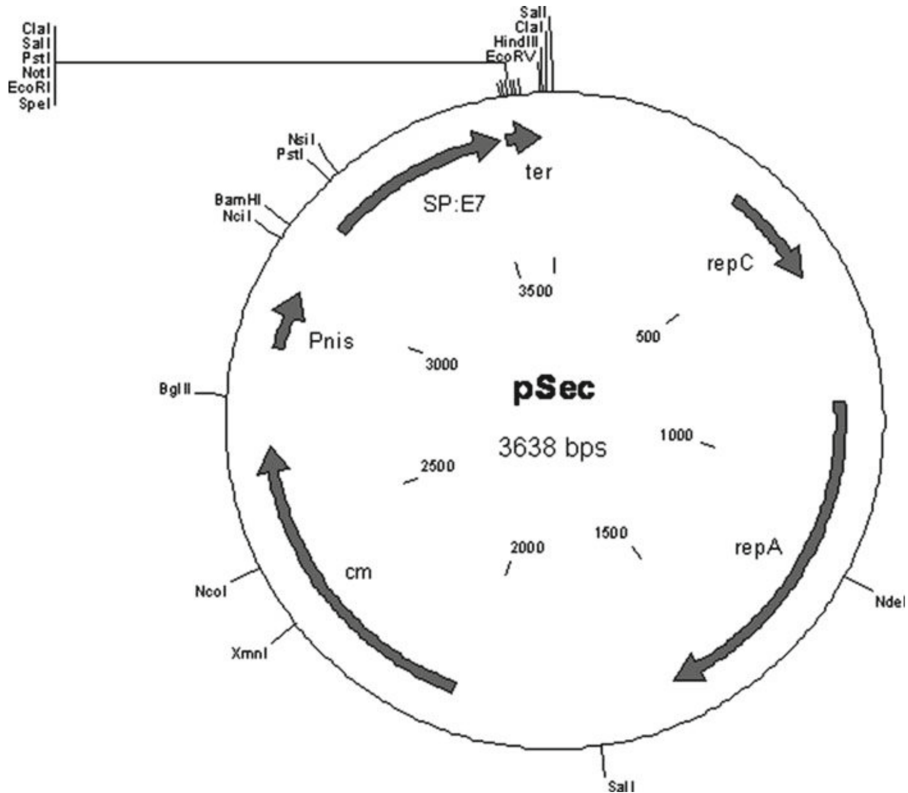


Fig. 1. Scheme of plasmid DNA used to secrete heterologous proteins in *L. lactis*. The figure represents the map of pSEC, a plasmid that replicates in *L. lactis* (*repA* replication origin). This high-copy number plasmid allows secretion of the heterologous protein generated by the fused in-frame *usp*45 signal peptide, the most secreted protein of *L. lactis.*

to the cell wall. The strain that had the E7 antigen attached to the cell wall produced the highest number of immune mice when they were challenged with a cell line derived from tumors that express E7 (therapeutic effect), even when each strain significantly expressed the same quantity of antigen (11) (Fig. 2). It is clear that for some proteins, such as cytokines or antibodies, the secretion path is the best choice, but for antigens and the creation of vaccines, there are decisions that need to be made before choosing the path to be followed.

1.5. Expressing Heterologous Proteins in L. lactis

After developing the construct that allows the expression of the target protein in *L. lactis*, assessment of the protein production must be addressed. Western blot analysis can determine qualitatively and quantitatively the production of the target protein. For this, several factors must be considered. Depending on the subcellular compartment where the protein is located, samples must be prepared.

Figure 3 shows the results of an SDS-PAGE and Western blot to detect a heterologous protein produced by *L. lactis*. Here, the *L. lactis* strain NZ9000 was transformed with a plasmid that contains the coding sequence of the usp45 signal peptide fused in-frame

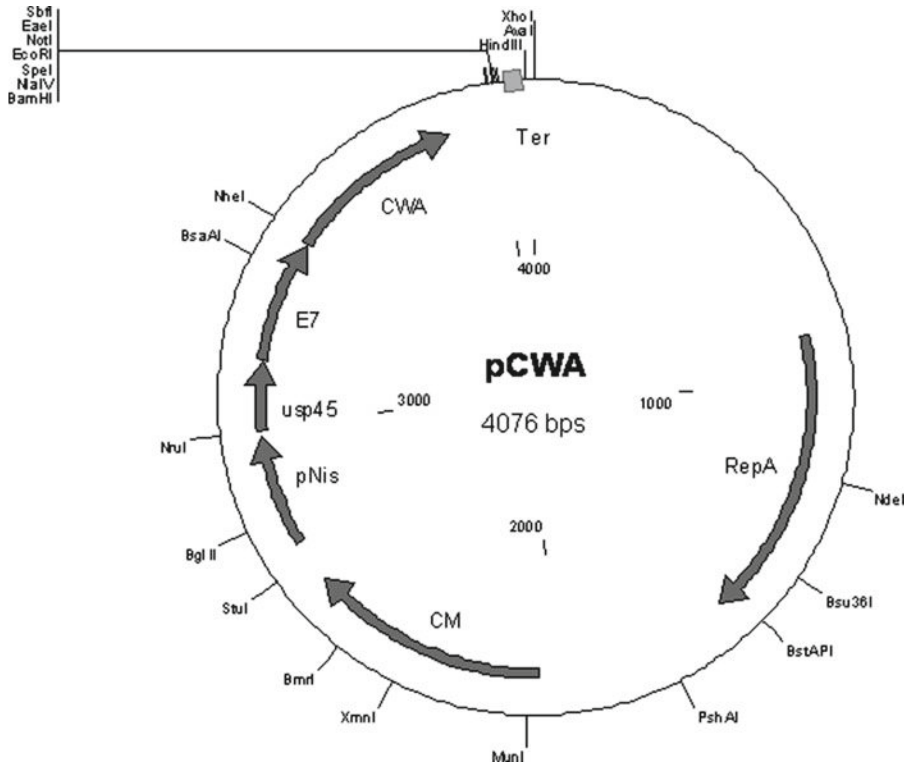


Fig. 2. Scheme of plasmid DNA used to express heterologous proteins in *L. lactis* attached to the cell wall. This plasmid contains the same features as the pSEC, plus the retention signal of the M6 protein of *Streptococcus pyogenes* (CWA). This protein associates with the cell wall of *L. lactis*, retaining the heterologous protein at this location and exposing it extracellularly.

to the coding sequence of the mutated E7 (E7m) of the human papilloma virus type-16 (HPV-16). Thus, the protein is expected to be secreted and found in the culture medium. Cell and cell-free (using a cell-free medium) samples were prepared and analyzed by SDS-PAGE and Western blot using specific anti-E7 polyclonal antibodies.

The electrophoresed gel with the samples was stained with a solution containing 0.25% Coomassie blue (Fig. 3a). Even though the heterologous protein is not visible by this type of staining in the gel, in the cell-free medium the mature protein usp45 (indicated with arrow) is visible, which is the most abundant protein secreted by *L. lactis*.

The developed Western blot (Fig. 3b) shows a clear single band in the sample corresponding to the strain transformed with the E7-expressing plasmid. This demonstrates the efficiency of the usp45 signal peptide because of the evident lack of other protein traces in the cell fraction (pSEC-E7m, cell). This same secretion efficiency has been observed in strains of *L. lactis* engineered to secrete the chemokine IP-10 (human and murine) and lymphotactin (7, 8, 12).

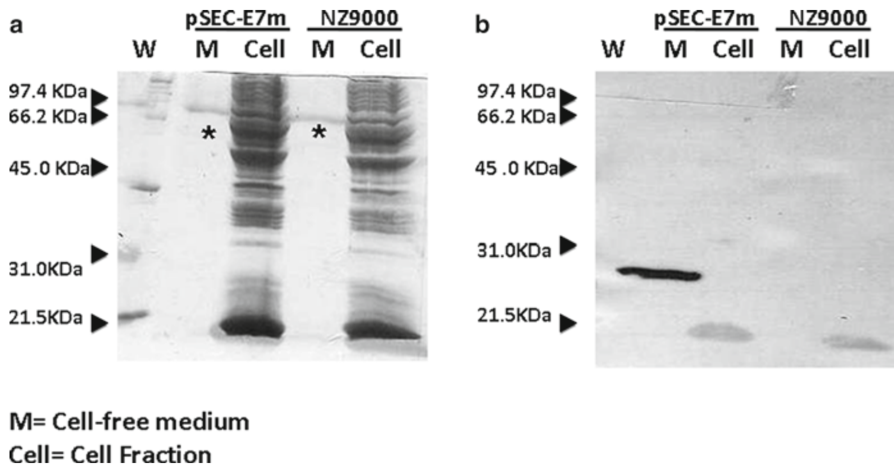


Fig. 3. SDS-PAGE and Western blot (WB) for one heterologous protein expressed by *L. lactis*. The wild-type and pSEC-E7m strains were analyzed by SDS-PAGE and WB. The pSEC-E7m strain was engineered to secrete the E7 protein of the human papilloma virus type-16. Cell and cell-free samples were analyzed. Samples from panel (**a**), from both strains, show a protein of approximately 66 kDa in the cell-free medium sample (M), corresponding to the usp45, the most secreted protein of *L. lactis* (indicated with asterisks). Panel (**b**) reveals a clear single band only in the cell-free medium fraction from the pSEC-E7 strain. This demonstrates that the E7 is efficiently secreted extracellularly because no signal is detected in the cell fraction. No band was detected in the wild-type strain.

In the following section, we describe a standardized protocol for both the detection of proteins located intra- and extracellularly by focusing on the preparation of the sample.

2. Materials

1. THMS buffer: 30 mM Tris–HCl, pH 8, 3 mM $MgCl_2$, 25% sucrose. Add 2 mg/mL lysozyme before use.
2. Wash buffer: 0.5 M sucrose and 10% glycerol.
3. M17 broth: 5 g pancreatic digest of casein, 5 g soy peptone, 5 g beef extract 2.5 g yeast extract, 0.5 g ascorbic acid, 0.25 g magnesium sulfate, and 19 g sodium β-glycerophosphate dissolved in 950 mL of water.
4. G-L-SGM17B: M17 broth containing 0.5 M sucrose, 2.5% glycine, and 0.5% glucose or 0.5% lactose (depending on the strain). Add the sucrose and glycine to the M17 and sterilize for 20 min at 121°C. Add sterile glucose or lactose after cooling.
5. 1,000× nisin stock solution: 1 mg/mL nisin in 0.05% acetic acid. Store at –20°C. Prepare work solution in water and dispose after use. Nisin is unstable when diluted on water.
6. TES-lysis buffer: 25% sucrose, 1 mM EDTA, 50 mM Tris–HCl, pH 8.0, and 10 mg/mL lysozyme supplemented with

1 mM phenylmethanesulfonyl fluoride (PMSF) and 10 mM dithiothreitol (DTT).

7. SDS-loading buffer: 100 mM Tris–HCl, pH 6.8, 200 mM DTT, 4% SDS, 0.1% bromophenol blue, and 10% glycerol.
8. 3 M potassium acetate, pH 5.5.
9. Ice-cold plasmid DNA wash solution: 30% mili-q water–70% ethanol.
10. Competent cells wash solution: 0.5 M sucrose, 10% glycerol.
11. 30% acrylamide–Bis solution 19:1.
12. TBS-Tween 20: 10× TBS (0.5 M Tris base, 9% NaCl, pH 8.4): Dilute this 1:10 and add 1% Tween 20.
13. Antibodies: Primary antibodies should be specific for the heterologous protein of interest and a proper secondary antibody must be selected based on the source of primary antibody (anti-goat, anti-rabbit, etc.).
14. PVDF membrane from BioRad.
15. Bacterial strains: The most frequently used strain is the NZ9000, which has cloned the *NisRK* genes chromosomically (3). This strain must be used when the NICE system is being used to express the heterologous protein. Wild-type strain MG1363 can also be used (does not contain the *NisRK* genes).
16. Vectors: This is a list of some vectors that can be used to express proteins (Figs. 1 and 2):
    (a) pVE3684 (nisin-inducible promoter) (3).
    (b) pMG36 (p32 constitutive promoter) (13).
    (c) pVE3684:CWA (M6 cell wall-attachment signal) (11).
17. Electroporation device: Bio-Rad Gene Pulser.
18. Electroporation cuvettes (e.g., Btx Inc).
19. Protein electrophoresis in acrylamide gels: Bio-Rad Mini Protean 3 cell.
20. Immunoblotting apparatus: Bio-Rad Semi-Dry Electrophoretic Transfer Cell.
21. Bio-Rad Power Pac 300 (Power supply).

3. Methods

3.1. Culturing L. lactis and Protein Production

L. lactis is a facultative anaerobic microorganism. It can be grown in several media, but it is important to complement it with a carbon source. The media normally used to culture *L. lactis* can be commercially acquired. MRS (*Lactobacillus* deMan Rogosa Sharpe) and brain heart infusion (BHI) can be used without the need to

add any carbon source, but the most-used medium is M17 (see Note 1). When M17 is prepared, 0.5% glucose (or lactose) should be added (see Note 2).

3.2. Plasmid DNA Isolation from L. lactis: Miniprep

1. Use 5 mL of a fully grown culture of *L. lactis*, preferably from an overnight culture.
2. Centrifuge for 10 min at 8,000 × *g* at room temperature (RT).
3. Resuspend pellet in 250 μL THMS buffer in an Eppendorf tube.
4. Incubate for 10 min at 37°C.
5. Add 500 μL 0.2 N NaOH and 1% SDS, and mix carefully by inversion (do not vortex).
6. Incubate for 5 min on ice.
7. Add 375 μL ice-cold 3 M potassium acetate, pH 5.5, and mix carefully by tapping.
8. Incubate for 5 min on ice.
9. Centrifuge for 5 min at 20,000 × *g* preferably at 4°C.
10. Recover supernatant and add it into a new Eppendorf tube.
11. Add 1 volume of 2-propanol.
12. Incubate for 5–10 min at RT. Alternatively, this could be also done in a −20°C freezer.
13. Centrifuge for 10 min at 20,000 × *g* at 4°C.
14. Wash pellet carefully with 500 μL of 70% ethanol.
15. Air dry the pellet for 10 min at RT.
16. Dissolve the pellet in 50 μL of 1× TE buffer or sterile water.

3.3. Preparation of Electrocompetent L. lactis Cells

The procedure is done over 3 days. Be sure to prepare all the solutions and the medium before starting.

1. On day 1, inoculate the strain of preference in 5 mL of G-L-SGM17B from a −80°C stock. Grow at 30°C overnight.
2. On day 2, add 5 mL of the previous culture in 50 mL G-L-SGM17B. Grow at 30°C (this step may be done overnight) (see Note 3).
3. On day 3, dilute the 50 mL culture in 400 mL G-L-SGM17B.
4. Grow until OD_{600} is 0.2–0.3 (approximately for 3 h) (see Note 4).
5. Centrifuge the cells for 20 min at 6,000 × *g*, at 4°C.
6. Wash cells with 400 mL of ice-cold wash solution.
7. Centrifuge the cells for 20 min at 6,000 × *g*, at 4°C.
8. Keep the cells on ice for 15 min in 200 mL 0.5 M sucrose, 10% glycerol, and 0.05 M EDTA at 4°C.
9. Centrifuge the cells for 20 min at 6,000 × *g*, at 4°C.

10. Wash cells with 100 mL ice-cold wash buffer.
11. Centrifuge the cells for 20 min at 6,000 × *g*, at 4°C.
12. Resuspend the cells in 4 mL 0.5 M sucrose and 10% glycerol (on ice). Cells may be stored at −80°C.

3.4. Transformation of L. lactis by Electroporation

1. Place 40-μL cells in a prechilled electroporation cuvette with at least 20 ng of DNA resuspended in no more than 5 μL of distilled water. Keep the cuvette on ice.
2. Use the electroporator with the following settings: 2,000 V, 25 μF, and 200 Ω.
3. Pulse: The "normal" reading you should get is 4.5–5 ms.
4. Immediately add 1 mL G-L-M17B, with 20 mM $MgCl_2$ and 2 mM $CaCl_2$.
5. Keep the cuvette for 5 min on ice and then incubate for 1–1.5 h at 30°C without shaking.
6. Plate 10, 100, and 900 μL of cells on M17 agar with glucose or lactose and antibiotics depending on the plasmid.
7. Incubate for 1–2 days at 30°C.

3.5. Working with NICE System

1. Grow 5 mL of an overnight culture of the interest strain at 30°C in GM17 with its respective antibiotic.
2. Add 500 μL of the previous culture to 10 mL GM17 fresh medium and grow cells at 30°C without shaking (see Note 5).
3. Grow to an OD_{600} 0.6–0.8, but avoid going over OD_{600} 0.8.
4. Add 1 ng/mL nisin A (Sigma) to one of the tubes containing 10 mL of culture. Use the other 10 mL as a negative control.
5. Incubate for 2–3 h, and measure the OD_{600} (see Note 6).
6. Pellet cells by centrifugation at 13,000 × *g* for 3 min at 4°C. Process the cell and cell-free samples separately. Keep them on ice.

3.6. Protein Analysis of Cellular Fractions

1. Obtain cell pellet from 1.35 mL of culture from previous section.
2. Centrifuge at 21,000 × *g* at 4°C for 5 min.
3. Resuspend cells with100 μL of TES-lysis buffer.
4. Incubate at 37°C for 1 h.
5. Add 50 μL of 20% SDS to break the cells.
6. Add the proper amount of loading buffer for further SDS-PAGE and Western blot analysis (see Note 7).

3.7. Protein Analysis from Cell-Free Samples (Medium)

In this section, we describe a protocol for protein precipitation from the cell-free medium using trichloroacetic acid (TCA). Other methods can be used, but are not described in this chapter.

1. Centrifuge 5 mL of the protein-containing culture at 21,000 × *g* at 4°C for 5 min.
2. Carefully collect the supernatant and filter it through a 0.22-μM filter to ensure a cell-free sample.
3. Add to 1.35 mL of the previous supernatant 150 μL of 100% TCA and incubate on ice for 10 min.
4. Centrifuge at 21,000 × *g* at 4°C for 15 min.
5. Carefully decant supernatant completely and air dry for 5 min at RT.
6. Proteins are contained in the pellet.
7. Resuspend pellet with 50 μL of 50 mM NaOH and 50 μL of loading buffer. Use different volumes for further SDS-PAGE and Western blot analysis.
8. Add 1 mM of PMSF and 10 mM DTT to avoid proteolysis.
9. Boil the samples at 100°C for 5 min.
10. Place the sample on ice until they are loaded on the gel.

3.8. SDS-PAGE and Western Blotting

1. Prepare 12–15% gels from 30% acrylamide–Bis solution 19:1 (v/v).
2. Load 20 μL of the previous samples and separate by SDS-PAGE at 100 V (constant for about 60 min).
3. Use PVDF membranes to transfer the proteins at 80 mA (constant) for 1 h.
4. Block the membrane with TBS-Tween 20 solution containing 10% skim milk.
5. Dilutions and incubation of the first and second antibodies are made as suggested by the manufacturer.
6. Develop signal according to the selection of secondary antibody.

4. Notes

1. Use the M17 medium to grow *L. lactis* on a regular basis; MRS and BHI media are used for mutant (sick) strains that are not able to grow on M17.
2. Use glucose for routine growth of *L. lactis*. After sterilization, add 0.5% of sterile glucose and mix. If the use of lactose is desired, the strain must have the *Lac* operon.
3. The culture grows in the presence of glycine. Glycine inhibits the synthesis of peptidoglycan, so the culture grows at a much lower rate. This step can be done overnight.
4. DO NOT allow the culture to exceed this OD_{600}. If it does, diluting the culture decreases the competence of the cells.

5. This amount of culture can be increased if necessary.
6. The cells might grow more slowly depending on the nature of the heterologous protein and if it is secreted or kept intracellular.
7. Keep samples on ice to avoid proteolysis.

References

1. Wells, J. M., and Mercenier, A. (2008) Mucosal delivery of therapeutic and prophylactic molecules using lactic acid bacteria. *Nat. Rev. Microbiol.* **6**, 349–362.
2. Van Huynegem, K., Loos, M., and Steidler, L. (2009) Immunomodulation by genetically engineered lactic acid bacteria. *Front. Biosci.* **14**, 4825–4835.
3. de Ruyter, P. G., Kuipers, O. P., and de Vos, W. M. (1996) Controlled gene expression systems for *Lactococcus lactis* with the food-grade inducer nisin. *Appl. Environ. Microbiol.* **62**, 3662–3667.
4. Steidler, L., Neirynck, S., Huyghebaert, N., Snoeck, V., Vermeire, A., Goddeeris, B., Cox, E., Remon, J. P., and Remaut, E. (2003) Biological containment of genetically modified *Lactococcus lactis* for intestinal delivery of human interleukin 10. *Nat. Biotechnol.* **21**, 785–789.
5. Bermudez-Humaran, L. G., Langella, P., Cortes-Perez, N. G., Gruss, A., Tamez-Guerra, R. S., Oliveira, S. C., Saucedo-Cardenas, O., Montes de Oca-Luna, R., and Le Loir, Y. (2003) Intranasal immunization with recombinant *Lactococcus lactis* secreting murine interleukin-12 enhances antigen-specific Th1 cytokine production. *Infect. Immun.* **71**, 1887–1896.
6. Steidler, L., Robinson, K., Chamberlain, L., Schofield, K. M., Remaut, E., Le Page, R. W., and Wells, J. M. (1998) Mucosal delivery of murine interleukin-2 (IL-2) and IL-6 by recombinant strains of *Lactococcus lactis* coexpressing antigen and cytokine. *Infect. Immun.* **66**, 3183–3189.
7. Villatoro-Hernandez, J., Loera-Arias, M. J., Gamez-Escobedo, A., Franco-Molina, M., Gomez-Gutierrez, J. G., Rodriguez-Rocha, H., Gutierrez-Puente, Y., Saucedo-Cardenas, O., Valdes-Flores, J., and Montes-de-Oca-Luna, R. (2008) Secretion of biologically active interferon-gamma inducible protein-10 (IP-10) by *Lactococcus lactis. Microb. Cell Fact.* 7, 22.
8. Zavala-Flores, L. M., Villatoro-Hernandez, J., Gamez-Escobedo, A., Franco-Molina, M., Rangel-Colmenero, B. R., Villanueva-Olivo, A., Gutierrez-Puente, Y., de Oca-Luna, R. M., Valdes-Flores, J., and Saucedo-Cardenas, O. (2009) Production of biologically active human lymphotactin (XCL1) by *Lactococcus lactis. Biotechnol. Lett.* **31**, 215–220.
9. van Asseldonk, M., de Vos, W. M., and Simons, G. (1993) Functional analysis of the *Lactococcus lactis* usp45 secretion signal in the secretion of a homologous proteinase and a heterologous alpha-amylase. *Mol. Gen. Genet.* **240**, 428–434.
10. Hanniffy, S. B., Carter, A. T., Hitchin, E., and Wells, J. M. (2007) Mucosal delivery of a pneumococcal vaccine using *Lactococcus lactis* affords protection against respiratory infection. *J. Infect. Dis.* **195**, 185–193.
11. Bermudez-Humaran, L. G., Cortes-Perez, N. G., Le Loir, Y., Alcocer-Gonzalez, J. M., Tamez-Guerra, R. S., de Oca-Luna, R. M., and Langella, P. (2004) An inducible surface presentation system improves cellular immunity against human papillomavirus type 16 E7 antigen in mice after nasal administration with recombinant lactococci. *J. Med. Microbiol.* **53**, 427–433.
12. Villatoro-Hernandez, J., Arce-Mendoza, A. Y., Rosas-Taraco, A. G., Esparza-Gonzalez, S. C., Guzman-Lopez, S., Elizondo-Omana, R. E., Chavez-Reyes, A., Saucedo-Cardenas, O., and de Oca Luna, R. M. (2009) Murine interferon-gamma inducible protein-10 (IP-10) secreted by *Lactococcus lactis* chemo-attracts human CD3+ lymphocytes. *Biotechnol. Lett.* **31**, 1795–1800.
13. van de Guchte, M., van der Vossen, J. M., Kok, J., and Venema, G. (1989) Construction of a lactococcal expression vector: expression of hen egg white lysozyme in *Lactococcus lactis* subsp. *lactis. Appl Environ Microbiol* **55**, 224–228.

Chapter 9

An Extended Suite of Genetic Tools for Use in Bacteria of the *Halomonadaceae*: An Overview

Montserrat Argandoña, Carmen Vargas, Mercedes Reina-Bueno, Javier Rodríguez-Moya, Manuel Salvador, and Joaquín J. Nieto

Abstract

Halophilic gammaproteobacteria of the family *Halomonadaceae* (including the genera *Aidingimonas*, *Carnimonas*, *Chromohalobacter*, *Cobetia*, *Halomonas*, *Halotalea*, *Kushneria*, *Modicisalibacter*, *Salinicola*, and *Zymobacter*) have current and promising applications in biotechnology mainly as a source of compatible solutes (powerful stabilizers of biomolecules and cells, with exciting potentialities in biomedicine), salt-tolerant enzymes, biosurfactants, and extracellular polysaccharides, among other products. In addition, they display a number of advantages to be used as cell factories, alternative to conventional prokaryotic hosts like *Escherichia coli* or *Bacillus*, for the production of recombinant proteins: (1) their high salt tolerance decreases to a minimum the necessity for aseptic conditions, resulting in cost-reducing conditions, (2) they are very easy to grow and maintain in the laboratory, and their nutritional requirements are simple, and (3) the majority can use a large range of compounds as a sole carbon and energy source. In the last 15 years, the efforts of our group and others have made possible the genetic manipulation of this bacterial group. In this review, the most relevant and recent tools for their genetic manipulation are described, with emphasis on nucleic acid isolation procedures, cloning and expression vectors, genetic exchange mechanisms, mutagenesis approaches, reporter genes, and genetic expression analyses. Complementary sections describing the influence of salinity on the susceptibility of these bacteria to antimicrobials, as well as the growth media most routinely used and culture conditions, for these microorganisms, are also included.

Key words: Halophilic bacteria, Genetic transfer, Reporter genes, Gene expression analysis, Mutagenesis, Halomonadaceae, Chromohalobacter, Halomonas

1. Introduction

Among moderately halophilic bacteria defined as those prokaryotes that grow optimally in media containing 3–15% NaCl (1), members of the family *Halomonadaceae* show a remarkable versatility with respect to their salt tolerance, showing suitable features to study the

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_9, © Springer Science+Business Media, LLC 2012

molecular basis of bacterial osmoadaptation. In fact, some species such as *Halomonas elongata* and, especially, *Chromohalobacter salexigens* have been extensively used in recent years for the study of molecular osmoadaptation mechanisms in halophilic bacteria (2), which have been recently reported to be very complex (3). This heterogeneous family currently comprises ten genera of gammaproteobacteria (*Aidingimonas*, *Camimonas*, *Chromohalobacter*, *Cobetia*, *Halomonas*, *Halotalea*, *Kushneria*, *Modicisalibacter*, *Salinicola*, and *Zymobacter)* and 76 species, mostly halophilic bacteria. Recently, the phylogenetic status of this family has been extensively examined and reevaluated based on complete 16S rRNA and 23 rRNA gene sequence analysis (4). It was shown that these genera are phylogenetically coherent, except *Halomonas*, which is polyphyletic, showing two major clusters of strains. These findings confirm the high phenotypic heterogeneity previously reported for members of the genus *Halomonas* (5).

Besides their ecological interest, these extremophiles have interesting applications and potentialities in biotechnology. For instance, some of them are good sources for salt-tolerant enzymes (i.e., amylases, proteases, and nucleases) useful in a variety of biotechnological processes. Other organisms from this group produce salt-resistant biosurfactants or extracellular polysaccharides, which have a great potential as enhancers of oil recovery processes and for the pharmaceutical or food industry. Some of them are able to degrade toxic compounds from hypersaline wastewaters such as phenol, organophosphorous compounds, and others that are generated during manufacture, i.e., of pesticides and herbicides, constituting an interesting potential for the decontamination of such polluted waters (1, 6, 7).

Overall, one of their most interesting current and potential applications is the commercial use of the organic osmolytes that they accumulate in their cytoplasm to cope with the salt stress, the so-called compatible solutes. Although there is a relative diversity of such compounds in moderate halophiles, they are mainly poliols (glucosylglycerol), aminoacids (proline) and diaminoacids, quaternary amines (betaines) and ectoines (ectoine and hydroxyectoine), which have been demonstrated to be the predominant osmolytes accumulated by bacteria of the *Halomonadaceae* (8, 9). Since compatible solutes display additional biological properties in the cell, such as the stabilization and protection of enzymes, nucleic acids, and organelles against different stress agents, they have received a great interest as general biostabilizers with a wide range of biotechnological applications. Thus, some of them, especially ectoines, can be used as powerful stabilizers of enzymes, membranes, nucleic acids, antibodies, cell tissues, and whole cells against a variety of denaturing stresses caused by urea, salt, heating, freezing, radiation, and desiccation. Whereas betaines are widespread in nature, ectoines can only be produced biotechnologically from some natural

halophilic producers such as some *Halomonas* strains, although recently many studies involving heterologous expression of ectoine synthesis pathways both in prokaryotic and eukaryotic organisms have been carried out. For this reason, they are receiving increasing attention by the scientific community because of their multiple applications such as their use in the cosmetic industry, optimization of PCR, etc., among multiple uses, and exciting potentialities as therapeutic agents in the treatment of some degenerative diseases, by preventing whole cell damage and loss of viability. Recently, we have extensively reviewed the current and potential applications of ectoines as well as the different microorganisms and processes used for their industrial scale production (10).

Finally, an emerging field is the use of the *Halomonadaceae* for the production of recombinant proteins, as cell factories alternative to conventional prokaryotic hosts like *E. coli* or *Bacillus*. These moderately halophilic bacteria offer a number of advantages to be used as host cells. Thus, their high salt tolerance minimizes the risk of contamination, and they are easy to grow fast and maintain in the laboratory, with simple nutritional requirements since the majority can use a large range of compounds as their sole carbon and energy source (1). Among moderate halophiles, members of the *Halomonadaceae*, especially *Chromohalobacter* and *Halomonas*, possess the additional advantage of relatively easy genetic manipulation (1). Among others, genetic tools such as transposons, plasmids, cloning and expression vectors are available, as well as gene transfer methodology have been developed for this group of bacteria (11). Both native and heterologous promoters have proved to be functional in these halophiles to express a number of heterologous proteins of commercial interest, such as the ice-nucleation protein from *Pseudomonas syringae* (12, 13) or some amylases from hyperthermophiles such as *Pyrococcus woesei* (14). Although few recombinant proteins have been produced yet, the results show that the *Halomonadaceae* have a great potential as alternative cell factories. Bacterial ice nuclei (formed by the aggregation of ice nucleation, InaZ proteins) can be useful in various biotechnological applications such as artificial snow-making, frozen food industry, or cryoconcentration. However, the plant pathogenic character of the ice nucleation bacteria hampers the commercialization of most of these applications. In this respect, the expression of ice nucleation genes in nonpathogenic bacterial hosts (i.e., *Halomonadaceae*) is of great importance. Notably, recombinant *Halomonas* strains expressing the *inaZ* gene released ice nuclei in their growth medium, whereas in their natural hosts they usually remain membrane associated. In addition, these ice nuclei were more heat resistant than the ice nuclei from all other bacterial sources tested so far (13). Therefore, *Halomonadacea*e are a good source of ice nucleation protein.

Owing to the inherent difficulty of cultivation of hyperthermophiles, proteins of hyperthermophilic origin need to be expressed in mesophilic hosts for biotechnological purposes. For instance, an extracellular α-amylase gene from the hyperthermophilic archaea *P. woesei* was cloned under the control of a native promoter and expressed in *H. elongata*. The recombinant pyrococcal enzyme purified from *H. elongata* displayed catalytic properties similar to the native enzyme purified from *E. coli* with respect to thermal stability, pH and optimal temperature, and effect of various reagents on its activity (14). Thus, *H. elongata* could be used at least as well as *E. coli* for the production of functional α-amylase in a recombinant form.

In this review, the most relevant and recent tools for the genetic manipulation of moderate halophiles of the family *Halomonadaceae*, especially those belonging to the genera *Chromohalobacter* and *Halomonas*, are described. Emphasis has been given to the following: nucleic acid isolation procedures, cloning and expression vectors, genetic transfer mechanisms, mutagenesis approaches, reporter genes, promoter mapping and transcriptional regulation, and gene expression analysis. Growth media for the genetic manipulation and maintenance of these bacteria, as well as culture conditions are described. Owing to its relevance for genetic studies, an additional section devoted to the influence of salinity on the susceptibility of moderate halophiles to antimicrobials has been also included.

2. Genetic Manipulation

2.1. Media and Growth Conditions

Members of the *Halomonadaceae* grow best in media containing 3–15% NaCl (1). Apart from this, they have simple growth requirements and complex and defined media are available for all species. In our laboratory, the pH of all media are adjusted to 7.2 with a solution of KOH 1 M. Solid media contain 2% (w/v) Bacto-Agar (Difco). Unless otherwise stated, cultures are incubated at 37°C; liquid cultures are incubated in an orbital shaker at 200 rpm.

2.1.1. Complex Media

Strains of *Halomonas* and *Chromohalobacter* are routinely grown in complex SW-10 medium, which contains 10% (w/v) total sea salts and 0.5% (w/v) yeast extract (15, see Note 1). Modified versions of this medium containing a lower percentage of total salts can be also used for specific purposes, such as when cells are grown for conjugation. Many other rich media have been reported in the literature, as the MY medium supplemented with sea salt (16), the casamino acid (CAS) medium supplemented with NaCl described by Vreeland et al., (17), the Artificial Organic Lake Peptone (AOLP) medium supplemented with a salt solution and phosphate

reported by James et al. (18), and the K10 medium described by Severin et al., (19), Besides, for minimizing EPS production in *Halomonas* strains for plasmids isolation experiments, the MM medium was also used (20), and the RM and the T media were used for enrichment before transformation by electroporation of *Zymobacter palmae* (21, 22).

It is interesting to mention those media used for isolation of some *Halomonas* strains after different enrichment strategies from the natural environment by using as the salt solution the filter-sterilised water collected from the sources where they were isolated (23, 24). Otherwise, to determine siderophore production by *C. salexigens* strains at different salinities, a modified procedure to prepare Chrome Azurol S (CAS) agar plates was used. Thus, M63 minimal medium (see below) was used as a basal medium in which $FeSO_4 \cdot 7H_2O$ was omitted and a reduced KH_2PO_4 concentration (<0.03%, w/v) was used to eliminate any interference of the phosphate present in M63 medium with iron (3).

We have also tested the growth of some *Halomonas* and *Chromohalobacter* strains in the *E. coli* medium LB, which contains 1% NaCl, as well as in LB containing 2% NaCl. Whereas some of the strains grew well in LB + 2% NaCl, the original LB medium was growth-inhibitory for all but three strains of *Halomonas*, indicating again the salt requirement of the *Halomonadaceae* (Table 1).

2.1.2. Defined Media

When supplemented with NaCl, M63 (a defined medium commonly used for growth of *Enterobacteria*, 25) supports the growth of all *Halomonas* and *Chromohalobacter* strains tested in our laboratory and is usually used as the minimal medium of election. Glucose at a final concentration of 20 mM is used as the carbon source. We have also found that members of the *Halomonadaceae* grow well in the defined medium CDMM described by Kamekura et al. (26) for the moderate halophile *Salinivibrio costicola*. Other defined media have been also described, as the minimal medium used by Vreeland et al. (17), the MBM medium reported by Cummings and Gilmour (27), or the glucose-mineral medium G10 described by Severin et al. (19). It is worth noticing that the defined medium described by Vreeland et al. (17) and the CDMM medium contain sodium glutamate. Thus, the amount of NaCl has to be adjusted to allow for the sodium contributed by the sodium glutamate.

2.2. Effect of Salinity on Antimicrobial Susceptibility of Moderate Halophiles

Antimicrobials are extensively used as selective agents in genetic exchange experiments. The early work reported by our group (28) demonstrated that when moderate halophiles are grown at their optimal salt concentration, usually 10%, they tolerate very high concentrations of the majority of antimicrobials. This fact makes very difficult not only the use of antimicrobials as genetic markers but also the design of suitable selection media when these halophiles

Table 1
Growth of some strains[a] of *Halomonadaceae* in five complex media after 48 h of incubation at 37°C[b]

	SW-10[c]	SW-2	SW-1	LB[d]	LB-2[e]
Halomonas elongata ATCC 33173[Tb]	++	++	+	+	++
Halomonas meridiana DSM 5425[T]	++	++	++	–	++
Halomonas eurihalina ATCC 49336[T]	++	+	–	–	+
Halomonas halophila DSM 4770[T]	++	++	–	–	–
Halomonas halmophila ATCC 19717[T]	++	–	–	–	–
Halomonas halodurans ATCC 29686[T]	++	++	++	+	++
Halomonas subglaciescola DSM 4683[T]	++	++	++	++	++
Halomonas boliviensis DSM15516[T]	++	+	+	+	+
Chromohalobacter marismortui ATCC 17056[T]	++	+	–	–	–
Chromohalobacter salexigens DSM 3043[T]	++	++	++	–	+
Chromohalobacter salexigens ATCC 33174	++	++	++	–	+
Chromohalobacter canadensis ATCC 43984[T]	++	–	–	–	–
Chromohalobacter israelensis ATCC 43985[T]	++	++	+	–	–
Chromohalobacter sarecensis LV4[T]	++	+	+	–	–
Chromohalobacter nigrandesensis CECT5315[T]	++	–	–	–	–
Cobetia marina DSM4741[T]	++	++	+	–	–
Salinicola socius DSM19940[T]	++	+	+	–	–
Modicisalibacter tunisiensis CIP 109206[T]	++	++	++	++	++

[a] Cells were grown on plates with solid media
[b] T, type strain
[c] Media SW-10, SW-2 and SW-1 contain 10%, 2%, and 1% (w/v) total salts, respectively, and 0.5% (w/v) yeast extract
[d] The original recipe of medium LB contains 1% (w/v) NaCl, 1% (w/v) tryptone, and 0.5% (w/v) yeast extract
[e] LB-2: LB supplemented with 2% NaCl

are used as recipient for plasmids or transposons. Other studies on the susceptibility of *H. elongata* to some antibiotics indicated that a decrease of the salinity resulted in an enhanced sensitivity to these antimicrobials (29).

Based on these findings, we performed a deeper analysis on the influence of salinity on the susceptibility of moderately halophilic bacteria to antimicrobials (30). For comparative purposes, the MICs for *E. coli* strains DH5α and S17-1, widely used in genetic transfer studies, were also determined at different salt concentrations. Three different patterns of tolerance were found for moderate halophiles when salinity was varied from 10 to 1% (w/v) total salts

in the test media. The first one included the responses to ampicillin and rifampicin, where only minimal effects on the susceptibility were found. All moderate halophiles showed a high sensitivity to rifampicin regardless of the salt concentration. On the contrary, many of the tested strains showed a high resistance to ampicillin, indicating the presence of β-lactamases that inactivated the antibiotic regardless of the salt concentration. In the second group, including the response to aminoglycosides (gentamycin, kanamycin, neomycin, streptomycin, and geneticin), a remarkable and gradual increase of the toxicity was detected at lower salinities. Thus, by using media with low salinity, genes encoding aminoglycoside-resistance become suitable as genetic markers for plasmids or transposons to be transferred to moderate halophiles. The two *E. coli* strains tested also showed a gradually increased tolerance to aminoglycosides at high salt concentration (see Note 2), indicating that this effect is not exclusive for moderate halophiles. The third group included the rest of antimicrobials assayed (trimethoprim, nalidixic acid, spectinomycin, and tetracycline), where the effect of salinity was moderate and dependent on both the individual strain and the antimicrobial tested. All these data greatly facilitated further genetic studies on moderate halophiles since they simplified the design of selection media for genetic exchange experiments and enabled the use of the genes encoding resistance to some antimicrobials, especially aminoglycosides, as genetic markers for this bacterial group. Table 2 summarizes the final concentrations of the antimicrobials used in our laboratory to grow *Halomonas* and *Chromohalobacter* strains on solid SW media with different osmolarity.

2.3. Nucleic Acid Isolation

Plasmid and total DNA can be isolated from *Halomonas* and *Chromohalobacter* strains by using standard methods. Small- to medium-sized plasmids can be isolated by using the alkaline lysis method (31, see Note 3), or the Wizard SV plus kit (Promega). However, the "boiling" method (32), widely used for minipreps from *E. coli*, is not a suitable procedure to isolate plasmid DNA from *Halomonadaceae*. For large-scale preparation, alkaline lysis-isolated plasmid DNA can be further purified trough a Qiagen columm (Qiagen). The presence of megaplasmids can be tested by using a modification of the procedure of the Eckhardt method described by Plazinski et al. (33) or by a modification of the method previously described by Wheatcroft et al., (34) adapted to the members of the family of *Halomonadaceae* in which the original resuspension solution was changed to one containing 2% (w/v) salts to prevent lysis of the halophilic bacteria (20). The CTBA-NaCl method described by Ausubel et al. (35) or the procedure described by Marmur (36) are both good choices to isolate total DNA from *Halomonas* and *Chromohalobacter* strains.

Table 2
Final concentration of antimicrobials used for SW media of different salinities

Antimicrobial	Final concentration[a] (μg/mL) SW-10	SW-7.5	SW-5	SW-3	SW-2	SW-1
Ampicillin (Amp)	1,000	500	500	200	100	100
Rifampicin (Rif)	25	25	25	25	25	25
Gentamycin (Gm)	500	500	200	50	20	20
Kanamycin (Km)	1,000	500	200	100	75	50
Neomycin (Nm)	1,000	1,000	300	50	50	50
Streptomycin (Sm)	1,000	1,000	500	200	200	200
Trimethoprim (Tp)	500	500	200	150	150	150
Nalidixic acid (Nd)	1,000	500	200	150	150	150
Spectinomycin (Sp)	1,000	1,000	500	500	200	200
Tetracycline (Tc)	1,000	1,000	500	200	125	125
Geneticin (Gn)[c]	500	150	75	25	25	25

This table is a general guideline to select the appropriate antimicrobial concentration for members of the *Halomonadaceae*, which may vary depending on the particular strain. For more precise information on a particular strain of *Halomonas* or *Chromohalobacter*, as well as on two strains of *E. coli*, see ref. 30

[a] Concentration are referred to solid media

[b] Media SW-10 to SW-1 contain 10% (w/v) to 1% (w/v) total salts

[c] Concentration are referred only to *Chromohalobacter salexigens*

For RNA total isolation standard methods like that described by Salser et al. (37) or the acid phenol method (38) were used on *C. salexigens*, being preferred the last one if a high purity of the sample is needed (39). We strongly recommend the use in *C. salexigens* of the high pure RNA isolation kit (Roche) when no contamination of DNA is required (RT-PCR, qPCR, primer extension analysis, etc.).

Modifications of the hot phenol extraction method (40) have been described for *H. elongata* (see Note 4) being further purified using the RNeasy kit from Qiagen according to the manufacturer's instruction (41, 42). The use of RNA II Nucleospin kit (Macherey-Nagel) followed by rigorous treatment with Turbo DNA-free RNase (Ambion) has been reported in *Halomonas maura* (43, 44).

2.4. Cloning Vectors

Undoubtedly, cloning vectors are indispensable tools for the genetic manipulation of any bacterial group. A number of native plasmids have been isolated from Gram negative moderate halophiles (Table 3), and cloning vectors based on two of the best characterized replicons have been constructed (Table 4). In addition, some broad-host range vectors for Gram negative bacteria,

Table 3
Native plasmids isolated from bacteria of the *Halomonadaceae*

Plasmid	Natural host strain	Features	References
pMH1	*H. elongata* ATCC 33173 *H. halmophila* ATCC 19717 *H. halophila* CCM 3662 *Salinivibrio costicola* NCMB 701	11.5 kb, Km^r, Nm^r, Tc^r; unique *Eco*RI, *Eco*RV, and *Cla*I sites; able to replicate in *E. coli*.	(45)
pHE1	*C. salexigens* ATCC 33174	4.1 kb; unique *Bgl*II, *Eco*RI, *Pst*I, *Xba*I, and *Xho*I sites; unable to replicate in *E. coli*; Mob$^+$; complete sequence available (EMBL accession No. AJ132759 and AJ243735)	(47–49)
pHI1	*H. israelensis* ATCC 43985	ca. 48 kb	(47)
pHS1	*H. subglaciescola* UQM 2927	ca. 70 kb	(47)
pCM1	*C. marismortui* ATCC 17056	17.5 kb; unique *Bam*HI, *Bgl*II and *Xba*I sites; unable to replicate in *E. coli*; minimal replicon sequenced (GenBank accession No. X86092)	(50)
pVE1 and pVE2	*H. eurihalina* strain F2-7	pVE1: 8.2 kb; unique *Sma*I site pVE2: 5.8 kb; unique *Bgl*II site	(51)
pMA21, pMA22	*Halomonas strain* MA2	11 kb and 5 kb resistance plasmids, encode chromate, cobalt, zinc, and cadmium transporters (pMA21) and specific zinc and RND efflux transporters (pMA22). Confer both toxic metal and antibiotic resistance. Amp^R Sm^R, Tc^R, Km^R, Pn^R	(52)

especially those of the IncQ and IncP incompatibility groups, have shown their utility as cloning and expression vectors for members of the *Halomonadaceae* (Table 4).

2.4.1. Native Plasmids Isolated from Gram Negative, Moderately Halophilic Bacteria

Our group reported in 1992 the first plasmid isolated from moderately halophilic bacteria, pMH1, harbored by three strains of *Halomonas* and one strain of *S. costicola* (45) (Table 3). Despite *S. costicola* does not belong to the *Halomonadaceae*, we found that this species is present in the same ecological niches than *Halomonas* and *Chromohalobacter*. Therefore, it might be possible that pMH1 encode functions which are common for these microorganisms. The occurrence of pMH1 in the original halophilic strains was confirmed by Southern hybridization. However, the plasmid could not be detected on agarose gels by ethidium bromide staining of standard plasmid DNA preparations from the parental strains, but

Table 4
Cloning and expression vectors useful for gram-negative moderately halophilic bacteria

Vector	Features[a]	Cloning sites	Insertional inactivation	Host range: *E. coli*	*Halomonas*	*Chromohalobacter*	*Salinivibrio*	*Zymobacter*	References
Native-plasmid-based									
pHS15	12.25 kb, pHE1-derivative, Amp^r, Sm^r, Sp^r, ColE1 *ori*, RK2 *oriT*, Tra$^-$ Mob$^+$	*Bam*HI, *Eco*RI, *Kpn*I, *Not*I, *Pst*I, *Sal*I, *Sma*I , *Spe*I	–	+	+	+	+	ND	(47)
pHS134	9 kb, pHE1-derivative, Amp^r, Sm^r, Sp^r, ColE1 *ori*, Tra$^-$, Mob$^+$	*Bam*HI, *Bgl*II, *Sal*I, *Eco*RV	–	+	ND[b]	+	ND	ND	(49)
pEE3	6.6 kb, pCM1-derivative, Amp^r, Tp^r, ColE1 *ori*, Tra$^-$, Mob$^+$	*Not*I, *Sac*I, *Sma*I	–	+	+	+	–	ND	(50)
pEE5	7 kb, pCM1-derivative, Amp^r, Tp^r, ColE1 *ori*, *lacZ*, Tra$^-$, Mob$^+$	*Bam*HI, *Sac*I, *Sal*I, *Sma*I, *Xba*I	LacZ (*Bam*HI, *Sac*I, *Sal*I, *Sma*I *Xba*I)	+	+	+	–	ND	(50)
Broad-host-range									
pVK102	23 kb, Tc^r, Km^r, Tra$^-$, Mob$^+$, IncP, cosmid vector	*Sal*I, *Hin*dIII, *Xho*I	Tc (*Sal*I) Km (*Hin*dIII, *Xho*I)	+	+	+	+	ND	(54, 55)
pMP220	10.5 kb, Tra$^-$, Mob$^+$, IncP, promoterless *lacZ*, promoter-probe vector	*Bgl*II, *Eco*RI, *Pst*I, *Sph*I, *Xba*I	–	+	+	+	+	ND	(56)
pKT230	11.9 kb, Km^r, Sm^r, Tra$^-$, Mob$^+$, IncQ	*Bam*HI, *Eco*RI, *Hin*dIII, *Sac*I, *Sma*I, *Sst*I, *Sst*II, *Xho*I	Km (*Hin*dIII, *Sma*I, *Xho*I) Sm (*Eco*RI, *Sac*I, *Sst*I, *Sst*II)	+	+	+	+	ND	(54, 57)

pML123	12.5 kb, Gm^r, Nm^r, Tra^-, Mob^+, IncQ, expression vector	*Bam*HI, *Cla*I, *Eco*RI, *Hin*dIII, *Sac*I *Xba*I, *Xho*I	–	+	+	+	+	ND	(58)
pGV1124	10.8 kb , Sm^r, Cm^r, Tra^-, Mob^+, IncW	*Bam*HI, *Eco*R, *Hin*dIII, *Sal*I	Cm (*Eco*RI)	+	+[c]	+[c]	–	ND	(54, 59)
pMP92	pMP220 derived vector, $lacZ^-$, IncP, Tc^R	*Bgl*II, *Eco*RI, *Pst*I, *Sph*I, *Xba*I	–	+	+	+	+	ND	(56)
pMFY31	13.2 kb, RSF1010 replication origin. Tc^R, Cm^r, Amp^r	–		ND	ND	ND	ND	+	(22)

[a]See Table 2 for abbreviations of antimicrobials
[b]ND, not determined
[c]Only *H. elongata* and *C. salexigens*, and with very low transfer frequencies

could be observed when these preparations were introduced into *E. coli* by transformation. The above-mentioned difficulties prompted us to look for other native plasmids that could be more easily manipulated. In a second screening three autochthonous cryptic plasmids were isolated from *C. salexigens* (formerly *H. elongata*) ATCC 33174 (pHE1), *H. israelensis* (pHI1), and *H. subglaciescola* (pHS1) by using a slightly modified alkaline lysis procedure (Table 3). However, when a variation of the Eckhardt method (46) was utilized, no megaplasmid was detected in any of the strains tested (47).

Because of its small size (4,185 bp), pHE1 was selected for further characterization and construction of a shuttle vector for moderate halophiles. Analysis of its complete sequence revealed that the plasmid basically is composed of two functional parts: the replication (*rep)* and the mobilization (*mob*) regions (48, 49). The basic replicon consists of 1.77 kb *Bgl*II-*Eco*RI fragment, which contains the genetic information required for autonomous replication and stable maintenance. Analysis of its sequence revealed the presence of two genes, *repAB*, probably organized in one operon. *repA* encodes the replication initiator protein (RepA), which appeared to have a high degree of homology to the theta-replicase proteins of ColE2-related plasmids. The *repB* encoded product showed certain similarity with the RepB proteins of the same family of replicons. Deletion analysis suggests that pHE1 origin of replication (*ori*) is located in an 800-bp region upstream of *repA*. A third putative gene, *incA*, was found in the complementary strand to the leader region of *repA* mRNA. This, together with the presence in the 5′ end of the *repA* mRNA of inverted repeats that could form stable stem-loops structures, suggests that the *IncA* gene encodes a small antisense RNA. A possible control mechanism of pHE1 replication involving an RNA molecule, which sequesters the translational initiation region of the replication protein RepA, was proposed (48). The *mob* region, located in the 2.4-kb *Eco*RI-*Bgl*II remaining half of pHE1, contains four genes (*mobCABD*), which showed a complex organization with two of them (*mobB* and *mobD*) entirely overlapped by a third (*mobA*). The deduced proteins shared a high degree of homology to Mob proteins of ColE1 and closely related plasmids. Upstream of the *mob* genes, an *oriT* region with a putative nick sequence highly homologous to that of ColE1 plasmids was identified. pHE1 was found to be mobilizable from *E. coli* to *H. elongata* assisted by the helper plasmid pRK600 (49). To our knowledge, this is the first mobilizable plasmid found in moderate halophiles.

In a different screening, Mellado et al. (50) isolated pCM1, a narrow-host-range plasmid from *C. marismortui* (Table 3). Plasmid regions competent for self-replication were identified by cloning into suicide vectors. The minimal replicon of pCM1 was localized in a 1.6 kb fragment and was found to contain two

functionally discrete regions, the *oriV* region and the *repA* gene. *oriV*, located on a 0.7-kb fragment, contains four iterons 20 bp length adjacent to a DnaA box that is dispensable but required for efficient replication of pCM1, and it requires *trans*-acting functions. The *repA* gene encodes a replication protein of 289 residues which is similar to the replication proteins of other Gram negative bacteria such as that of plasmid pFA3 from *Neisseria gonorrhoeae*. However, despite pCM1 and pHE1 were isolated from two species of *Chromohalobacter*, their Rep proteins do not show significant homology to each other (48). This is not surprising, since although both plasmids have been proposed to have a theta-type replication mechanism, pHE1 replicon seems to be regulated via antisense RNA (48), whereas pCM1 replicon shares common features of iteron-regulated plasmids (50). In agreement with this, pHE1 and pMC1 have been shown to be compatible, allowing simultaneous introduction into the same host of vectors derived from their replicons for complementation studies or strain improvement.

H. eurihalina strain F2-7 produces large amounts of an anionic exopolysaccharide called polymer V2-7, with a number of promising applications in industry. The strain was found to contain two plasmids, pVE1 and pVE2 (Table 3), which were characterized with respect to their restriction maps and analysis of homologous sequences in other moderately halophilic strains. After treatment at high temperature, cured derivative strains lacking one or two plasmids were constructed. However, they produced the same mucoid phenotype than the parent strain, indicating that none of the plasmids is involved in polysaccharide production. Attempts to correlate pVE1 and pVE2 with other phenotypes, i.e., morphological, physiological, biochemical, and nutritional aspects, and antimicrobial and heavy-metal resistance were unsuccessful. Owing to their small size, the plasmids pVE1 and pVE2 are good candidates to develop cloning and expression vectors for Gram negative moderate halophiles, including the *Halomonadaceae* (51).

Osman et al. isolated two plasmids (pMA21 and pMA22) from *Halomonas strain* MA2, which showed resistance to both high copper concentration and other metals such as nickel, cobalt, zinc and a group of antibiotics (52). Plasmid curing resulted in a strain that was sensitive to ampicillin, streptomycin, tetracycline, kanamycin, penicillin, cobalt and chromate but not cooper, nickel and zinc, and showed a weak growth in the presence of 0.5–1 M NaCl. Partial sequencing of both plasmids led to identification of different toxic metals transporters. Thus, pMA21 encoded chromate, cobalt, zinc and cadmium transporters, whereas pMA22 encoded specific zinc and RND (resistance, nodulation, cell division) efflux transporters as well as different kinds of metabolic enzymes. These plasmids have a promising future as toxic metal- and antibiotic-resistance vectors as well as salt-tolerant agents.

2.4.2. Cloning and Expression Vectors

Two of the best characterized native plasmids have been used as a base to create cloning vehicles, pHE1 from *C. salexigens* (formerly *H. elongata*) ATCC 33174 and pCM1 from *C. marismortui*. The first shuttle vector for Gram negative moderate halophiles, pHS15 (Table 4), was constructed from the *C. salexigens* plasmid pHE1 (47). Since pHE1 appeared to be unable to replicate into *E. coli* cells, a number of mobilizable pHE1-derived hybrid plasmids were constructed that could be selected and maintained both in *E. coli* and *C. salexigens*. For this purpose, the *E. coli* pKS plasmid, the omega cassette (encoding resistance to streptomycin and spectinomycin) the *oriT* region (encoding the mobilization functions of plasmid RK2), and the entire pHE1 plasmid individually digested with *Pst*I, *Eco*RI, or *Bgl*II were assembled to give pHS7, pHS9, and pHS13, respectively. Plasmid pHS13 was improved by deletion of some repeated restriction sites to create pHS15. This vector contains a number of restriction sites commonly used for cloning and encodes resistance to streptomycin, spectinomycin, and ampicillin. It should be noted that, with some exceptions, the ampicillin resistance gene is not very useful for the *Halomonadacea*e since most of the strains have been shown to be resistant to this antibiotic (30).

It was fortunate that pHS13 was selected as the precursor to construct the shuttle vector pHS15. In pHS13, the native plasmid pHE1 was inserted as one *Bgl*II fragment. Analysis of the pHE1 sequence, subsequent to pHS15 construction, revealed that *Bgl*II recognition site is located in the intergenic region between the *repAB* and *mobCABD* genes of pHE1, and therefore, this site does not interrupt plasmid replication or mobilization functions. At the moment of pHS15 construction, it was unknown that pHE1 was mobilizable itself, and the mobilization functions were provided by cloning the *oriT* of plasmid RK2. Thus, pHS15 contains two *mob* regions, one from RK2 and one from pHE1. A simplified version of pHS15, vector pHS134 (Table 4), has been constructed by joining the *E. coli* plasmid pKS, the omega cassette and the plasmid pHE1 (47). This vector can be used directly for cloning or may constitute the basis for specialized pHE1-derived mobilizable vectors.

Two useful vectors, pEE3 and pEE5, have been constructed based on the minimal replicon of the *C. marismortui* plasmid pCM1 (Table 4). To generate pEE3, the trimethoprim resistance gene from plasmid pAS396 and the *oriV* of pCM1 were cloned in the plasmid pUC18mob, a derivative of pUC18 with the RK2 mob DNA (53). pEE3 carries unique *Not*I, *Sac*I, and *Sma*I sites, useful for cloning purposes. The characteristics of pEE3 as cloning vector were improved by removing superfluous restriction sites, and introducing the α-peptide of *lacZ* and a multiple cloning site. This new vector, pEE5 (53, Table 4), offers the possibility of white/blue colonies selection for cloning experiments. Both pEE3 and pEE5 carry a selection marker different than that of pHS15,

the trimethoprim resistance gene. Trimethoprim has been shown to be highly toxic for most species of the *Halomonadaceae*, regardless of the salt concentration of the culture media. This makes it a useful genetic marker for this bacterial group, especially in experiments in which high salinities need to be used.

Besides shuttle vectors, which are generally high copy number and restricted in host range, broad host range for Gram negative bacteria (usually low-copy number, apart from IncQ plasmids) also represent an important alternative for genetic studies. Our results indicate that IncQ and IncP plasmids can be particularly useful as cloning and expression vectors for the *Halomonadaceae*. Some of these plasmids have been included in Table 4.

2.5. Genetic Transfer Mechanisms

Conjugation, transformation, and transduction have been described as available genetic transfer mechanisms for some *Halomonadaceae*.

2.5.1. Conjugation

Conjugation has been successfully used to transfer both vectors derived from native plasmids (47, 52), and broad-host range vectors (29, 54) from *E. coli* to *Chromohalobacter*, and *Halomonas*. Several factors affecting the efficiency of conjugation (cell growth phase, mating time, donor–recipient ratio, and composition and salinity of the mating medium) were evaluated to optimize the conditions for conjugation between *E. coli* and moderate halophiles. In addition, comparative studies on plasmid host-range, stability and compatibility in *Halomonas* and *Chromohalobacter* strains were carried out. Table 4 summarizes the host range of the main cloning vectors reported. Whereas pHS15 (47), the IncP plasmids pVK102 (55) and pMP220 (56), and the IncQ plasmids pKT230 (57) and pML123 (58) were able to replicate in all strains tested, pGV1124 (IncW, 59) was maintained only in *H. elongata* ATCC 33173 and *C. salexigens* (formerly *H. elongata*) ATCC33174, and pCU1 (IncN, 55) could not be established in any of the host tested (54).

On the contrary, conjugation between moderate halophiles was demonstrated by using the self-transmissible IncP plasmid RK2 (54). Transfer of this plasmid between *Halomonas* species in SW-2 (2% total salts) medium was found at frequencies up to 1.2×10^{-3} to 2.8×10^{-4}, which were the highest conjugation frequencies achieved. When the salinity of the mating medium was raised up to 5% or 7%, conjugation was observed but the transfer frequencies were lower. No genetic transfer of RK2 was detected at 10% total salts, which are actually the salinity conditions that predominate in the natural habitats where moderate halophiles grow. The reason for this is unknown, but it might be related to a failure of the conjugation machinery encoded by RK2 to function (i.e., assemble or retrieve the sexual *pilus*) at high salt. At this respect, it would be interesting to test if any of the relatively large plasmids isolated in our laboratory, such as pHI1 or pHS1 (see Table 3) are

conjugative and, if so, whether or not they are able to mediate genetic exchange in hypersaline environments.

For optimal use of the cloning vector it is desirable that, once introduced, it is stably maintained in the host cell, even in a nonselective medium. The stability of the vectors pEE5 (53), and pHS15, pVK102, pKT230, and pGV1124 (54) has been assayed and compared up to approximately 80 generations. Out of the host strains tested, pEE5 was stable in *H. subglaciescola* and *H. halophila* for 80 generations, whereas it demonstrated to be highly unstable in *C. marismortui* and, to a lower extend, in *H. eurihalina*. The hybrid plasmid pHS15 (47) could not be maintained, in the absence of selective pressure, for long time in any of the moderate halophiles tested, including its parental strain. In any case, the highest stability of pHS15 occurred in the parental strain (60% loss after 80 generations) and the lowest in *H. halophila* (complete loss after 60 generations). The IncQ plasmid pKT230 was found to be stably inherited in all strains tested. Conversely, pVK102 (IncP) was highly unstable in most moderate halophiles assayed, with the exception of *C. marismortui*. Maintenance of pGV1124 in *H. elongata* was observed to be very unstable under nonselective conditions. Shuttle vector pHS15 was found to be compatible with pEE5 and with each of the broad-host-rage vectors assayed.

2.5.2. Transformation

Although natural transformation in this bacterial group has not been so far reported, a method for transforming *Zymobacte*r cells by using electroporation was described by Yanase et al. (21). Thus, the broad host range vector plasmids pRK290, pMFY31 and pMFY40 were transferred to *Z. palmae* by this method. Electroporation was performed using the procedure described in Note 5. High frequency of transformation was obtained at pulse voltages above 1.8 kV followed by 1 h of incubation. The frequency was 7.5×10^3 transformants/g of DNA for pRK290, 9.2×10^5 transformants/g of DNA for pMFY31, and 1×10^6 transformants/g of DNA for pMFY40, respectively, and was dependent on the molecular size of the plasmids, with the largest plasmid (pRK290) yielding the lowest frequency (21).

2.5.3. Transduction

Transduction assays have been described by Seaman and Day (60) using the *Myoviridae* ΦgspB and ΦgspC, the first virus reported to infect the genus *Halomonas* (see Note 6). Both viruses were investigated for their ability to transfer rifampicin resistance between hosts by transduction. ΦgspC was able to transfer rifampicin resistance between *H. salina* GSP21cells and also between them and *H. venusta* GSP4 cells, their alternative host. However, transduction frequency was below the level of detection with ΦgspB and either host.

The phage infection process showed a high level of tolerance toward temperature, pH, and salinity. Temperatures below 20°C

resulted in a fivefold reduction in infection for ΦgspB and ΦgspC. However, whereas ΦgspB infection was unaffected by pHs ranging from 6 to 9, ΦgspC showed optimum infection at pH 7 and 30°C, with infection rates dropping eightfold or more when incubated at pHs either side of this in combination with temperatures between 15 and 25°C. However, when exposed to a combination of pH 8–9 and temperatures greater than 30°C, ΦgspC exhibited the ability to lyse cells without replication, indicating a capability to infect cells, but not to replicate. Both phages showed high levels of infectivity across salinities ranging from 5 to 15%. However, once sodium chloride concentrations reached 20%, ΦgspB infectivity dropped eightfold, whereas, at this concentration, ΦgspC again exhibited lysis-from-without (60).

2.6. Mutagenesis

Mutagenesis and complementation analysis are powerful techniques in the characterization of operons, genes, and proteins. The creation of mutants is an essential step in determining gene function. In general, two main approaches are available to generate bacterial mutants: (1) in vivo random mutagenesis using chemical agents or transposons or (2) gene replacement of previously cloned DNA that has been manipulated in vitro by introducing insertions, deletions or base replacements. We describe below these kind of approaches carried out in *Halomonadaceae*.

2.6.1. Chemical Mutagenesis

To our knowledge, hydroxylamine (HA), methylmetane sulfonate, and nitrosoguanidine are so far the only chemical mutagens used in *Halomonadaceae*. Thus, mutants of *H. subglaciescola*, capable of growing without NaCl, were obtained after random mutagenesis using methyl methanesulfonate or nitrosoguanidine treatment as described in Note 7 (61). Otherwise, HA was used as a mutagen since it induces a high mutagenicity, causing a change in base pairing by favoring a tautomeric shift. The effect of different concentrations of HA on survival of *C. salexigens* DSM 3043 and *H. meridiana* as a function of the time exposure was investigated (62). A concentration of 0.2 (for *C. salexigens*) or 0.1 M (for *H. meridiana*) of mutagen was selected for further experiments because it offered a sufficient number of survivors for reliable detection of mutants. In addition, a clear effect of decreasing salinity on the enhancing of the killing action of HA in both species was found. From these experiments, SW-5 (5% total salts) was chosen as the treating medium for the two strains. Subsequently, the induced mutagenicity caused by treatment with HA was measured as the frequency of streptomycin resistant mutants isolated when cells of *C. salexigens* and *H. meridiana* were exposed to 0.2 M or 0.1 M, respectively, at different times of exposure. As a consequence of these results, the optimal conditions of HA mutagenesis for the two species were established (see Note 7). Finally, these conditions were used to isolate a number of auxotrophic mutants of *C. salexigens*

as well as different salt-sensitive mutants of *C. salexigens* and *H. meridiana*. Some of these latter mutants appeared to be affected in the synthesis of compatible solutes (62).

On the contrary, nonmucoid mutants of *H. eurihalina* were obtained after HA treatment. Different conditions of temperature and time of exposure were investigated, along with different concentrations of HA. A concentration of 0.02–0.04 M of mutagen for 30 min at 32°C was selected as optimal mutagenic conditions (see Note 8) (63).

2.6.2. Transposon Mutagenesis

2.6.2.1. Transposon Tn*1732*

Tn*1732* (64), a derivative of Tn*1721* (65, 66), which in turn is a member of the Tn*21* subgroup of the Tn*3* family, was used for the first time in 1995 by Kunte and Galinski (29) for transposon mutagenesis in *H. elongata*. Tn*1732* consists of the "basic transposon" of Tn*1721* (Tn*1722*) that encodes the transposition functions (transposase, *tnpA*, and resolvase, *tnpR*, genes), and a kanamycin resistance marker for positive selection of mutants. In view of the problems found with the use of Tn*5* in this species (see below), Tn*1732* was assayed as an alternative transposon since it was successfully used to overcome similar problems with Tn*5* mutagenesis of *Xanthomonas*. The suicide plasmid pSUP102-Gm (Cm^r) (see Note 9) was used to introduce Tn*1732* into *H. elongata* via *E. coli* SM10 mediated conjugation (29). Southern hybridization analysis of a number of transconjugants indicated the presence of single, randomly distributed, insertion sites within the chromosome. Furthermore, the possible cointegration of the vector pSUP102 in *Halomonas* DNA was excluded, as shown by hybridization experiments and the finding that Cm-resistant transconjugants were not observed. The phenotypic analysis of more than 3,000 Tn*1732* transconjugants proved the usefulness of the system for insertion mutagenesis in *Halomonas*. Thus, several auxotrophic mutants as well as different salt-sensitive mutants were isolated. These latter mutant strains occurred with a very high frequency (2%), confirming that Tn*1732* is a useful tool for transposon mutagenesis in this species. Since then, Tn*1732* has been successfully used for generalized insertion mutagenesis in different strains of *Halomonas* and *Chromohalobacter* (see Note 9). These included *H. elongata* (67) and *C. salexigens* mutants affected in the synthesis of the compatible solute ectoine (68), *H. elongata* mutants impaired in ectoine and hydroxyectoine uptake (69), *H. meridiana* mutants defective in extracellular amylolytic activity (70) and *H. eurihalina* Km resistant mutants (71).

2.6.2.2. Other Transposons

Tn*5*, which has been widely used for mutagenesis in Gram negative bacteria, was tested by Kunte and Galinski (29) for insertional mutagenesis in *H. elongata*. Transposon was successfully transferred from *E. coli* SM10 containing the suicide vector pSUP101::Tn*5* to *H. elongata* by biparental mating. However,

analysis of Tn*5* transconjugants showed that transposon and vector DNA cointegrated in a nonrandom manner into the chromosome. This was suggested to be due to the presence of hot spots for Tn*5* integration in *H. elongata* DNA, rather than the result of homologous recombination via vector DNA (29).

Mini-Tn5 has been used for insertional mutagenesis in *H. eurihalina*, obtaining both exopolysaccharide-defective mutants (71) and mutants that were dependent on arginine and uracil for growth (72), and in *H. maura*, obtaining mutants incapable of anaerobic respiration with nitrates (44) (see Note 10). However, we have attempted to use mini-Tn*5* for insertional mutagenesis in *C. salexigens* and found also evidence for the presence of hot spots into the chromosome. Attempts made by our group to use transposons belonging to the Tn*7* and Tn*10* families for insertional mutagenesis in *C. salexigens* have been likewise unsuccessful (personal communication). These findings cause severe limitations to the use of transposons in *Halomonadaceae*. One disadvantage of Tn*1732*, compared with transposons of the Tn*3* family, is its predilection for plasmid versus chromosomal DNA. That is, if one (or more) plasmid is present in the strain used in the mutagenesis, a high frequency of plasmid insertions will occur in detriment of chromosome insertions. On the contrary, new Tn*1721*-based transposons with distinct selection markers would be desirable, for instance, if multiple insertions in the same strains are required. Indeed, there are a number of Tn*1721*-derived transposons carrying antibiotic resistance genes other than kanamycin (64, 73). However, all these transposons contain the cognate transposase gene between their inverted repeated sequences. This fact implies that a recipient cell already containing one of these Tn*1732*-derived transposons becomes immune to further transposition rounds. New systems to deliver Tn*1721*-derived transposons in *Halomonadaceae* are therefore necessary, as the equivalent to the mini-Tn*5* transposons described by De Lorenzo et al. (74), which carry the transposase gene into the suicide vector (*in cis*), but external to the mobile element. In this way, a single recipient cell can be used for repeated insertion events with differentially marked minitransposons.

2.6.3. In Vitro Site-Directed Mutagenesis Followed by Gene Replacement

In vitro mutagenesis is an invaluable technique for studying protein structure–function relationships and for identifying intramolecular regions or amino acids, both of which may mediate gene expression and vector modification. Conventional DNA manipulation techniques (i.e., deletion of a specific DNA region by digestion and religation, insertion of a cassette carrying one antibiotic resistance marker or a reporter gene, among others) or PCR based site-directed mutagenesis approaches can be used to generate the mutation of interest. Following in vitro manipulation of the DNA, the mutated region must be transferred to the

host parental strain and gene replacement must occur by homologous (double) recombination.

Grammann et al. (69) reported the used of the splicing-by-overlap-extension (SOE) PCR technique (75) to generate a number of deletion mutants of *H. elongata* affected in the synthesis or transport of ectoine. Kraegeloh et al. (41) used also this technique to identify and characterize the K^+ uptake systems TrkH and TrkI from *H. elongata*. For this purpose, the DNA regions upstream and downstream of the gene to be deleted were joined together by applying the SOE-PCR methodology. The resulting PCR fragments containing the deleted gene were cloned into the suicide vector pKS18*mobsac* (76). This is a derivative of the *E. coli* plasmid pK18 carrying the *sacB* gene from *Bacillus subtilis*, which is inducible by sucrose and is lethal when expressed in Gram negative bacteria. The incorporation of this conditionally lethal gene greatly simplifies the identification of double recombination events (those leading to gene replacement) among the vast majority of single cross-overs (those which still contain the vector). This avoids time-consuming replica-plating or toothpicking. Constructs carrying the deletions were transferred into *H. elongata* by *E. coli* S17-1-mediated conjugation. Deletion mutants arising after double cross-over were selected on complex medium containing sucrose. The deletion sites were verified by PCR and DNA sequencing.

A similar methodology to generate deletion mutants, the Quick Change site-directed mutagenesis kit developed by Stratagene, a widely used PCR-based system that eliminates the necessity to subclone the amplified fragment, was used by Tokunaga et al. (77) to determine the role of the amino acid Glu134 in the structure of the nucleoside diphospate kinase from *Halomonas* sp. 593. Likewise, this methodology has been employed in our laboratory to obtain mutants affected in the region encoding ectoine synthesis (*ectABC genes*) in *C. salexigens.* Thus, following in vitro mutagenesis, a fragment containing a promoterless *lacZ* gene and a streptomycin resistant marker (the Ω cassette, 78) was inserted at the deletion point. The reporter *lacZ* was included to monitor the activity of the of the *ectABC* promoter. The Ω cassette served to facilitate the selection of mutants. The assembled Δ*ectABClacZ*Sm fragment was subcloned into the vector pJQ200SK (79), a suicide mobilizable vector similar to pKS18*mobsac* that also carries the conditional lethal *sacB* gene. Gene replacement was achieved after transferring the deleted fragments to a spontaneous rifampicin resistant mutant of *C. salexigens* by triparental matings assisted by the helper plasmid pRK600 (80). Selection was on SW-2 medium with streptomycin (100 μG/mL) and sucrose (10% w/v). Deletion was confirmed by Southern hybridization using an internal *ectB* probe, as well as by the absence of the compatible solute ectoine in the mutants.

Table 5
Intrinsic β-galactosidase and alkaline phosphatase activity[a] of some strains of *Halomonas* and *Chromohalobacter*

Strain	β-Galactosidase[b]	Alkaline phosphatase
Halomonas elongata ATCC 33173[Tc]	++[d]	++
Halomonas meridiana DSM 5425[T]	–	+
Halomonas eurihalina ATCC 49336[T]	–	+
Halomonas halophila DSM 4770[T]	–	++
Halomonas halmophila ATCC 19717[T]	+	–
Halomonas halodurans ATCC 29686[T]	+	+
Halomonas subglaciescola DSM 4683[T]	++	++
Chromohalobacter marismortui ATCC 17056[T]	+	+
Chromohalobacter salexigens DSM 3043[T]	–	–
Chromohalobacter salexigens ATCC 33174	++	++
Chromohalobacter canadensis ATCC 43984[T]	(+)	+
Chromohalobacter israelensis ATCC 43985[T]	(+)	++

[a] Activity was observed after 48 h of incubation at 37°C on plates of minimal medium M63 with 10% NaCl and 40 μg/mL of the substrate

[b] The chromogenic substrates used were 5-bromo-4-chloro-3-indolyl-β-d-galactopyranoside to test β-galactosidase activity and 5-bromo-4-chloro-3-indolyl-phosphate to test alkaline phosphatase activity

[c] T, type strain

[d] ++, strong positive; +, positive; (+) weak positive; –, negative

2.7. Reporter Genes

When reporter genes are used, the presence of cellular intrinsic activity leads to interferences in the assay interpretation. In this respect, many species of the family *Halomonadaceae* have intrinsic β-galactosidase or alkaline phosphatase activities, two of the most widely used reporter systems (see Table 5). For Lac^- strains, as *C. salexigens* or *H. meridiana*, a chromogenic assay with 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal, see Note 11) can be used on plates to qualitatively monitor gene expression. We have made transcriptional fusions of osmoregulated promoters from *C. salexigens* and the *E. coli lacZ* gene, encoding β-galactosidase. Our results indicate that salinities as high as 10% (1.72 M NaCl) do not interfere with the in vivo assay on plates, probably because the substrate is taken up by the cells and the *lacZ*-encoded enzymatic reaction occurs within cell cytoplasm. However, in our hands, the quantitative in vitro assay for β-galactosidase activity, which requires cell lysis previous to substrate (o-nitrophenyl-beta-D-galactopyranoside) addition, is unviable at salinities higher than 4.3% (0.75 M) NaCl. Several variations of the standard method,

with additional washing steps before cell lysis, or spheroplasts preparation (to remove NaCl that might have bound to the outer membrane) have been tested in our laboratory, but yielded no reproducible results when cells are grown at salinities above 0.75 M NaCl. This, which might be related to a rapid inactivation of the enzyme by salt, is an important constraint to the use of the *E. coli lac*Z gene as a reporter for members of the *Halomonadaceae*, where salinities up to 2.5 M (14.5%) NaCl are used to assay gene expression.

The above mentioned difficulties prompted us to look for alternative reporter systems for moderate halophiles. The ice nucleation (*inaZ*) and green fluorescent protein (*gfp*) genes are totally absent from halophilic microorganisms. Therefore, there were tested as potential reporter systems for members of the *Halomonadaceae*. The signal generated from the product of these two genes is not due to enzymatic catalysis but to a physical phenomenon. Both assays are in vivo tests. It was shown that both reporters were efficiently expressed in moderate halophiles of the genera *Halomonas* and *Chromohalobacter* under the control of either heterologous or native promoters (12, 81).

The product of the *inaZ* gene is the ice nucleation protein (about 185 kDa). Expression of the gene can be easily detected in cultures of cells growing at temperatures below 24°C by a freezing-droplet test using a supercool bath at temperatures ranging from −3 to −10°C (82). The *inaZ* gene from the plant pathogen *P. syringae* was the first reporter system described in Gram negative moderately halophilic bacteria (12). A promoterless version of *inaZ* was introduced into two different restriction sites (*Eco*RI and *Pst*I) of the pHE1 part of vector pHS15. One orientation of both recombinant constructs expressed high levels of ice nucleation activity in *Halomonas*, indicating that *inaZ* was probably introduced in the correct orientation downstream of putative native promoters. A recombinant construct carrying a tandem duplication of *inaZ* in the same orientation yielded significantly higher ice nucleation activity, showing that *inaZ* is appropriate for gene dosage studies. i*naZ* was also expressed in *Halomonas* under the control of heterologous promoters, such as the β-lactamase (Pbla) promoter of *E. coli* and the pyruvate decarboxylase promoter (Ppdc) of *Zymomonas mobilis* (12). Remarkably, the salt concentration of cultures growing on SW-10 medium (10% w/v total salts) did not cause any apparent inhibition of the ice nucleation activity. For this reason, SW10 was the choice medium to make the *InaZ* assays.

The existence of a promoter located upstream of the *Pst*I site of native plasmid pHE1 was further confirmed by the construction of a transcriptional fusion of the 1.3 kb *Eco*RI-*Pst*I fragment of pHE1 (containing the putative promoter) and *inaZ* (83). Subsequent analysis of the complete sequence of pHE1 (48, 49) revealed that this region contained the promoter of the pHE1 *mob*

operon (*mobCABD*) encoding mobilization functions (45). The exact localization of Pmob, which is functional in *E. coli* (83), needs experimental confirmation. However, the first 90 pb of the 1.3 kb region must be necessary for promoter function, as judged by the fact that deletion of this DNA sequence leads to a complete loss of *inaZ* activity (83).

The green fluorescent protein (GFP), a 27 kDa protein from the marine bioluminescent jellyfish *Aequorea victoria* (84), is a unique marker that can be detected by noninvasive methods. No substrates, other enzymes or cofactors are required for GFP assay. To monitor gene expression in *H. elongata*, Douka et al. (81) used a mutated version of this gene (85), which allows the establishment of GFP as a convenient expression marker for bacteria, because it can fold correctly and remain soluble in the cells, resulting in 100-times increased fluorescence as compared to the wild type GFP when expressed in *E. coli*. To investigate *gfp* expression in *H. elongata*, the shuttle vector pHS15 was used in which the *gfp* gene was placed under a native (Pmob) and an heterologous (Ppdc) promoter, giving the constructs pHS15G1 and pHS15G2, respectively. Both recombinant constructs were transferred to *H. elongata* by conjugation and GFP activity was measured at various concentrations of NaCl (2, 5, and 10%). Although the *gfp* gene was expressed in all transconjugants, as shown by RT-PCR analysis, expression from Pmob raised as the salt concentration in the growth medium increased from 2 to 10%, suggesting that the mobilization functions of the native plasmid pHE1 might be osmoregulated (81).

Currently our group is developing *inaZ* and GFP-based monitoring gene expression systems using IncP low-copy and autochthonous plasmids (Argandoña et al., unpublished data). The GFP-based system was also used to check whether there was a promoter upstream of the *cfuABC-fur-hisI-orf6* operon in *C. salexigens*. For that purpose, a transcriptional fusion of the region upstream of *cfuA* with the *gfp* reporter gene was constructed in the low-copy number plasmid pMP92 and transferred by triparental mating to the wild type strain. As a negative control, the promotorless vector was also transferred and promoter activity was measured in cells grown in M63 minimal medium containing 0.75 M NaCl supplemented with tetracycline (3).

2.8. Promoter Mapping and Transcriptional Regulation

The two methods conventionally used to determine transcriptional starting sites and transcriptional regulation are S1 nuclease mapping and radioactive primer extension experiments. Nuclease protection assays (S1 nuclease protection and RNase protection) are extremely sensitive procedures for detection and quantization of mRNA species in complex mixtures of total cellular RNA and are well suited for mapping positions of external and internal junctions in RNA, such as transcription initiation and termination sites. Besides, since the size of the probes used in nuclease protection

assays can be chosen by the investigator, probes may be designed to protect fragments of different sizes. This feature permits the simultaneous analysis of several different mRNAs in the same total RNA sample (86).

The S1 nuclease protection assay was developed in *C. salexigens* to map the promoter regions upstream of the genes *ectA* and *ectB* (39, see Note 12). Thus, four promoters upstream of *ectA* were suggested. This assay not only showed the number of transcription initiation sites of each promoter region ($pectA_{1-4}$ and $pectB_5$) but also allowed to analyze the transcriptional regulation of $pectA_{1-4}$ and $pectB_5$ at low (0.75 M NaCl) and high salinity (2.5 M NaCl). In fact, the data strongly suggested the involvement of the general stress σ^S in *ectABC* transcription in *C. salexigens*. Although the analysis of the same promoters using *lacZ* fusions were also carried out at increasing NaCl concentrations, we were not able to make reliable measurements of β-galactosidase activity in cells grown with >0.75 M NaCl, most probably due to the salt-inactivation of the enzyme (39).

Primer extension reaction is another alternative to determine the transcription initiation site and level of a target 5′end of mRNA. The assay uses a reverse transcriptase enzyme (AMV) and an end-labeled primer, usually ^{32}P end-label, to reverse-transcribe the target RNA into cDNA. The amount of cDNA synthesized is proportional to the amount of target RNA, and the length of the cDNA product reflects the distance from the primer to the 5′end of the RNA. Primer extension products are typically analyzed on denaturing polyacrylamide gels and the use of labeled DNA markers allow the size of the primer extension product to be determined. This protocol were successfully used in *H. maura* to localize a single transcription initiation site of the *epsABCDJ* operon, involved in the synthesis of the exopolysaccharide (EPS) mauran (43). Besides, it was used to map the promoter region upstream of *narGHJI* genes involved in anaerobic respiration with nitrate, showing a unique transcription initiation site and the existence of a regulation mechanism by anaerobiosis and nitrate (44, see Note 13).

Traditionally, the primer extension protocols are ^{32}P- or ^{35}S-based methods, making them potentially hazardous, time consuming, and specifically laboratory requirements for radioisotope research. In recent years, modifications of the primer extension methods have been developed using various fluorescently labeled primers and commonly available DNA sequencing instruments thus avoiding the necessity for handling radioactivity and polyacrylamide. These fluorescence-based primer extension reactions have been shown to be as sensitive as radioactive protocols. Fekete et al. (87) totally developed this innovative protocol, termed as fluorescently labeled oligonucleotide extension (FLOE), using 5 V-FAM-labeled primers, which (1) have a high level of precision and allows quantification of results, and (2) FAM-labeled cDNA

products are compatible with ABI fluorescent sequencing technology and can be electrophoresed on an ABI PRISMR DNA sequencer (88). The promoter corresponding to the major transcriptional start site will produce higher levels of mRNA transcript, resulting in a stronger FAM signal, correlating the intensity of the FAM signal with the amount of cDNA present in the sample. If multiple promoters are present in a sample, the relative strength of each promoter can be inferred from the peak heights of the primer extension products (88). The FLOE approach has been only used in a few recent studies in promoter mapping of bacteria (88–90). In our laboratory, we have successfully developed this approach to map the promoter region upstream of the gene *ectD* of *C. salexigens* involved in the synthesis of hydroxyectoine and also to analyze the transcriptional regulation under different stress conditions of salinity and temperature using 10–15 μg of total RNA and 6-FAM end-labeled primers (Reina-Bueno et. al, unpublished results).

Another alternative method for mapping promoters in which neither radioactive nor fluorescent labeled products is used has been recently developed, the 5′RACE-PCR. A modified protocol has been developed in *H. elongata* to study the involvement of the universal stress protein TeaD in regulating the osmoregulated ectoine transporter TeaABC (91). Thus, the transcriptional initiation sites were mapped in the upstream region of *teaABC* genes. In that protocol, cDNA is synthesized from total RNA using reverse transcriptase (Superscript III, Invitrogen) and a reverse primer. Applying terminal deoxynucleotide transferase, a poly (C) tail was attached to the 3′end and the modified cDNA was PCR-amplified with forward primer AAP (Abriged Anchor Primer) and reverse primer. Finally, PCR-amplified cDNA was sequenced. This approach revealed the presence of a σ^{70}-dependent promoter (P1) and another σ^{S}-dependent promoter (P2) (91).

2.9. Gene Expression Analysis

RT-PCR and Northern blot are techniques that are usually developed to detect the presence of a specific mRNA in a sample. RT-PCR is a semiquantitative technique that is generally used to determine the expression of a gene and also to analyze the gene organization in operons. This technique was recently used to confirm the organization of *cfuABCfurhisI* genes (involved in iron homeostasis) in a single operon in *C. salexigens* (3) or the organization of *epsABCDJ* genes in *H. maura* (43). This alternative protocol to Northern blot experiments could be used when the size of the mRNA is too large and more suitable to mechanical broken. Both Northern blot and RT-PCR were carried out to analyze the transcriptional organization of the *teaABCD l*ocus in *H. elongata*, confirming that *teaD* gene is cotranscribed with *teaABC* (91, see Note 14).

In *H. maura*, a semiquantitative RT-PCR based protocol was also carried out to undertook transcriptional analyses for the *narGHIJ* genes (responsible for anaerobic respiration with nitrate) under different nitrate respiration conditions where the cells were induced by nitrate (0.1, 0.5, and 1% w/v for 1 h) under O_2 limitation and also to confirm whether *narGHJI* genes were transcribed together (44). Reverse transcription-PCR analyses were also done to test the heterologous expression of *crtY*, *crtI*, and *crtB* genes of *Pseudomonas agglomerans* in *H. elongata* (92). Kim et al. carried out RT-PCR experiments in *Chromohalobacter* sp.HS-2 to investigate genes involved in benzoate and p-hydroxybenzoate catabolism and if degradative genes (*benABC*, *catBCA*, *pobA*, and *pcaGH*) are induced specifically by benzoate and *p*-hydroxybenzoate (93). Similar experiments have been recently performed in *C. salexigens* to analyze the transcriptional induction of genes involved in the isothionate formation from taurine (*tauAB1B2C*, *Csal_0158*, and *isfE*) (94).

As we mention above there are a variety of methods for the quantification of mRNA, like Northern blotting and ribonuclease protection assays (RPAs), but require more RNA than is sometimes available. PCR methods are therefore particularly helpful when amounts of RNA are low, since the fact that PCR involves an amplification step means that it is more sensitive. Thus, real-time PCR (qPCR) was developed because of the need to quantify differences in mRNA expression, also in samples with only small amounts of mRNA. In contrast to regular reverse transcriptase-PCR and analysis by agarose gels, qPCR gives quantitative results and additionally, it is relative simply and not time-consuming compared to some older methods. Despite these advantages, only a few gene expression studies have been carried out in moderate halophilic bacteria using qPCR. In *C. salexigens* this approach have been very helpful to quantify and compare the different expression levels of the *ectA* gene, involved in ectoines biosynthesis, and the *cfuA* gene (involved in iron homeostasis), under different conditions such as low and high salinity and low and excess of iron conditions both in the wild type and an iron homeostasis regulator mutant (*fur*) (3). Similar analysis have been developed in our laboratory to compare expression of *ectD* and *ectE*, which codify for ectoine hydroxylase paralogs, at different salinity and temperature stress conditions both in the wild type strain and in a specific regulator mutant backgrounds (Reina-Bueno et al., unpublished results). The possibility that qPCR offers to compare and quantify the expression of more than one specific mRNA at the same time in several samples and conditions from wild type strain and mutants is a powerful tool to unify expression data. Additionally, it allows us to measure and compare expression data obtained at low and high salt conditions, fact that is always a handicap in halophilic bacteria.

3. Notes

1. SW-10 is a saline complex medium containing 10% (w/v) total sea salts solution to which 0.5% (w/v) yeast extract is added (15). The composition of the sea salt solution 30% (from which the corresponding dilutions can be performed to reach the different SW media at different salinities) was (in g/L): NaCl, 234; $MgCl_2 \cdot 6H_2O$, 41.6; $MgSO_4 \cdot 7\ H_2O$, 59.82; $CaCl_2 \cdot 2H_2O$, 1.1; KCl, 6; $NaHCO_3$, 0.2; NaBr, 0.7; and $FeCl_3 \cdot 6H_2O$ solution at 0.5% (w/v), 0.65 mL (16). The pH of the medium is adjusted to 7.2 with KOH. The composition of MM medium is as follows: 0.4 g/L yeast extract, 2 g/L peptone no. 3, supplemented with a salt solution (SW) (16) (except for $MgSO_4$) to reach a final concentration of 2% (w/v) and 14.26 g sodium thioglycolate sterilized by filtration (20). The composition of modified MY medium was (in g/L): glucose, 10; proteose-peptone No.3 (Difco), 5; malt extract (Difco) 3; yeast extract (Difco) 3, in a 7.5% (w/v) salt solution (16). RM medium contains (in g/L): glucose, 20; Bacto yeast extract (Difco), 10, and potassium phosphate, 2, at pH 6.0. T medium contains (in g/L): glucose, 20; Bacto yeast extract (Difco) 10; potassium phosphate, 10; $(NH_4)_2SO4$, 2; and $MgSO_4 \cdot 7H_2O$, 0.5, at pH 6.0 (21).
2. Although *E. coli* is a nonhalophilic microorganism, it has the ability to adapt rapidly to moderate salinities provided that compatible solutes are supplied with the grown medium. Therefore it should be considered as a slightly halotolerant bacterium. We were able to grow *E. coli* in the same media than moderate halophiles do (SW complex medium) up to 5% total salts (30).
3. Prior to the alkaline lysis method, the cells can be washed with 0.1% SDS in TE buffer (10 mM Tris–HCl, 1 mM EDTA, pH 8.0). This washing step greatly improves the isolation method, yielding cleaner plasmid DNA preparations, suitable for digestion with restriction enzymes.
4. Some modifications of the hot phenol procedure were performed as follows: 100 mL of exponentially growing cells (OD_{600} of 0.6–0.7) were harvested by centrifugation, and approximately 100 mg of the pellet was resuspended in 4 mL of buffer A (50 mM Na acetate, 10 mM EDTA) containing 0.5 mL of 10% SDS, to lyse the cells. After addition of 5 mL of hot phenol (65°C), the sample was incubated at the same temperature (65°C) for 4 min and then frozen in liquid nitrogen for 2 min. The frozen sample was then thawed (37°C) in a water bath and centrifuged for 10 min (2,700 × *g*) to enhance phase separation. The aqueous top layer (400 μL) was removed

and mixed with an equal volume of phenol–chloroform–isoamyl alcohol and centrifuged at 8,000 × *g* for 5 min at 4°C. The RNA in the aqueous top layer was subsequently precipitated with 40 μL of 3 M Na acetate solution and 400 μL of pure ethanol at −70°C. Prior to use for Northern hybridization or reverse transcriptase PCR (RT-PCR), RNA was further purified by using the RNeasy mini-kit (Qiagen) according to the manufacturer's instructions.

5. To transform *Z. palmae* the strain was initially grown statically in 5 mL of RM/T medium at 30°C for 15 h. The preculture was added to 50 mL of fresh RM/T medium and incubated at 30°C for 90 min. The culture broth was then chilled on ice and centrifuged in a microfuge at 850 × *g* for 10 min at 4°C. The harvested cells were washed with 25 mL of ice-cold 10% glycerol and resuspended in 1 mL of ice-cold glycerol. An aliquot (0.2 mL) of cell suspension was transferred to a 0.2 cm electroporation cuvette and 10 μL of plasmid DNA solution was added. Immediately after pulsing of the cell suspension using a Bio-Rad gene pulser apparatus (1.8 kV, 200 Ω, 25 μF) 1 mL of prewarmed (30°C) RM/T medium was added and incubated statically at 30°C for 1 h. Finally, the cells were spread onto selection plates and incubated at 30°C for 1–3 days (21, 22).

6. To transfer rifampicin resistance through transduction, *H. salina* GSP21 was used as the donor and transducing particles were produced by infecting this strain with either ΦgspB or ΦgspC by the overlay method. Phages were purified from the plates after overnight incubation and to ensure that the transducing particles contained only donor strain DNA, a second round of high titre phage lysate was produced with the same donor strain. Previously, phage lysates were digested with 50 U of DNase I per mL to remove any free DNA that would be available for transformation. The bacterial hosts *H. salina* GSP21 and *H. venusta* GSP4 were used as recipients for transduction. 0.5 mL of phage lysate was then mixed with 0.5 mL of a mid-log phase host culture. The final washed cell pellet was resuspended in 0.5–1.5 mL of SP-1 broth, and plated onto selective SP-1 plates containing 50 mg/L rifampicin.

7. Optimal experimental condition of methylmethane sulfonate and nitrosoguanidine mutagenesis for *H. subglaciescola* were as follows: cellular suspensions containing about 10^8 log-phase cells/mL were harvested by centrifugation (5,000 × *g* and 4°C for 30 min) and washed twice with saline. 0.1 volume of 1.0 mM mutagen (Sigma) was added to the bacterial suspension and incubated at 4°C for 3 h. The treated cells were then spread on an agar medium plate without NaCl and incubated at 30°C for their growth.

8. Optimal experimental condition of HA mutagenesis: (a) for *C. salexigens*: Cellular suspensions containing about 10^8 log-phase cells/mL were centrifuged at 30,000×*g* for 1 h and washed three times with sterile SW-5 solution. Then, they were exposed to 0.2 M HA (Sigma) in SW-5 medium, pH 5.2, for 90 min at 37°C; (b) for *H. meridiana*: same as above but exposed to 0.1 M HA in SW-5 medium, pH 5.2, for 2 h at 37°C, (c) for *H. eurihalina*: The cells were previously grown until the early exponential phase only (10^6 cells/mL) to prevent the overproduction of extracellular polymers (EPS), which can hinder the effect of the mutagen. The cells were centrifuged at 30,000×*g* for 1 h and washed three times with sterile SW-7.5 solution to ensure that no EPS remained clinging to the cells. Then, about 10^9 cells were suspended in 1 mL of MY medium and transferred to tubes containing a concentration of 0.02–0.04 M HA, incubated at 32°C for 30 min, and later on incubated in MY medium without the mutagenic agent.
9. The use of *E. coli* vectors carrying a conjugal transfer (*mob*) sequence is a standard way to introduce transposons into a variety of bacteria, where such vectors cannot replicate. For this reason they are termed "suicide" vectors. We have tried to use other suicide Tn*1732* donors based on unstable (thermosensitive) replicons (64) but failed to obtain Tn*1732*-induced mutants. Therefore, the choice of the appropriate delivery vector for Tn*1732* into the cell is crucial for the success of mutagenesis. In our laboratory, Tn*1732* induced mutagenesis is performed by conjugal transfer of pSUP102-Gm::Tn*1732* from *E. coli* SM10 to a spontaneous Rif^r mutant of the wild type strain to be mutagenized as described by Cánovas et al. (68). Matings are carried out by mixing the donor and recipient cultures at a ratio of 1:4 (100 μL of donor, 400 μL of recipient). The mixed cultures are washed with sterile SW-2 medium to eliminate the antibiotics. The pellet is then suspended in 100 μL of SW-2 and placed on a 0.45-μM pore filter onto SW-2 solid medium (which allows the growth of *E. coli*). After overnight incubation at 30°C, cells are resuspended in 20% (v/v) sterile glycerol and, after appropriate dilutions, inoculated onto SW-2 plus 25 mg/L rifampicin and 75 mg/L kanamycin plates at a density resulting in about 100–200 colonies per plate. Colonies from these master plates are subsequently transferred with sterile toothpicks to the appropriate media for selection of mutants.
10. Mini-Tn5 induced mutagenesis was performed by conjugal transfer of pUT mini-Tn5 Km2/Cm from *E. coli* to the correspondent receptor strain as described by Llamas et al., (71).

The pellets were suspended in 100 μL of the same medium, mixed at a ratio of 1:1, 1:2, or 2:1, and placed onto a nitrocellulose membrane (pore size 0.22, 0.45, or 0.7 μm, Millipore) laid over a plate of the selected conjugation medium. The plates were incubated at 30, 32, and 37°C for 16, 20, and 24 h, and the appropriate dilutions were made and inoculated onto SW-2 added with the corresponding antibiotics. After 3 days of incubation at 30°C, the transconjugants were picked onto master plates of the same composition. The mating mixture was also spread onto SW-2 agar plus 50 mg/L rifampicin, to determine the rate of recipient-cell survival after the transposon experiment.

11. X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) is added to the plates at a concentration of 40 μg/mL.

12. In a first step of the S1 nuclease protection assay in *C. salexigens*, a complementary ssDNA probes hybridizing to the 5′ region of the mRNA is generated by lineal PCR (95) using linearized plasmids which contain the *ectA* and *ectB* genes as templates. The primers used are ^{32}P- end-labeled oligonucleotides that are complementary to the mRNA. 25 μg of total RNA and an excess of a ^{32}P-end-labeled single-stranded DNA (ssDNA), previously denaturalized, is used for hybridization reaction. Finally, the mix is digested with nuclease S1 and the obtained fragments are analyzed in an acrylamide gel.

13. In *H. maura*, primer extension experiments were performed using a specific oligonucleotide complementary to the sense strand and radioactively labeled at its 5′end with [γ-^{32}P] ATP and T4 polynucleotide kinase. The end-labeled primer was hybridized to 20–50 μg of total RNA and the extension reaction was carried out using AMV reverse transcriptase (Roche). The cDNA product and sequencing reactions were analyzed on a 6.5% (w/v) urea-polyacrylamide sequencing gel.

14. In *H. elongata* the optimal protocol for Northern experiments was as follows: A *teaA*-specific-RNA antisense probe was constructed from genomic DNA and used to amplify by PCR an intragenic *teaA* DNA fragment using a reverse primer, which carried the promoter sequence for T7 RNA polymerase. This DNA template was applied to in vitro transcription (Roche Molecular Biochemicals) to generate DIG-11UTP-labeled antisense *teaA* RNA. 5 μg of total RNA was electrophoretically separated on a denaturing gel (0.8%), transferred to a nylon membrane and covalently bound by UV irradiation. Bound RNA was hybridized with DIG-labeled antisense *teaA* RNA probe, and finally probe hybridization was detected by chemiluminescence via X-ray film.

Acknowledgments

Research in the authors' laboratory was financially supported by grants from the Spanish *Ministerio de Ciencia e Innovación* (grant BIO2008-04117) and *Junta de Andalucía* (grant P08-CVI-03724).

References

1. Ventosa, A., Nieto, J. J., and Oren, A. (1998) Biology of moderately halophilic aerobic bacteria. *Microbiol. Mol. Biol. Rev.* **62**, 504–544.
2. Vargas, C., Argandoña, M., Reina-Bueno, M., et al., (2008). Unravelling the adaptation responses to osmotic and temperature stress in *Chromohalobacter salexigens*, a bacterium with broad salinity tolerance. *Saline Systems* **4**, 14.
3. Argandoña, M., Nieto, J.J., Iglesias-Guerra, F., et al., (2010). Interplay between iron homeostasis and the osmotic stress response in the halophilic bacterium *Chromohalobacter salexigens. Appl. Environ. Microbiol.* **76**, 3575–3589.
4. De la Haba, R.R., Arahal, D.R., Márquez, M.C., et al., (2010). Phylogenetic relationships within the family *Halomonadaceae* based on 23S and 16S rRNA gene sequence analysis. *Int. J. Syst. Evol. Microbiol.* **60***,* 737–748.
5. Arahal, D.R., Ludwig, W., Schleiffer, K.H., et al., (2002). Phylogeny of the family *Halomonadaceae* based on 23S and 16S rRNA sequence analyses. *Int. J. Syst. Evol. Microbiol.* **52***,* 241–249.
6. Ventosa, A. and Nieto, J.J. (1995) Biotechnological applications and potentialities of halophilic microorganisms. *World J. Microbiol. Biotechnol.* **11**, 85–94.
7. Margesin, R. and Schinner, F. (2001) Potential of halotolerant and halophilic microorganisms for biotechnology. *Extremophiles* **5**, 73–83.
8. Da Costa, M. S., Santos, H., and Galinski, E. A. (1998) An overview of the role and diversity of compatible solutes in *Bacteria* and *Archaea. Adv. Biochem. Eng. Biotechnol.* **61**, 117–153.
9. Roberts, M.F. (2005). Organic compatible solutes of halotolerant and halophilic microorganisms. *Saline Systems* **1**, 5.
10. Pastor, J.M., Salvador, M., Argandoña, M., et al., (2010). Ectoines in cell stress protection: uses and biotechnological production. *Biotechnol. Adv.* **28**, 782–801.
11. Afendra, A.S., Vargas, C., Nieto, J.J., and Drainas, C. (2004). Gene transfer and expression of recombinant proteins in moderately halophilic bacteria. In *Methods in Molecular Biology*, vol. 267: *Recombinant Gene Expression: Reviews and Protocols*, 2nd ed. P. Balbás and A. Lorence (eds), pp. 209–223. Humana Press Inc., Totowa, NJ.
12. Arvanitis, N., Vargas, C., Tegos, G., et al., (1995) Development of a gene reporter system in moderately halophilic bacteria by employing the ice nucleation gene of *Pseudomonas syringae. Appl. Environ. Microbiol.* **61**, 3821–3825.
13. Tegos, G., Vargas, C., Perysinakis, A., et al., (2000) Release of cell-free ice nuclei from *Halomonas elongata* expressing the ice nucleation gene *inaZ* of *Pseudomonas syringae. J. Appl. Microbiol.* **89**, 785–792.
14. Frillingos, S., Linden, A., Niehaus, F., et al., (2000) Cloning and expression of α-amylase from the hyperthermophilic archaeon *Pyrococcus woesei* in the moderately halophilic bacterium *Halomonas elongata. J. Appl. Microbiol.* **88**, 495–503.
15. Nieto, J. J., Fernández-Castillo, R., Márquez, M. C., et al., (1989) A survey of metal tolerance in moderately halophilic eubacteria. *Appl. Environ. Microbiol.* **55**, 2385–2390.
16. Rodríguez-Valera, F., Ruiz-Berraquero, F., and Ramos-Cormenzana, A. (1981) Characteristics of the heterotrophic bacterial populations in hypersaline environments of different salt concentrations. *Microb. Ecol.* **7**, 235–243.
17. Vreeland, R. H., Anderson, R., and Murray, R.G.E. (1984) Cell wall and phospholipid composition and their contribution to the salt tolerance of *Halomonas elongata. J. Bacteriol.* **160**, 879–883.
18. James, S. R., Dobson, S. J., Franzmann, P. D., et al., (1990) *Halomonas meridiana*, a new species of extremely halotolerant bacteria isolated from Antartic saline lakes. *Syst. Appl. Microbiol.* **13**, 270–278.
19. Severin, J., Wohlfarth, A., and Galinski, E. A. (1992) The predominant role of recently discovered tetrahydropyrimidines for the osmoadaptation of halophilic eubacteria. *J. Gen. Microbiol.* **138**, 1629–1638.

20. Argandoña, M., Martínez-Checa, F., Llamas, I., et al., (2003). Megaplasmids in Gram-negative, moderately halophilic bacteria. *FEMS Microbiol. Lett.* **227**, 81–86.
21. Yanase, H., Sato, D., Yamamoto, K., et al., (2007). Genetic engineering of *Zymobacter palmae* for production of ethanol from xylose. *Appl. Environ. Microbiol.* **73**, 2592–2599.
22. Yanase, H., Yamamoto, K., Sato, D., et al., (2005). Ethanol production from cellobiose by *Zymobacter palmae* carrying the *Ruminocuccus albus* beta-glucosidase gene. *J. Biotechnol.* **118**, 35–43.
23. Joshi, A.A., Kanekar, P.P., Kelkar, A.S., et al.,(2007). Moderately halophilic, alkalitolerant *Halomonas campisalis* MCM B-365 from Lonar Lake, India. *J. Basic Microbiol.* **47**, 213–221.
24. Mwirichia, R., Muigai, A.W., Tindall, B., et al., (2010). Isolation and characterisation of bacteria from the haloalkaline Lake Elmenteita, Kenya. *Extremophiles* **14**, 339–348.
25. Cohen, G. N. and Rickenberg, R. H. (1956). Concentration specifique reversible des aminoacides chez *E. coli*. *Ann. Ins. Pasteur Paris* **91**, 693–720.
26. Kamekura, M., Wallace, R., Hipkiss, A. R., et al., (1985) Growth of *Vibrio costicola* and other moderate halophiles in a chemically defined minimal medium. *Can. J. Microbiol.* **31**, 870–872.
27. Cummings, E. P. and Gilmour, D. J. (1995) The effect of NaCl on the growth of a *Halomonas* species: accumulation and utilization of compatible solutes. *Microbiology* **141**, 1413–1418.
28. Nieto, J. J., Fernández-Castillo, R., García, M. T., et al., (1993) Survey of antimicrobial susceptibility of moderately halophilic eubacteria and extremely halophilic aerobic archaeobacteria: utilization of antimicrobial resistance as a genetic marker. *Syst. Appl. Microbiol.* **16**, 352–360.
29. Kunte, H. J. and Galinski, E. A. (1995) Transposon mutagenesis in halophilic eubacteria: conjugal transfer and insertion on transposon Tn*5* and Tn*1732* in *Halomonas elongata*. *FEMS Microbiol. Let.* **128**, 293–299.
30. Coronado, M. J., Vargas, C., Kunte, H. J. et al., (1995) Influence of salt concentration on the susceptibility of moderately halophilic bacteria to antimicrobials and its potential use for genetic transfer studies. *Curr. Microbiol.* **31**, 365–371.
31. Morelle, G. (1989) A plasmid extraction procedure on a miniprep scale. *BRL Focus* **11**, 7–8.
32. Sambrook, J. and Russell, D.W. (2001) *Molecular Cloning: A Laboratory Manual, 3 rd ed.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
33. Plazinski, J., Cen, Y. H., and Rolfe, B. G. (1985) General method for the identification of plasmid species in fast-growing soil microorganisms. *Appl. Environ. Microbiol.* **48**, 1001–1003.
34. Wheatcroft, R., McRae, D.G. and Miller, R.W. (1990) Changes in the *Rhizobium meliloti* genome and the ability to detect supercoiled plasmids during bacteroid development. *Mol. Plant Microbe Interact.* **3**, 9–17.
35. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1989). *Current Protocols in Molecular Biology*. Greene Publishing Associates, John Wiley & Sons, NY.
36. Marmur, J. (1961) A procedure for the isolation of deoxyribonucleic acid from microorganism. *J. Mol. Biol.* **3**, 208–218.
37. Salser, W., Gesteland, R.F. and Bolle, A. (1967) In vitro synthesis of bacteriophague lisozyme. *Nature* **215**, 588–591.
38. Monsalve, M., Mencía, M., Rojo, F., et al., (1995) Transcriptional regulation in bacteriophague phi 29: expression of the viral promoters throughout in the infection cycle. *Virology* **207**, 23–31.
39. Calderón, M.I., Vargas, C., Rojo, F., et al., (2004) Complex regulation of the synthesis of the compatible solute ectoine in the halophilic bacterium *Chromohalobacter salexigens* DSM 3043 T. *Microbiology* **150**, 3051–3063.
40. Edmonds, M., and Caramela, M.G. (1969) The isolation and characterization of adenosine monophosphate-rich polynucleotides synthesized by Ehrlich ascites cells. *J. Biol. Chem.* **244**, 1314–1324.
41. Kraegeloh, A., Amendt, B., and Kunte H,J. (2005) Potassium transport in a halophilic member of the bacteria domain: identification and characterization of the K+ uptake systems TrkH and TrkI from *Halomonas elongata* DSM 2581 T. *J. Bacteriol.* **187**, 1036–1043.
42. Schweikhard, E.S., Kuhlmann, S.I., Kunte, H.J., et al., (2010) Structure and function of the universal stress protein TeaD and its role in regulating the ectoine transporter TeaABC of *Halomonas elongata* DSM 2581(T). *Biochemistry* **49**, 2194–204.
43. Arco, Y., Llamas, I., Martínez-Checa, F., et al., (2005) *epsABCJ* genes are involved in the biosynthesis of the exopolysaccharide mauran produced by *Halomonas maura*. *Microbiology* **151**, 2841–51.

44. Argandoña, M., Martínez-Checa, F., Llamas, I., et al., (2006) A membrane-bound nitrate reductase encoded by the narGHJI operon is responsible for anaerobic respiration in *Halomonas maura*. *Extremophiles* **10**, 411–419.
45. Fernández-Castillo, R., Vargas, C., Nieto, J. J, et al., (1992) Characterization of a plasmid from moderately halophilic eubacteria. *J. Gen. Microbiol.* **138**, 1133–1137.
46. Eckhardt, T. (1978) A rapid method for the identification of plasmid DNA in bacteria. *Plasmid* **1**, 584–588.
47. Vargas, C., Férnandez-Castillo, R., Cánovas, D., et al., (1995) Isolation of cryptic plasmids from moderately halophilic eubacteria of the genus *Halomonas.* Characterization of a small plasmid from *H. elongata* and its use for shuttle vector construction. *Mol. Gen. Genet.* **246**, 411–418.
48. Vargas, C., Tegos, G., Drainas, C., et al., (1999) Analysis of the replication region of the cryptic plasmid pHE1 from the moderate halophile *Halomonas elongata*. *Mol. Gen. Genet.* **261**, 851–861.
49. Vargas, C., Tegos, G., Vartholomatos, G., et al., (1999) Genetic organization of the mobilization region of the plasmid pHE1 from *Halomonas elongata*. *Syst. Appl. Microbiol.* **22**, 520–529.
50. Mellado, E., Asturias, M.A., Nieto, J. J., et al., (1995) Characterization of the basic replicon of pCM1, a narrow-host-range plasmid from the moderate halophile *Chromohalobacter marismortui*. *J. Bacteriol.* **177**, 3433–3445.
51. Llamas, I., Del Moral, A., Béjar, V., et al., (1997) Plasmids from *Halomonas eurihalina*, a microorganism which produces an exoplysaccharide of biotechnological interest. *FEMS Microbiol. Lett.* **156**, 251–257.
52. Osman, O., Tanguichi, H., Ikeda, K., et al., (2010) Copper-resistant halophilic bacterium isolated from the polluted Maruit Lake, Egypt. *J. Appl. Microbiol.* **108**, 1459–7140.
53. Mellado, E., Nieto, J.J., and Ventosa, A. (1995) Construction of novel shuttle vectors for use between moderately halophilic bacteria and *Escherichia coli. Plasmid* **34**, 157–164.
54. Vargas, C., Coronado, M. J., Ventosa, A., et al., (1997) Host range, stability and compatibility of broad-host-range plasmids and a shuttle vector in moderately halophilic bacteria. Evidence of intragenic and intergenic conjugation in moderate halophiles. *Syst. Appl. Microbiol.* **20**, 173–181.
55. Knauf, V. C. and Nester, E. W. (1982) Wide host range cloning vectors: a cosmid clone bank of an *Agrobacterium* Ti plasmid. *Plasmid* **8**, 45–54.
56. Spaink, H. P., Okker, R. J. H., Wiffelman, C. A., et al., (1987). Promoters in the nodulation region of the *Rhizobium leguminosarum* Sym plasmid pRLJI. *Plant Mol. Biol.* **9**, 27–39.
57. Bagdasariam, M., Lurz, R., Rückert, B., et al., (1981) Specific purpose plasmid cloning vectors, II. Broad-host-range, high-copy-number RSF 1010-derived vectors and host-vector system for gene cloning in *Pseudomonas. Gene* **16**, 237–247.
58. Labes, M., Pühler, A., and Simon, R. (1990). A new family of RSF1010-derived expression and *lac*-fussion broad-host-range vectors for Gram-negative bacteria. *Gene* **89**, 37–46.
59. Leemans, J., Langenakens, J., De Greve, H., et al., (1982). Broad-host-range cloning vectors derived from the W-plasmid Sa. *Gene* **19**, 361–364.
60. Seaman, P. F., and Day, M. J. (2007) Isolation and characterization of a bacteriophage with an unusually large genome from the Great Salt Plains National Wild life Refuge, Oklahoma, USA. *FEMS Microbiol. Ecol.* **60**, 1–13
61. Ryu, H. J., Jeong, Y. J., and Park, D. H. (2004) Growth and physiological properties of wild type and mutants of *Halomonas subglaciescola* DH-1 in saline environment. *J. Microbiol.* **42**, 174–180.
62. Cánovas, D., Vargas, C., Ventosa, A., et al., (1997) Salt-sensitive and auxotrophic mutants of *Halomonas elongata* and *H. meridiana* by use of hydroxylamine mutagenesis. *Curr. Microbiol.* **34**, 85–90.
63. Llamas, I., Béjar, V., Argandoña, M., et al., (1999) Chemical mutagenesis of *Halomonas eurihalina* and selection of exopolysaccharide-deficient variants. *Biotechnol. Lett.* **21**, 367–370.
64. Ubben, D. and Schmitt, R. (1986) Tn*1721* derivatives for transposon mutagenesis, restriction mapping and nucleotide sequence analysis. *Gene* **41**, 145–152.
65. Schmitt, R., Bernhard, E., and Mattes, R. (1979) Characterization of Tn*1721*, a new transposon containing a tetracycline resistance gene capable of amplification. *Mol. Gen. Genet.* **172**, 53–65.
66. Allmeier, H., Cresnar, B., Greck, M., et al., (1992) Complete nucleotide sequence of Tn*1721*: gene organization and a novel gene product with features of a chemotaxis protein. *Gene* **111**, 11–20.
67. Göller, K., Ofer, A., and Galinski, E.A. (1998) Construction and characterization of an NaCl-sensitive mutant of *Halomonas elongata*

impaired in ectoine biosynthesis. *FEMS Microbiol. Let.* **161**, 293–300.

68. Cánovas, D., Vargas, C., Iglesias-Guerra, F., et al., (1997) Isolation and characterization of salt-sensitive mutants of the moderate halophile *Halomonas elongata* and cloning of the ectoine synthesis genes. *J. Biol. Chem.* **272**, 25794–25801.
69. Grammann, K., Volke, A., and Kunte, H. J. (2002) New type of osmoregulated solute transporter identified in halophilic members of the *Bacteria* domain: TRAP transporter TeaABC mediates uptake of ectoine and hydroxyectoine in *Halomonas elongata* DSM 2581. *J. Bacteriol.* **184**, 3078–3085.
70. Coronado, M.J., Vargas, C., Mellado, E., et al., (2000) The α-amylase gene of the moderate halophile *Halomonas meridiana*: cloning and molecular characterization. *Microbiology* **146**, 861–868.
71. Llamas, I., Argandoña, M., Quesada, E., et al., (2000) Transposon mutagenesis in *Halomonas eurihalina*. *Res. Microbiol.* **150**, 13–18.
72. Llamas. I., Suárez, A., Quesada, E., et al., (2003) Identification and characterization of the *carAB* genes responsible for encoding carbamoylphosphate synthetase in *Halomonas eurihalina*. *Extremophiles* 7, 205–211.
73. Ubben, D. and Schmitt, R. (1987) A transposable promoter and transposable promoter probes derived from Tn*1721*. *Gene* **53**, 127–134.
74. De Lorenzo, V., Herrero, M., Jakubzik, U., et al., (1990) Mini-Tn*5* transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in Gram-negative bacteria. *J. Bacteriol.* **172**, 6568–6572.
75. Horton, R. M., Hunt, H. D., Ho, S. N., et al., (1989) Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* 77, 61–68.
76. Schäfer, A., Tauch, A., Jäger, W, et al., (1994) Small mobilizable multipurpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene* **145**, 69–73.
77. Tokunaga, H., Ishibashi, M., Arisaka, F., et al., (2008) Residue 134 determines the dimer–tetramer assembly of nucleoside diphosphate kinase from moderately halophilic bacteria. *FEBS Letters* **582**, 1049–1054.
78. Prentki, P. and Krisch, H. M. (1984) In vitro insertional mutagenesis with a selectable DNA fragment. *Gene* **29**, 303–313.
79. Quandt, J. and Hynes, M. (1993) Versatile suicide vectors which allow direct selection for gene replacement in Gram-negative bacteria. *Gene* **127**, 15–21.
80. Kessler, B., De Lorenzo, V., and Timmis, N. K. (1992) A general system to integrate *lacZ* fusions into the chromosome of Gram negative bacteria: regulation of the Pm promoter of the TOL plasmid studied with all controlling elements in monocopy. *Mol. Gen. Genet.* **233**, 293–301.
81. Douka, E., Christogianni, A., Koukkou, A. I., et al., (2001) Use of a green fluorescent protein as a reporter in *Zymomonas mobilis* and *Halomonas elongata*. *FEMS Microbiol. Lett.* **201**, 221–227.
82. Lindgren, P. B., Frederick, R., Govindarajan, A. G., Panopoulos, N. J., Staskawicz, B. J., and Lindow, S. E., et al., (1989) An ice nucleation reporter gene system: identification of inducible pathogenicity genes in *Pseudomonas syringae* pv. *Phaseolicola*. *EMBO J.* **8**, 2990–3001.
83. Tegos, C., Vargas, C., Vartholomatos, G., et al., (1997) Identification of a promoter region on the *Halomonas elongata* plasmid pHE1 employing the *inaZ* reporter gene of *Pseudomonas syringae*. *FEMS Microbiol. Lett.* **154**, 45–51.
84. Chalfie, M., Tu, Y., Euskirchen, G., et al., (1994) Green fluorescent protein as a marker for gene expression. *Science* **263**, 802–805.
85. Cormark, B. P., Valdivia, R. H., and Falkow, S. (1996) FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173**, 33–38.
86. Goldrick, M., and Kessler, D. (2003) RNA analysis by nuclease protection, in *Current Protocols in Neuroscience* (Gerfen, C., ed.), John Wiley and Sons, Inc., pp. 5.1–5.30.
87. Fekete, R.A., Miller, M.J. and Chattoraj, D.K (2003) Fluorescently labeled oligonucleotide extension, a rapid and quantitative protocol for primer extension. *Biotechniques* **35**, 90–94.
88. Lloyd, A.L., Marshall, B.J. and Mee, B.J. (2004) Identifying cloned *Helicobacter pylori* promoters by primer extension using a FAM-labeled primer and GeneScanR analysis. *J. Microbiol. Meth.* **60**, 291–298.
89. Ducey, T.F., Jackson, L., Orvis, J. et al., (2009) Transcript analysis of *nrrF*, a Fur repressed sRNA of *Neisseria gonorrhoeae*. *Microb. Pathog.* **46**, 166–170.
90. Tobias, N.J., Seemann, T., Pidot, S.J., et al., (2009) Mycolactone gene expression is controlled by strong SigA-like promoters with utility in studies of *Mycobacterium ulcerans* and Buruli Ulcer. *PLoS Negl. Trop. Dis.* **3**, e553.
91. Schweikhard, E.S., Kuhlmann, S.I., Kunte, H.J., et al., (2010) Structure and function of the universal stress protein TeaD and its role in

regulating the ectoine transporter TeaABC of *Halomonas elongata* DSM 2581(T). *Biochemistry* **49**, 2194–2204.

92. Rodríguez-Sáiz, M., Sánchez-Porro, C., De La Fuente, J., et al., (2007) Engineering the halophilic bacterium *Halomonas elongata* to produce β-carotene. *Appl. Microbiol. Biotechnol.* 77, 637–643.
93. Kim, D., Kim, S.W., Choi, K.Y., et al., (2008) Molecular cloning and functional characterization of the genes encoding benzoate and *p*-hydroxybenzoate degradation by the halophilic *Chromohalobacter* sp. strain HS-2. *FEMS Microbiol. Lett.* **280**, 235–241.
94. Krejcik, Z., Hollemeyer, K., Smits, T.H. et al., (2010) Isethionate formation from taurine in *Chromohalobacter salexigens*: purification of sulfoacetaldehyde reductase. *Microbiology* **156**, 1547–1555.
95. Yuste, L., Canosa, I. and Rojo, F. (1998) Carbon-source-dependant expression of the *PalkB* promoter from *Pseudomonas oleovorans* alkane degradation pathway. *J. Bacteriol.* **180**, 5218–5226.

Chapter 10

Regulated Recombinant Protein Production in the Antarctic Bacterium *Pseudoalteromonas haloplanktis* TAC125

Valentina Rippa, Rosanna Papa, Maria Giuliani, Cinzia Pezzella, Ermenegilda Parrilli, Maria Luisa Tutino, Gennaro Marino, and Angela Duilio

Abstract

This review reports results from our laboratory on the development of an effective inducible expression system for the homologous/heterologous protein production in cold-adapted bacteria. Recently, we isolated and characterized a regulative genomic region from *Pseudoalteromonas haloplanktis* TAC125; in particular, a two-component regulatory system was identified. It is involved in the transcriptional regulation of the gene coding for an outer membrane porin (PSHAb0363) that is strongly induced by the presence of L-malate in the growth medium.

We used the regulative region comprising the two-component system located upstream the *PSHAb0363* gene to construct an inducible expression vector – named pUCRP – under the control of L-malate. Performances of the inducible system were tested using the psychrophilic β-galactosidase from *P. haloplanktis* TAE79 as model enzyme to be produced. Our results demonstrate that the recombinant cold-adapted enzyme is produced in *P. haloplanktis* TAC125 in good yields and in a completely soluble and catalytically competent form. Moreover, an evaluation of optimal induction conditions for protein production was also carried out in two consecutive steps: (1) definition of the optimal cellular growth phase in which the gene expression has to be induced; (2) definition of the optimal inducer concentration that has to be added in the growth medium.

Key words: Cold-adapted bacteria, Regulated gene expression system, *Pseudoalteromonas haloplanktis*, Malate induction, Minimal medium

1. Introduction

Incorrect folding of the nascent polypeptide chains is one of the main problems occurring during heterologous protein production in bacteria. A key role in this process is played by intermolecular hydrophobic interactions taking place among partly folded intermediates

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_10, © Springer Science+Business Media, LLC 2012

that cause protein molecules to stick together thus driving them away from the productive folding pathway (1). Since formation of inclusion bodies often impairs the recombinant production of valuable proteins, many experimental approaches have been explored to minimize this undesirable effect, including expression of chimerical proteins (2) and co-expression with chaperonines (3). Expression of "difficult" proteins has also been carried out by lowering the temperature at the physiological limit allowed for the growth of mesophilic host organisms (between 15 and 18°C for *Escherichia coli*). Lowering the temperature, in fact, has a pleiotropic effect on the folding process, destabilizing the hydrophobic interactions occurring during intermediates aggregation (4). Although in some cases this approach has been reported to increase yields of soluble and active recombinant protein products, the exploitation of an industrial process performed at suboptimal growth condition of the expression host might hardly be considered. A rational alternative to mesophilic organisms is the use of naturally cold-adapted bacteria as hosts for protein production at low temperature (even at around 0°C). *Pseudoalteromonas haloplanktis* TAC125 is a Gram-negative bacterium isolated from an Antarctic coastal seawater sample collected in the vicinity of the French Antarctic station Dumont D'Urville, Terre Adélie. It can be classified as a Eurypsychrophile i.e. a bacterium growing in a wide range of low temperatures (5); and it was the first Antarctic Gram-negative bacterium of which the genome was fully sequenced and carefully annotated (6). Genomic and metabolic features of this bacterium, accounting for its remarkable versatility and fast growth compared with other bacteria from aqueous environments, were discovered by combining genome sequencing and further in silico and in vivo analyses. *P. haloplanktis* TAC125 is able to duplicate in a wide range of temperatures (0–30°C), with an apparent optimal growth temperature at 20°C, where the observed duplication time in rich medium is 31 min (7). However, the bacterium still duplicates at fast speed even at lower temperatures (at 4°C, one cell division is completed in about 100 min; unpublished results) and, when provided with sufficient nutrients and aeration, it grows to very high density (up to $OD_{600} = 20$) under laboratory settings, even at 0°C. This growth performance makes it one of the faster growing psychrophiles so far characterized. Fast growth rates, combined with the ability of *P. haloplanktis* TAC125 to reach very high cell densities even under laboratory growth conditions and to be easily transformed by intergeneric conjugation (8), made this bacterium an attractive host for the development of an efficient gene expression system at low temperatures.

By combining mesophilic and psychrophilic genetic signals a collection of psychophilic gene expression vectors was set up to produce recombinant proteins in *P. haloplanktis* TAC125. The mesophilic signals consist of the pUC18-derived origin of replication (*OriC*)

and a selection marker gene (a β-*lactamase* encoding gene), allowing the plasmid to replicate either in *E. coli* or in the psychrophilic host. Another crucial mesophilic signal is represented by the *OriT* sequence, the conjugational DNA transfer origin from the broad host range plasmid pJB3 (9). Structural and functional studies led to the isolation of the psychrophilic origin of replication (*OriR*) from the *P. haloplanktis* TAC125 endogenous plasmid pMtBL (10).

This system has made it possible the isolation of constitutive psychrophilic promoters (11). The structural/functional characterization of *P. haloplanktis* TAC125 promoters was carried out by random cloning of genomic DNA fragments and identification of promoter sequences by evaluating their capability to express a promoter-less reporter gene (11). By this promoter-trap strategy, a collection of constitutive psychrophilic promoters showing different strengths at different temperatures was identified. The implementation of the above described psychrophilic promoters in the pMtBL-derived shuttle vectors resulted in the set up of cold-adapted gene-expression systems, characterized by the constitutive production of the recombinant protein.

The described expression system represented the first example of heterologous protein production based on a true cold-adapted replicon (11). However, the development of an effective cold expression system needs to be finely tuned possibly using ad hoc promoters. Physical separation between bacterial growth phase and expression of the desired proteins, in fact, cannot only improve the productivity of the entire system but can also play an important role in the production of proteins toxic for the host cells. Sometimes efficient production can only be achieved by fine tuning the recombinant gene expression. This goal can be reached by using regulated promoters and efficient induction strategies.

Recently, using a proteomic approach and taking advantage from the genome sequence of *P. haloplanktis* TAC125 (6) we isolated and characterized a functionally active two-component system. The regulatory system (encompassing the genes *PSHAb0361–PSHAb0362*) is involved in the transcriptional regulation of the gene coding for an outer membrane porin (PSHAb0363), and it is strongly induced by the presence of L-malate in the medium (12). We used the regulative region comprising the two-component system located upstream the *PSHAb0363* gene to construct an inducible expression vector, named pUCRP that is under the control of L-malate (13). Performances of the inducible system were tested for protein production using a psychrophilic protein as model system. Moreover, an evaluation of optimal induction conditions for protein production was also carried out. Data presented demonstrated that psychrophilic β-galactosidase from *P. haloplanktis* TAE79 (14) is produced in *P. haloplanktis* TAC125 in good yields and in a completely soluble and catalytically competent form.

2. Materials

2.1. Bacterial Strains

1. *P. haloplanktis* TAC125 (see Note 1).
2. *E. coli* DH5α [*supE*44, Δ*lac*U169 (ϕ80 *lacZ*ΔM15) *hsdR*17, *recA*1, *endA*1, *gyrA*96, *thi*-1, *relA*1).
3. *E. coli* strain S17-1(λ*pir*) [*thi*, *pro*, *hsd* (r^- m^+) *recA*::RP4-2-TC^r::Mu Km^r::Tn7 Tp^r Sm^r λ*pir*) (15) was used as donor in intergeneric conjugation experiments.

2.2. Liquid Media

1. LB medium (16): 10 g Bacto-tryptone, 5 g bacto-yeast, 10 g NaCl, and 950 mL deionized H_2O. Add agar 1.5% (w/v) to the media for production of agar plates. Sterilize by autoclaving for 20 min at 1 atm on liquid cycle. Add 100 μg/mL ampicillin if required.
2. TYP medium (17): 16 g Bacto-tryptone, 16 g bacto-yeast, and 10 g marine mix. Shake until the solutes have dissolved. Adjust the pH to 7.5 with 5 M NaOH, and adjust the volume of the solution to 1 L with deionized H_2O. Add agar 1.5% (w/v) to the media for production of agar plates. Sterilize by autoclaving for 20 min at 1 atm on liquid cycle. Add 100 μg/mL ampicillin if required.
3. Schatz medium (12): 1 g KH_2PO_4, 1 g NH_4NO_3, 10 g NaCl, 0.2 g $MgSO_4 \cdot 7H_2O$, 0.01 g $FeSO_4 \cdot 7H_2O$, 0.01 g $CaCl_2 \cdot 2H_2O$. Shake until the solutes have dissolved. Adjust the pH to 7.0 with 5 M NaOH. Adjust the volume of the solution to 1 L with dH_2O. Sterilize by filtering through a 0.22-μm filter and supplement with casamino acid 0.5% (w/v) and L-malate 0.2% (w/v) for promoter induction. Add 100 μg/mL ampicillin if required.
4. Bacteria can be stored indefinitely in cultures containing glycerol at low temperature (either −20 to −70°C stocks) (see Note 2).

2.3. Kits and Reagents for Molecular Biology

1. Phusion® High-Fidelity DNA Polymerase (Finnzymes).
2. Restriction enzymes (Promega).
3. Calf Intestinal Alkaline Phosphatase, CIP (Roche).
4. T4 DNA Ligase (Promega).
5. Lambda DNA/*Eco*RI + *Hind*III marker (Fermentas).
6. ChargeSwitch® gDNA Mini Bacteria Kit (Invitrogen).
7. High Pure PCR Product Purification Kit (Roche).
8. QIAprep® Spin Miniprep Kit (Qiagen).
9. QIAquick® Nucleotide Removal Kit (Qiagen).
10. Bradford Protein Assay Dye Reagent Concentrate 5× (Bio-Rad).
11. Albumin from bovine serum, BSA (Sigma-Aldrich).

2.4. Buffers and Solutions

1. 100 mg/mL ampicillin stock solution: Dissolve 1 g of Ampicillin powder in 8 mL of dH_2O. Adjust the volume of the solution to 10 mL with deionized H_2O and sterilize by filtration through a 0.22-μm sterile filter. Split the obtained stock solution in ten aliquots of 1 mL each in sterile polypropylene tubes and store them at −20 C.
2. 20% (w/v) L-Malate stock solution: Dissolve 10 g of L-malic acid in 40 mL of deionized H_2O. Adjust the pH to 7.0 with 5 M NaOH. Adjust the volume of the solution to 50 mL with ddH_2O and sterilize by filtration through a 0.22-μm sterile filter. Store at 4°C up to 2 months.
3. 50× TAE stock solution: 242 g of Tris base, 57.1 mL of glacial acetic acid, and 100 mL of 0.5 M EDTA pH 8.0. Adjust the volume of the solution to 1 L with dH_2O. Store at room temperature (RT) up to 1 year.
4. 1× TAE buffer: 40 mM Tris-acetate pH 8.0, 1 mM EDTA pH 8.0. Make a 50× TAE stock solution by mixing 0.5 M EDTA pH 8.0: dissolve 186.1 g of EDTA in 800 mL of deionized H_2O. Adjust the pH to 8.0 with NaOH (about 20 g of NaOH pellets) and adjust the volume of the solution to 1 L with dH_2O.
5. Agarose gel-loading buffer: 0.25% Bromophenol blue, 0.25% xylene cyanol FF, and 30% glycerol (Fermentas).
6. 10 mg/mL Ethidium bromide stock solution: add 1 g of ethidium bromide to 100 mL of dH_2O. Stir on a magnetic stirrer for several hours to ensure that the dye has dissolved. Wrap the container in aluminium foil and store at 4°C.
7. Lysis buffer: 100 mM Sodium phosphate pH 7.8, 5 mM EDTA pH 8.0, 2% Triton X-100. Prior to use, add: 1 mM PMSF, 1 mM dithiothreitol (DTT), and 5 mg/mL lysozyme.
8. 0.5 M Sodium phosphate buffer: Dissolve 68.9 g of $NaH_2PO_4 \cdot H_2O$ in 900 mL of dH_2O. Adjust the pH to 7.8 with NaOH and add dH_2O to make up a final volume of 1 L.
9. 10 mM PMSF: Dissolve PMSF in ethanol at a concentration of 1.74 mg/mL. Divide the solution into aliquots and store at −20°C.
10. 1 M DTT: Dissolve 3.09 g of DTT in 20 mL of deionized H_2O. Sterilize by filtration. Dispense into 1 mL aliquots and store them at −20°C.
11. 1× SDS-PAGE loading buffer: 62.5 mM Tris–HCl pH 6.8, 10% glycerol, 2% SDS, 100 mM DTT, and 0.1% Bromophenol blue. This buffer lacking DTT can be stored at RT. DTT should then be added just before that the buffer is used from a 1 M stock.
12. 0.5 M Tris–HCl pH 6.8: Dissolve 60.55 g of Tris base in 800 mL of deionized H_2O. Adjust the pH to 6.8 with HCl and add dH_2O to make up a final volume of 1 L.

13. 0.5% (w/v) Bromophenol blue: Dissolve 0.25 g bromophenol blue powder in 45 mL of dH_2O. Shake well to dissolve the dye and then adjust the volume of the solution to 50 mL with dH_2O. Store at RT.
14. 5× Running buffer: Dissolve 15.1 g of Tris base, 94 g of glycine, and 5 g of SDS in 900 mL of deionized H_2O. Adjust the volume of the solution to 1 L with dH_2O.
15. Buffer for β-galactosidase assays: 100 mM sodium phosphate pH 7.3, 1 mM $MgCl_2$, 100 mM β-mercaptoethanol, 0.9 mg/mL 2-nitrophenyl-β-D-galactopyranoside (ONPG).
16. 0.5 M Phosphate buffer: Dissolve 68.9 g of $NaH_2PO_4 \cdot H_2O$ in 900 mL of dH_2O. Adjust the pH to 7.3 with NaOH and add H_2O to make up a final volume of 1 L.
17. 1 M $MgCl_2$: Dissolve 9.5 g of $MgCl_2$ in 90 mL of deionized H_2O. Then add H_2O to make up a final volume of 100 mL.

3. Methods

3.1. Development of a Cold Regulated Expression System

3.1.1. Construction of the pUCRP Expression Vector

The pUCRP inducible expression vector is constructed by cloning the DNA fragment corresponding to the cold-active P *(PSHAb0363)* promoter region (isolated as described in 12) that is under the control of L-malate into pUCLT/Rterm vector (Fig. 1, Table 1, 13). The pUCLT/Rterm plasmid, deriving from the pUC18 plasmid, is characterized by the presence of: (1) the pJB3-derived *OriT* (9), a DNA fragment responsible for the initiation of the conjugative transfer between *E. coli* S17-1 λ*pir* strain (donor) and the psychrophilic cells (acceptor); (2) a pUC18 derived polylinker wherein the target gene can be cloned; (3) *E. coli blaM* gene, encoding a mesophilic β-lactamase, which is used for the selection of the recombinant clones; (4) *OriC*, the origin of replication allowing the plasmid to replicate in *E. coli*; (5) the *T/R* box, a DNA fragment containing the cold-adapted origin of replication (*OriR*); (6) the T*aspC*, the transcription termination signal of the aspartate aminotransferase gene (*aspC*) isolated from *P. haloplanktis* TAC125.

1. The P *(PSHAb0363)* promoter region is amplified from *P. haloplanktis* TAC125 DNA genomic by PCR using Phusion® High-Fidelity DNA Polymerase according to the manufacturer's protocols (Finnzymes) (see Note 3). PCR reactions carried out by using specific oligonucleotide pairs listed in Table 1 generate three DNA fragments:
   (a) The insert *a* (~1,700 bp) is generated by using the primers PSHAb0363a Fw and PSHAb0363a Rv containing the proper overhangs of *Hind*III and *Cla*I restriction sites (underlined in Table 1);

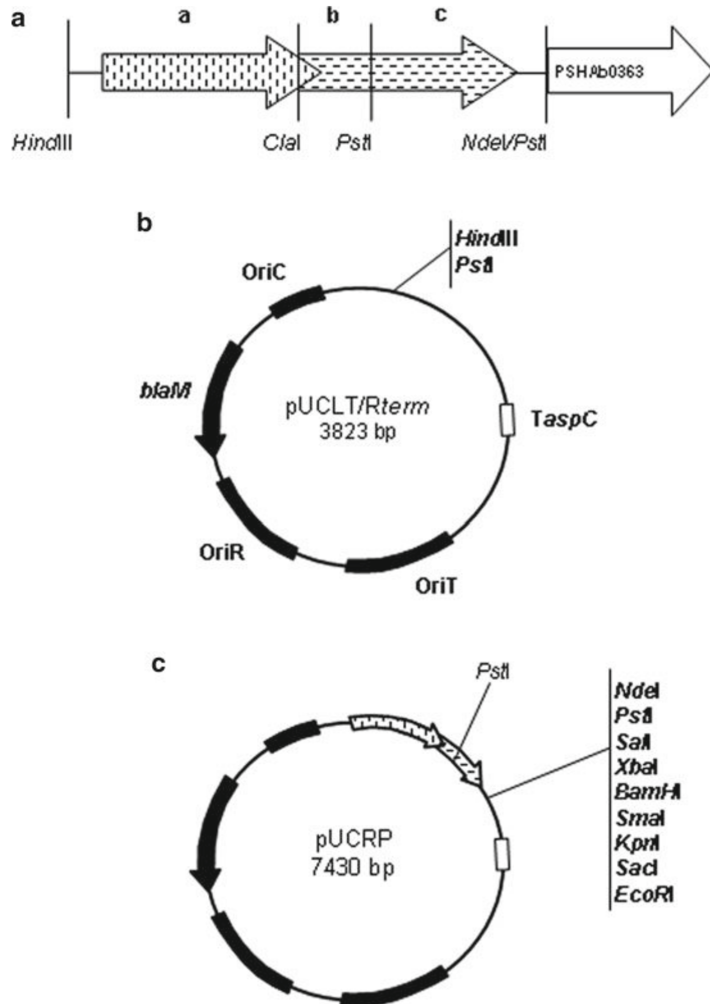


Fig. 1. Construction of pUCRP expression vector. (**a**) *P. haloplanktis* TAC125 DNA genomic fragment corresponding to the cold-active promoter of P(PSHAb0363) (13). (**b**) pUCLT/R*term* promoter-less vector. (**c**) pUCRP expression vector. Adapted from ref. 13.

(b) The DNA fragment b (750 bp) is amplified by using PSHAb0363b Fw and PSHAb0363b Rv as primers;

(c) The fragment c containing the 3′ region of P(PSHAb0363) promoter is amplified using PSHAb0363c Fw and PSHAb0363c Rv as primers.

The primers PSHAb0363a Rv, PSHAb0363b Fw, PSHAb0363b Rv, and PSHAb0363c Fw are designed to introduce an internal restriction site to join the fragments, while PSHAb0363a Fw and PSHAb0363c Rv contain the restriction sites for their cloning into pUCLT/R*term* vector.

1. The three amplification products are purified using a commercial purification kit (Roche) and then digested with the opportune restriction enzymes (Promega) underlined in Table 1 (see Note 4).

Table 1
Plasmids and oligonucleotides

Plasmids	Description	References
pSP79β-gal	Genomic clone containing the *P. haloplanktis* TAE79 *lacZ* gene and its flanking regions	(14)
pUCLT/R*term*	Vector deriving from the pUC18 plasmid, containing the T/R box and the transcription termination signal from *P. haloplanktis* TAC125 *aspC* gene	(19)
pUCRP	pUCLT/R*term* containing PSHAb0363 promoter region	(13)
pUCRPLACZ	pUCRP containing *P. haloplanktis* TAE79 *lacZ* gene	(13)
Oligonucleotides		
PSHAb0363a Fw	5′ CAAAGCTAGGCAAAGCTTAATTATAC 3′	
PSHAb0363a Rv	5′ CCTGGATCCAATATCGATAGTTTTACG 3′	
PSHAb0363b Fw	5′ GATGGACGTCTAGAACTATCGATATTAG 3′	
PSHAb0363b Rv	5′ CCTTCAATCTAGATATCTGCAGGAGTATC 3′	
PSHAb0363c Fw	5′ GATACTCCTGCAGATATATTAATTG 3′	
PSHAb0363c Rv	5′ GTTTAACGTGTCTGCAGTTTATCATATGGTGTCC 3′	
BG1	5′ CATTCTAGATATCGATTTATGCAAGGAATAAACATG 3′	
BG2	5′ CCCGGATCCAATAACGACTCATCGACC 3′	

2. The pUCLT/R*term* vector is digested with *Hind*III and *Pst*I.
3. The 5′ phosphate groups of the cleaved vector are dephosphorylated by treatment with calf intestinal alkaline phosphatase (CIP, Roche) (0.5 U/pmol of 5′phosphate ends) (see Note 5) for 15 min at 37°C and 45 min at 55°C by using the appropriate buffer delivered with the enzyme. The CIP is heat-inactivated at 75°C for 10 min.
4. The dephosphorylated DNA is then loaded on a 1% agarose gel (containing ethidium bromide as fluorescent marker for the migrating DNA). The DNA is cut out of the gel and purified using a commercial gel-purification kit (Qiagen) following the manufacturer's instructions.
5. The cleaved dephosphorylated vector is then ligated to the digested amplification products by using two consecutive ligation reactions, by the means of T4 DNA ligase (Promega) according to the supplier's instructions.
6. The ligation reaction mixture is used directly for transformation of the chemically competent bacteria (*E. coli*, strain DH5α) according the procedure described by Duilio and co-workers (8).

7. Recombinant clones are selected on LB agar plates containing 100 μg/mL ampicillin as selective marker.
8. Plasmids are isolated from amp^R clones (QIAprep® Spin Miniprep kit, Qiagen) and the presence of the appropriate insert is verified by restriction digestion analysis (see Note 6).
9. Finally the nucleotide sequences of the inserts are checked by DNA sequencing to rule out the occurrence of any mutation during synthesis.
   The resulting vector pUCRP (see Note 7) is shown in Fig. 1.

3.1.2. Procedure to Grow Bacterial Cells

The following procedure is used to perform cellular growths of *P. haloplanktis* TAC125 recombinant strain in liquid culture at 15°C in minimal medium (Schatz) containing 100 μg/mL ampicillin.

1. Using a sterile wire, streak *P. haloplanktis* TAC125 recombinant strain directly from a frozen stock (stored at −80°C in freezing medium) onto the surface of a TYP agar plate supplemented with 100 μg/mL ampicillin. Incubate the plate for 16 h at 15°C.
2. Transfer one well-isolated colony into 3 mL of TYP medium pH 7.5 containing 100 μg/mL ampicillin in a loosely capped 15-mL tube. Grow the culture under continuous rotary shaking in aerobic conditions at 15°C for 24 h. Carry out a tenfold dilution of the bacterial culture in Schatz medium supplemented with 100 μg/mL ampicillin and casamino acid 0.5% (w/v).
3. Grow the cells for 24 h at 15°C, measure the optical density at 600 nm and use the starter culture to inoculate 50 mL of Schatz containing 100 μg/mL ampicillin, casamino acid 0.5% (w/v), and L-malate 0.2% to obtain an initial OD_{600} of about 0.08 (see Note 8). Incubate the culture under continuous rotary shaking at 15°C. To monitor the cellular growth, determine the OD_{600} every 30 min.

3.1.3. Preparation of Protein Extracts

1. Harvesting of bacteria
   (a) During the growth, harvest cells equivalent to a total OD_{600} of 25.
   (b) Centrifuge the samples at 200 × *g* for 15 min at 4°C in a microfuge.
   (c) Remove the medium by aspiration, leaving the bacterial pellet as dry as possible.
   (d) Store the pellets at −80°C.
2. Lysis of bacteria
   (a) Resuspend the bacterial pellet in 1 mL of lysis buffer (100 mM sodium phosphate pH 7.8, 5 mM EDTA pH 8.0, 2% Triton X-100, 5 mg/mL lysozyme, 1 mM DTT, and 1 mM PMSF) by vortexing (see Notes 9 and 10).

(b) Incubate the samples for 20 min in a water bath set at 15°C.

(c) Centrifuge at 8,200 × *g* for 20 min at 4°C in a microfuge. Transfer the supernatant to a fresh tube.

(d) Determine protein concentration with the Bio-Rad protein assay (18), using bovine serum albumin as standard.

(e) Store the protein extracts at −80°C (see Note 11).

3.2. Recombinant β-Galactosidase Protein Production

The *P. haloplanktis* TAE79 *lacZ* gene encoding the β-galactosidase (14) is inserted into the pUCRP expression vector by using the following procedure:

3.2.1. Production and Detection of Recombinant Psychrophilic β-Galactosidase

1. A 280-bp fragment is amplified by PCR by using the primers BG1 and BG2 and the genomic subclone pSP79β-gal as template, Table 1, (14). The amplification product is digested with *Xba*I and *Hsp*92I.
2. A 3.2-kb fragment, containing the 3′ region of *P. haloplanktis* TAE79*lac*Z gene, is recovered from pSP79β-gal clone (14) by a double *Hsp*92I/*BamH*I digestion.
3. The two fragments are ligated into the pUCRP vector, previously digested with *Xba*I and *BamH*I, generating the plasmid pUCRPLACZ.
4. A DNA sequencing reaction is performed on the pUCRPLACZ amplified region to rule out the occurrence of mutations introduced by the PCR synthesis.

The production of the recombinant β-galactosidase is performed using the following procedure:

1. The resulting vector pUCRPLACZ is mobilized into *P. haloplanktis* TAC125 cells by interspecific conjugation (8) (see Note 12).
2. The psychrophilic transconjugants are grown in liquid culture in minimal medium in the presence and in the absence of L-malate at 15°C up to the stationary phase.
3. An aliquot of each culture is collected and the soluble protein content is analyzed by SDS gel electrophoresis.

In these experiments we found that under the control of the inducible expression system, β-galactosidase is produced in *P. haloplanktis* TAC125 only in the presence of L-malate, indicating an undetectable basal activity of the inducible P(PSHAb0363) promoter. Figure 2 shows the Coomassie blue-stained gel corresponding to the protein extracts from *P. haloplanktis* TAC125 recombinant cells grown in the absence and in the presence of L-malate. The presence of a strong 118-kDa extra-band could clearly be detected in the cellular extract grown in the presence of the inducer (lane C). In the absence of L-malate, only a tiny band with the same electrophoretic mobility was observed (lane B). The unambiguous identification and a detailed structural characterization of the expressed β-galactosidase were

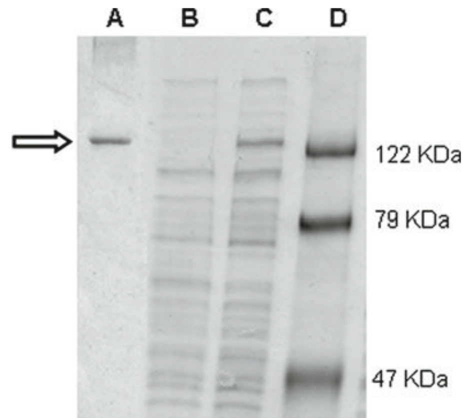


Fig. 2. Recombinant production of the thermally labile β-galactosidase from *Ph*TAE79 in *P. haloplanktis* TAC125 cells. 7.5% SDS PAGE gel electrophoresis of protein extracts from *P. haloplanktis* TAC125 cells harboring pUCRPLACZ and grown in minimal medium in the absence (lane B) and in the presence (lane C) of L-malate, in comparison with β-galactosidase from *P. haloplanktis* TAE79 used as control (lane A); molecular weight marker (lane D). Adapted from ref. 13.

obtained by mass spectrometric analysis of the corresponding Coomassie blue-stained protein band (Fig. 2, lane C) following in situ digestion with trypsin. This structural characterization demonstrated that the recombinant product is the expected β-galactosidase from *P. haloplanktis* TAE79.

The catalytic activity of the recombinant β-galactosidase is tested by the following procedure:

1. *P. haloplanktis* TAC125 recombinant cells harbouring pUCRPLACZ are grown in liquid culture at 4 and 15°C both in the absence and in the presence of L-malate up to the stationary phase.
2. Since wild type *P. haloplanktis* TAC125 does not show any detectable β-galactosidase activity, the lysate from recombinant cells harbouring the void vector is used as negative control.
3. Cellular pellets (corresponding to 25 OD_{600nm}) are collected during the growth.
4. The cells are lysated and the protein extracts are saved.
5. Recombinant cold-active β-galactosidase is spectrophotometrically assayed at 25°C as previously reported (14, see Notes 13–15). The protein activity is calculated on the basis of the extinction coefficient for *o*-nitrophenol of 3.5 mM^{1} cm^{-1} at 410 nm and specific activity of 138.2 U/mg purified enzyme (14).

The specific activity of the recombinant enzyme is reported in Table 2. Both at 4 and 15°C enzymatic activity greatly increases in the presence of L-malate, being the largest increase detected at 15°C.

Table 2
β-Galactosidase-specific activity (U/mg) from *P. haloplanktis* TAC125 recombinant cells at 4 and 15°C grown in minimal medium (MM) in the presence and in the absence of L-malate, respectively

	β-Galactosidase-specific activity (U/mg) ± SD	
	4°C	15°C
MM	0.21 ± 0.05	0.86 ± 0.02
MM + L-malate	2.12 ± 0.05	11.42 ± 1.05

Data shown are the average of 12 measurements. *SD* standard deviation

These data demonstrated that a soluble psychrophilic β-galactosidase is efficiently produced in *P. haloplanktis* TAC125 in a catalytically competent form (Table 2).

3.2.2. Induction Conditions

Once the production of a recombinant biologically active β-galactosidase in *P. haloplanktis* TAC125 was defined, different induction conditions were tested to optimize the productivity of the pUCRP inducible system (see Note 16).

This procedure consists of the two following steps:

1. Definition of the optimal cellular growth phase in which expression of the heterologous gene has to be induced (definition of t_0);
2. Definition of the optimal inducer concentration to obtain higher levels of expression of the psychrophilic gene (definition of C_0).

1(a). *P. haloplanktis* TAC125 recombinant cells are grown in minimal medium at 15°C.

1(b). Induction of psychrophilic gene expression is tested by adding L-malate 0.2% at four different phases during cellular growth: early (t_0 = 5 h; OD_{600} = 0.32 ± 0.03) and mid (t_0 = 7 h; OD_{600} = 0.53 ± 0.04) exponential phase, early (t_0 = 10 h; OD_{600} = 1.31 ± 0.08) and mid (t_0 = 23 h; OD_{600} = 1.92 ± 0.10) stationary phase.

1(c). Cellular pellets are collected at 2, 8, and 24 h from induction.

1(d). The cells are lysated and the protein extracts are saved.

1(e). Recombinant β-galactosidase activity is monitored before the addition of the inducer and after the induction (t_1).

Figure 3a shows the yield of β-galactosidase, expressed as milligram of protein/L (14), and the Induction Ratios (IR) obtained at the four different phases calculated as a ratio of β-galactosidase yield at a specific time after the induction (t_1) with respect to that before induction (t_0).

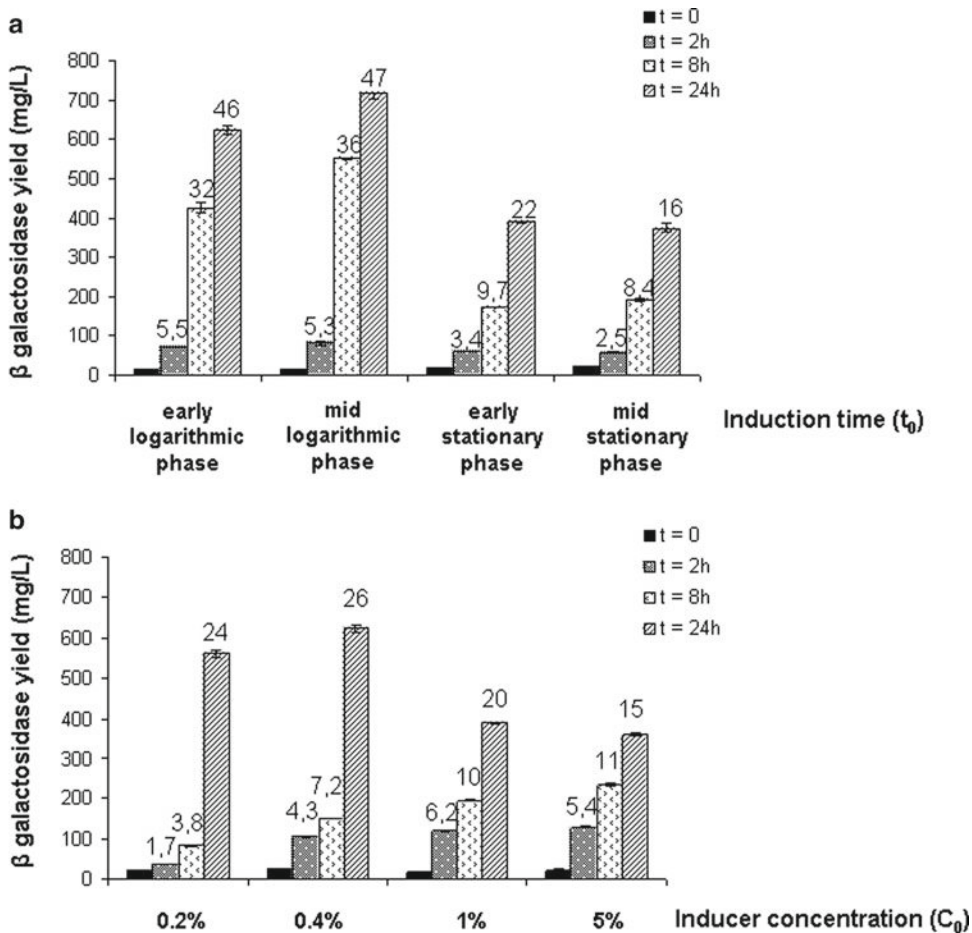


Fig. 3. Definition of optimal conditions of induction for the β-galactosidase production. (**a**) Definition of induction time (t_0). β-galactosidase production yield in *P. haloplanktis* TAC125 cells harbouring pUCRPLACZ grown in minimal medium (MM) obtained by adding L-malate 0.2% at different time (t_0) during cellular growth curve. (**b**) Definition of inducer concentration (C_0). β-galactosidase production yield in *P. haloplanktis* TAC125 cells harbouring pUCRPLACZ grown in minimal medium (MM) obtained by adding L-malate at mid-exponential phase at four different concentrations (C_0). β-galactosidase activity was monitored before the addition of inducer (as negative control of the experiment), and 2-8-24 h after the induction (t_1). *Numbers above bars* refer to induction ratios (t_1/t_0). Means and standard deviations have been calculated from four independent assays. Adapted from ref. 13.

The best induction ratios (t_1/t_0) are observed when L-malate is added during the exponential phase, with the highest ratio detected in the mid-exponential phase. High levels of recombinant enzyme can be already detected on cell lysates after 8 h from induction, but maximal β-galactosidase production is recorded at $t_1 = 24$ h from induction (716 mg/L).

The second step in the optimization procedure consists in the definition of the optimal concentration of inductor (C_0).

2(a). *P. haloplanktis* TAC125 recombinant cells are grown in minimal medium at 15°C.

2(b). Four different concentrations of L-malate (0.2%, 0.4%, 1% and 5%) are added to recombinant *P. haloplanktis* TAC125 cells, at mid-exponential phase.

2(c). Cell lysates are collected at 2, 8, and 24 h from induction.

2(d). The β-galactosidase activity is assayed before the addition of the inducer and after the induction (t_1).

Yields of β-galactosidase and calculated induction ratios in these experiments are reported in Fig. 3b. Lower concentrations of inducer yielded the highest IR values with the maximum effect obtained using 0.4% L-malate. In these experiments, the results obtained at 0.2% L-malate are slightly different from those reported in Fig. 3a possibly due to a slightly higher basal β-galactosidase activity, although the final enzyme production at 24 h is very similar (561 mg/L). Higher concentration of L-malate resulted in a strong reduction in the IR value, probably due to an effect of the dicarboxylic acid on the osmotic equilibrium of the cell; in particular when 5% L-malate is added to recombinant cells, a decrease of cellular growth is observed (data not shown). Also in this case, the highest production is obtained when cell lysates were collected 24 h after the induction (620 mg/L).

4. Notes

1. Strain was kindly provided by C. Gerday (University of Liege, Belgium). The strain was isolated from the sea water in the surrounding of the Dumont d'Urville Antarctic station (66°40′ S, 40°01′ E) during the 1988 summer campaign of the "Expeditions Polaires Française" in Terre Adélie.
2. To store bacteria, pick a single, well-isolated colony growing on the surface of an agar plate, with a sterile inoculating needle, dip the needle into the liquid broth and grow the bacterial culture up to medium-late exponential phase. To 0.85 mL of bacterial culture, add 0.15 mL of sterile glycerol (sterilize by autoclaving for 20 min at 1 atm on liquid cycle). Vortex the culture to ensure that the glycerol is evenly dispersed. Transfer the culture to a labelled storage tube, freeze the culture in ethanol-dry ice or in liquid nitrogen, and then transfer the tube to −70°C for long-term storage.
3. We recommend a DNA polymerase such as Phusion® High-Fidelity DNA Polymerase (Finnzymes) to minimize the occurrence of mutations during PCR.

4. When using enzymes that exhibit star activity, use the minimum reaction time required for complete digestion given that prolonged incubation may result in increased star activity.
5. To prevent re-circularization of the vector in cloning strategies, the DNA needs to be dephosphorylated by treatment with alkaline phosphatase.
6. Alternatively, screen for recombinant plasmid clones by using PCR to directly amplify the insert from each bacterial recombinant colony.
7. The pUCRPLACZ expression vector DNA can be kept for several years in TE buffer at −20°C and the corresponding bacterial clone can be preserved as a 15% glycerol stock at −80°C for long-term storage.
8. It is necessary that the starter culture is grown until to logarithmic phase given that the recombinant construct shows lower expression if the starter culture reached the stationary phase.
9. Add DTT fresh immediately before use because it tends to undergo oxidation in the presence of air.
10. Add PMSF to the lysis buffer to prevent protein degradation.
11. Avoid repeated freeze/thaw cycles of the protein extracts.
12. The purity of the reagents and the cleanliness of the plasticware is very critical.
13. Ensure that the rate of the reaction is directly proportional to the amount of enzyme, i.e., that the substrate is in excess and that the activities obtained are within the linear range of the enzyme under the conditions employed. Miller reported that β-galactosidase activity is within the linear range as long as the absorbance at 410 nm produced by enzymatic hydrolysis of ONPG remains below 0.6–0.9. Thus, if assays produce activity values outside the linear range, they should be repeated using a smaller amount of protein extract or a shorter reaction time.
14. β-Mercaptoethanol is added to the reaction buffer to stabilize the β-galactosidase enzyme. The important part of β-mercaptoethanol is a reactive thiol (SH group). Thiols react with oxygen in the air and oxidize (inactivate) over time. Therefore, try not to make more buffer than you will use in a few days. Store the unused portion at 4°C.
15. The ONPG substrate should be dissolved fresh each time.
16. The amount of L-malate for induction of protein expression using the pUCRP vector depends on the protein that is to be expressed.

References

1. Georgiou, G., and Valax, P. (1996) Expression of correctly folded proteins in *Escherichia coli. Curr. Opin. Biotechnol.* **7**, 190–197.
2. Mitra, A., Chakrabarti, K. S., Shahul Hameed, M. S., et al. (2005) High level expression of peptides and proteins using cytochrome b5 as a fusion host. *Protein Expr. Purif.* **41**, 84–97.
3. Luo, Z. H., and Hua, Z. C. (1998) Increased solubility of glutathione S-transferase-P16 (GST-p16) fusion protein by co-expression of chaperones GroES and GroEL in *Escherichia coli. Biochem. Mol. Biol. Int.* **46**, 471–477.
4. Jeon, Y. H., Negishi, T., Shirakawa, M., et al. (1995) Solution structure of the activator contact domain of the RNA polymerase alpha subunit. *Science* **270**, 1495–1497.
5. Atlas, R. M., and Bartha, R. (1993) Microbial ecology: fundamentals and applications, 3rd ed. Benjamin/Cummings, Redwood City California.
6. Medigue, C., Krin, E., Pascal, G., et al. (2006) Coping with cold: the genome of the versatile marine Antarctica bacterium *Pseudoalteromonas haloplanktis* TAC125. *Genome Res.* **15**, 1325–1335.
7. Tutino, M. L., Duilio, A., Fontanella. B., et al. (1999) Cold-adapted organisms: ecology, physiology, enzymology and molecular biology. Springer, Berlin.
8. Duilio, A., Tutino, M. L, and Marino, G. (2004) Recombinant protein production in Antarctic Gram-negative bacteria. *Methods Mol Biol.* **267**, 225–237.
9. Blatny, J. M., Brautaset, T., Winther-Larsen, et al. (1997) Construction and use of a versatile set of broad-host-range cloning and expression vectors based on the RK2 replicon. *Appl Environ Microbiol.* **63**, 370–379.
10. Tutino, M. L., Duilio, A., Parrilli, E., et al. (2001) A novel replication element from an Antarctic plasmid as a tool for the expression of proteins at low temperature. *Extremophiles* **5,** 257–264.
11. Duilio, A., Madonna. S., Tutino, M. L., et al. (2004) Promoters from a cold-adapted bacterium: definition of a consensus motif and molecular characterization of UP regulative elements. *Extremophiles* **8**, 125–132.
12. Papa, R., Glagla, S., Danchin, A., et al. (2006) Identification of a two-component regulatory system in *Pseudoalteromonas haloplanktis* TAC125 by using a proteomic approach. *Extremophiles* **10**, 483–491.
13. Papa, R., Rippa, V., Sannia, G., et al. (2007) An effective cold inducible expression system developed in *Pseudoalteromonas haloplanktis* TAC125. *J. Biotechnol.* **127**, 199–210.
14. Hoyoux, A., Jennes, I., Dubois, P., et al. (2001) Cold-adapted beta-galactosidase from the Antarctic psychrophile Pseudoalteromonas haloplanktis. Appl. Environ. Microbiol. **67**, 1529–1535.
15. Tascon, R. I., Rodriguez-Ferri, E. F., Gutierrez-Martin, C. B., Rodriguez-Barbosa, I., Berche, P., and Vazquez-Boland, J. A. (1993) Transposon mutagenesis in Actinobacillus pleuropneumoniae with a Tn10 derivative. J. Bacteriol. **175**, 5717–5722.
16. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
17. Tutino, M. L., Duilio, A., Moretti, M. A., Sannia, G., and Marino, G. (2000) A rolling-circle plasmid from Psychrobacter sp. TA144: evidence for a novel Rep subfamily. Biochem. Biophys. Res. Commun. **274**, 488–495.
18. Bradford, M. M. (1976) A rapid and sensitive method for the quantisation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**, 248–54.
19. Tutino, M. L., Parrilli, E., Giaquinto, L., Duilio, A., et al. (2002) Secretion of α-amylase from *Pseudoalteromonas haloplanktis* TAB23: two different pathways in different hosts. *J. Bacteriology* **184**, 5814–5817.

Chapter 11

A Novel Strategy for the Construction of Genomic Mutants of the Antarctic Bacterium *Pseudoalteromonas haloplanktis* TAC125

Maria Giuliani, Ermenegilda Parrilli, Cinzia Pezzella, Valentina Rippa, Angela Duilio, Gennaro Marino, and Maria Luisa Tutino

Abstract

The sequencing and the annotation of the marine Antarctic *Pseudoalteromonas haloplanktis* TAC125 genome has paved the way to investigate on the molecular mechanisms involved in adaptation to cold conditions. The growing interest in this unique bacterium prompted the developing of several genetic tools for studying it at the molecular level. To allow a deeper understanding of the *Ph*TAC125 physiology a genetic system for the reverse genetics in this bacterium was developed. In the present work, we describe a practical technique for allelic exchange and/or gene inactivation by in-frame deletion and the use of a counterselectable marker in *P. haloplanktis*. The construction of suitable non-replicating plasmid and methods used to carry out a two-step integration–segregation strategy in this bacterium are reported in detail.

Furthermore two examples, in which the developed methodology was applied to find out gene function or to construct genetically engineered bacterial strains, were described.

Key words: *Pseudoalteromonas haloplanktis* TAC125, Allelic exchange, Counterselectable marker, Suicide vector

1. Introduction

A significant fraction of life develops in the sea at temperatures lower than 15°C. However, little is known about the adaptive features selected under those conditions.

The fast-growing psychrophilic bacterium *Pseudoalteromonas haloplanktis* TAC125 (*Ph*TAC125) is the first Antarctic bacterium whose genome has been sequenced and carefully annotated (1). Genomic and metabolic features of this γ-proteobacterium, accounting

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_11, © Springer Science+Business Media, LLC 2012

for its remarkable versatility and fast growth compared with other bacteria from aqueous environments, were discovered by combining genome sequencing and further in silico and in vivo analyses (1). *Ph*TAC125 has found much unexpected solutions to cope with life in the cold (1), not only it grows fast under such conditions, but it displays remarkable resistance to reactive oxygen species (ROS). Moreover, it provides a way to resist to the ageing caused by asparagine cyclization and deamidation. Indeed the *Ph*TAC125 proteome reveals a concerted amino acid usage bias specific to psychrophiles, making it an organism of choice for foreign protein production when deamidation ought to be put to a minimum (1).

Fast growth rates, combined with the ability of *Ph*TAC125 to reach very high cell densities and to be easily transformed by intergeneric conjugation (2), made this bacterium not only a model for the study of adaptation to cold conditions but also an attracting host for the development of an efficient gene-expression system at low temperatures.

The growing interest in this unique bacterium for potential applications prompted the developing up of several genetic tools to allow the recombinant protein production in *Ph*TAC125 (3), although it is necessary to continue to expand the repertoire of useful genetic tools for studying this bacterium at the molecular level. Indeed, if a deeper understanding of the *Ph*TAC125 physiology is to be gained, a genetic system for the reverse genetics in this bacterium must be developed.

Reverse genetics is a powerful approach for the identification of gene function, in which the gene of interest is mutated or inactivated to study the resulting effects on the microorganism (4). The setting up of methods for the reverse genetic approaches in *Ph*TAC125 is essential to determine gene function, and to create genetically engineered bacterial strains, possibly displaying an improved feature as host for recombinant protein production.

The replacement of a wild type chromosomal allele of a gene of interest with a cloned allele that has been manipulated in vitro (termed allelic exchange) is one essential genetic technique, indeed for some experiments it may be desirable to replace the wild type allele with a mutant allele rather than completely knockout a gene; thus, a means of allelic exchange rather than insertion inactivation is important.

This technique can be used to transfer various types of mutant alleles of a gene of interest, including, for example, in-frame deletions and point mutations, to the wild type chromosomal locus. One advantage of allelic exchange is the placement of a mutant allele in its natural chromosomal context, an optimal situation for genetic analysis of the effects of the mutation. Furthermore, if markerless exchange is used, multiple mutations can be serially introduced into the same strain without the addition of distinct antibiotic resistance markers that might not be available or might otherwise interfere with subsequent genetic analysis.

Allelic exchange is typically performed by using a two-step integration–segregation strategy, in which a non-replicating plasmid harbouring a cloned copy of the gene with the desired mutation is first integrated into the chromosome via Campbell-type homologous recombination (5) with selection for a plasmid-encoded marker. After removal of the selective condition, the resulting merodiploid recombinants can resolve the duplication via a second recombination event between the duplicated chromosomal segments, yielding a recombinant carrying either the wild type allele or the mutated allele, depending on the site of recombination.

Although allelic exchange is easy to perform with many bacteria, it remains very difficult or impractical with others. The classical method of using a suicide plasmid that is unable to replicate in the studied strain to deliver an inactivated allele of the gene in the chromosome is often not efficient because the frequency of double crossover events may be low and because illegitimate recombination may occur (6). Consequently, allelic exchange mutants may represent only a small fraction of the transformants and may be difficult to isolate. The use of counterselectable marker (7) are often instrumental for the construction of such mutants, especially in microorganisms for which the genetic scheme is poorly developed. Under appropriate growth conditions, a counterselectable gene promotes the death of the microorganisms harbouring it (7). Hence, transformants which have integrated a suicide vector containing a counterselectable marker, either by a single event of homologous or illegitimate recombination, retain a copy of the counterselectable marker in the chromosome and are therefore eliminated in the presence of the counterselective compound. Consequently, counterselectable markers have been used for to facilitate the identification of recombinants that have lost the integrated plasmid. This strategy permits selection to be used for isolation of recombinants that have resolved the merodiploid state.

The evaluation of possible counterselectable markers that can be successfully used in *Ph*TAC125 has been recently reported (8, 9). Although several counterselectable markers were tested, the only gene able to allow a counterselection against the Antarctic bacterium was *pheS*^{Gly294} mutated gene. The *pheS* gene encodes the α subunit of phenylalanyl-t-RNA synthetase (10). Mutation of G294A in this protein results in an enzyme with relaxed substrate specificity that tolerates incorporation of the toxic phenylalanine analogue *p*-chlorophenylalanine. Expression, therefore, of a mutated *pheS* gene in the presence of *p*-chlorophenylalanine is toxic to the host cell (10). This strategy has successfully been used for counterselection in Gram-negative (10, 11) and Gram-positive bacteria (12).

In this chapter we describe a practical technique for allelic exchange and/or gene inactivation by in-frame deletion and the use of *pheS*^{Gly294} mutated gene as counterselectable marker in *Ph*TAC125. The construction of suitable non-replicating plasmid (suicide vector) and

methods used to perform two-step integration–segregation strategy in the Antarctic bacterium are reported in detail. Moreover several examples, in which the developed methodology was applied to determine gene function or to create genetically engineered bacterial strains, are described.

2. Materials

2.1. Bacterial Strains

1. *Pseudoalteromonas haloplanktis* TAC125 (see Notes 1 and 2).
2. *Escherichia coli* DH5α [*supE*44, Δ*lac*U169 (ϕ80 *lacZ*ΔM15) *hsdR*17, *recA*1, *endA*1, *gyrA*96, *thi*-1, *relA*1] was used as host for the gene cloning.
3. *E. coli* strain S17-1(λ*pir*) [*thi*, *pro*, *hsd* (r^- m^+) *recA*::RP4-2-TC^r::Mu Km^r::Tn7 Tp^r Sm^r λ*pir*] was used as donor in intergeneric conjugation experiments (13).

2.2. Liquid Media

1. 100 mg/mL Ampicillin stock solution: Dissolve 1 g of ampicillin powder in deionized H_2O (dH_2O) and sterilize by filtration through a 0.22-μm sterile filter. Split the obtained stock solution in aliquots of 1 mL each in sterile polypropylene tubes and store them at −20°C.
2. 30 mg/mL Carbenicillin stock solution: Dissolve 300 mg of carbenicillin in 10 mL dH_2O and sterilize by filtration through a 0.22-μm sterile filter. Split the obtained stock solution in aliquots of 1 mL each in sterile polypropylene tubes and store them at −20°C.
3. 40% (w/v) D-galactose stock solution: Dissolve 20 g of D-galactose in 40 mL of dH_2O. After the powder has dissolved adjust the volume of the solution to 50 mL with dH_2O and sterilize by filtration through a 0.22-μm sterile filter. Store at 4°C up to 2 months.
4. 100× KH_2PO_4 stock solution: Dissolve 10 g of KH_2PO_4 in 90 mL of deionized water. Adjust the pH to 7 with NaOH and add dH_2O up to 100 mL. Sterilize by filtration through a 0.22-μm sterile filter. Store at 4°C up to 6 months.
5. 300× $MgSO_4$ stock solution: Dissolve 6 g of $MgSO_4{\cdot}7H_2O$ in 90 mL of deionized water. Adjust the pH to 7 with NaOH and add dH_2O up to 100 mL. Sterilize by filtration through a 0.22-μm sterile filter. Store at 4°C up to 6 months.
6. 100× $FeSO_4$ stock solution: Dissolve 183 mg of $FeSO_4{\cdot}7H_2O$ in 90 mL of deionized water. Adjust the pH to 7 with NaOH and add dH_2O up to 100 mL. Sterilize by filtration through a 0.22-μm sterile filter. Store at 4°C up to 1 year.

7. 3,000× $CaCl_2$ stock solution: Dissolve 1.5 g of $CaCl_2 \cdot 2H_2O$ in 90 mL of deionized water. Adjust the pH to 7 with NaOH and add dH_2O up to 100 mL. Sterilize by filtration through a 0.22-μm sterile filter. Store at 4°C up to 6 months.
8. 50× TAE stock solution: Mix 242 g of Tris base, 57.1 mL of glacial acetic acid, and 100 mL of 0.5 M EDTA pH 8 and adjust the volume of the solution to 1 L with dH_2O. Store at room temperature (RT) up to 1 year.
9. 0.5 M EDTA pH 8: Dissolve 186.1 g of EDTA in 800 mL of deionized H_2O. Adjust the pH to 8 with NaOH (about 20 g of NaOH pellets) and adjust the volume of the solution to 1 L with dH_2O.
10. 6×/Agarose gel-loading buffer: 0.25% bromophenol blue, 0.25% xylene cyanol FF, and 30% glycerol (Fermentas).
11. 10 mg/mL Ethidium bromide stock solution: Add 1 g of ethidium bromide to 100 mL of dH_2O. Stir on a magnetic stirrer for several hours to ensure that the dye has dissolved. Wrap the container in aluminium foil and store it at 4°C.
12. Luria-Bertani (LB) media: 10 g Bacto-tryptone, 5 g yeast extract, and 10 g NaCl. Dissolve in 900 mL dH_2O and shake until the solutes have dissolved. Adjust the volume of the solution to 1 L with dH_2O. Sterilize by autoclaving for 20 min at 1 atm on liquid cycle. Let it cool down and store at RT. When required, add 1 mL of sterile ampicillin stock solution (see Note 3). To prepare solid medium, add 15 g/L bacto-agar just before autoclaving.
13. TYP medium (14): 16 g Bacto-tryptone, 16 g yeast extract and 10 g marine mix. Dissolve in 900 mL dH_2O and shake until the solutes have dissolved. Adjust the pH to 7.5 with 5 N NaOH and adjust the volume to 1 L with dH_2O. Sterilize by autoclaving for 20 min at 1 atm on liquid cycle. When required, add 1 mL of sterile carbenicillin stock solution (see Note 3). To prepare solid medium, add 15 g/L bacto-agar just before autoclaving.
14. Counterselection medium (15): 1 g NH_4NO_3, 10 g NaCl, 1.6 g L-alanine, 1.9 g, L-glutamate, 1.2 g L-leucin. Only when required add also 2 g p-chlorophenylalanine (see Note 4). Dissolve in 900 mL dH_2O and shake until the solutes have dissolved and sterilize by autoclaving for 20 min at 1 atm on liquid cycle. Allow the solution to cool down at 50°C or less, and then add: 10 mL of 100× KH_2PO_4, 3.34 mL of 300× $MgSO_4$, 10 mL of 100× $FeSO_4$, 0.33 mL 3,000× $CaCl_2$ (see Note 5) and 12.5 mL of 50× D-galactose from sterile solution stock solutions. Store at 4°C up to 3 months. When required, add 1 mL of sterile carbenicillin stock solution (see Notes 3 and 9). To prepare solid medium, add 15 g/L bacto-agar just before autoclaving.

2.3. Reagents for Molecular Biology

1. Phusion™ DNA Polymerase (Finnzymes).
2. Go-Taq DNA Polymerase (Promega).
3. Restriction enzymes (Fermentas).
4. Calf Intestinal Alkaline Phosphatase, CIAP (Fermentas).
5. T4 DNA Ligase (Promega).
6. ChargeSwitch® gDNA Mini Bacteria Kit (Invitrogen).
7. High Pure PCR Product Purification Kit (Roche).
8. QIAprep® Spin Miniprep Kit (Qiagen).
9. QIAquick® Nucleotide Removal Kit (Qiagen).
10. pGEM®-T Easy Vector System I (Promega).

3. Methods

3.1. Construction of the *P. haloplanktis* TAC125 Suicide Vector

A two-step integration–segregation strategy is achieved by means of a non-replicating plasmid (suicide vector, pVS), suitably constructed for *P. haloplanktis* TAC125. The pVS vector is characterized by the presence of: (1) the pJB3-derived *oriT* (3), a DNA fragment responsible for the initiation of the conjugative transfer between an *E. coli* S17-1 λ*pir* strain (donor) and the psychrophilic cells (acceptor); (2) a pGEM7Z derived polylinker, wherein a mutated portion of the target gene of interest (*goi*) can be cloned; (3) the *E. coli blaM* gene, encoding a mesophilic β-lactamase, which is used as selection gene to isolate the first site-specific integration event; (4) *pheS*^{Gly294}, which encodes a mutated version of the *E. coli* α-subunit of Phe-tRNA synthase (10), which renders bacteria sensitive to *p*-chlorophenylalanine. This phenylalanine analogue is used as counterselective agent for the isolation of those strains in which a second recombination event occurred. To assure a proper level of *pheS*^{Gly294} expression, its transcription was subjected to the control of a psychrophilic synthetic promoter (P13) (3).

1. The *oriT*, responsible for the initiation of the conjugative transfer, is amplified on pJB3 vector by using the primers oriTEcoRIfw (5′-TTGAATTCTCGCACGATATACAGG-3′) and oriTSacIrv (5′-AAGAGCTCTTGAAGACGAAAGGG-3′) containing the proper overhangs of *Eco*RI and *Sac*I restriction sites (underlined above). PCR reaction is carried out by Phusion® High-Fidelity DNA Polymerase following the manufacturer's protocols (see Note 6). The amplified product is purified using a commercial purification kit and then digested with *Eco*RI and *Sac*I. The same enzymes are used to linearize pGEM7Z. Restriction hydrolyses are performed by using 5 enzyme U/μg of DNA, in the reaction conditions defined by the manufacturer (see Note 7). Afterwards, the cleaved vector is dephosphorylated with calf intestinal

alkaline phosphatase, CIAP (0.5 U/pmol of 5′ phosphate ends) for 15 min at 37°C and 45 min at 55°C (see Note 8). The CIP is heat-inactivated at 75°C for 10 min and the entire reaction volume is then loaded on a 1% agarose gel (containing ethidium bromide as fluorescent marker for the migrating DNA). The linearized vector is cut out of the gel and purified using a commercial gel-purification kit following the manufacturer's instructions.

The cleaved dephosphorylated vector is then ligated to the digested *oriT* fragment by the means of T4 DNA ligase and finally transformed into chemically competent bacteria (*E. coli* strain *Top10*) according to the procedure described by Duilio (2). Recombinant clones are selected on LB-agar plates containing ampicillin (final concentration 100 μg/mL) as selective marker. Plasmids are isolated from ampicillin-resistant clones and sequences are checked by sequencing.

2. The *pheS*^{Gly294} gene, which encodes a mutated version of the *E. coli* α subunit of Phe-tRNA synthase, is amplified using pKSS (10) vector as template. The PCR reaction is carried out using the oligonucleotide pair PheSSNfw (5′-TTGTCGACATATGTCACATCTCGCAGAAC-3′) and PheSXrv (5′-CCTCTAGAGAATTTCATAATCTATTCCTGCC-3′), designed to introduce *Nde*I and *Xba*I restriction sites (underlined above).

   In order to assure a proper level of *pheSG*^{ly294} expression in the psychrophilic host, its transcription is subjected to the control of a psychrophilic synthetic promoter (P13). Thus, the amplified DNA fragment is subjected to double *NdeI/XbaI* digestion and cloned into pPM13 plasmid (3) corresponding sites generating pPM13-*pheSG*^{ly294} vector. The DNA fragment, containing P13 promoter and *pheSG*^{ly294} gene, is recovered from pPM13-*pheSG*^{ly294} vector by *SmaI/EcoR*V digestion, and cloned into the pGEM7Z-*oriT NaeI* restriction site, resulting in the construction of the pVS vector. Plasmids are isolated from ampicillin-resistant clones and sequences are checked by sequencing. The molecular biology procedures are similar to those described in step 1.

3. Two *goi* fragments, a and b, are amplified by PCR on bacterial genomic DNA as template, by the means of specific oligonucleotide pairs: A–B and C–D respectively (Fig. 1A). The primers are designed to introduce an internal restriction site (in B and C) to join the fragments, together with the restriction sites (in A and D) for their cloning in pVS polylinker. The amplified products correspond to two internal gene fragments (300–600 bp) that are sequential but not adjacent in the gene sequence (being at least 200 bp apart), in order to provide a deletion in the *goi*. The translation frame-shift of the resulting deleted gene must

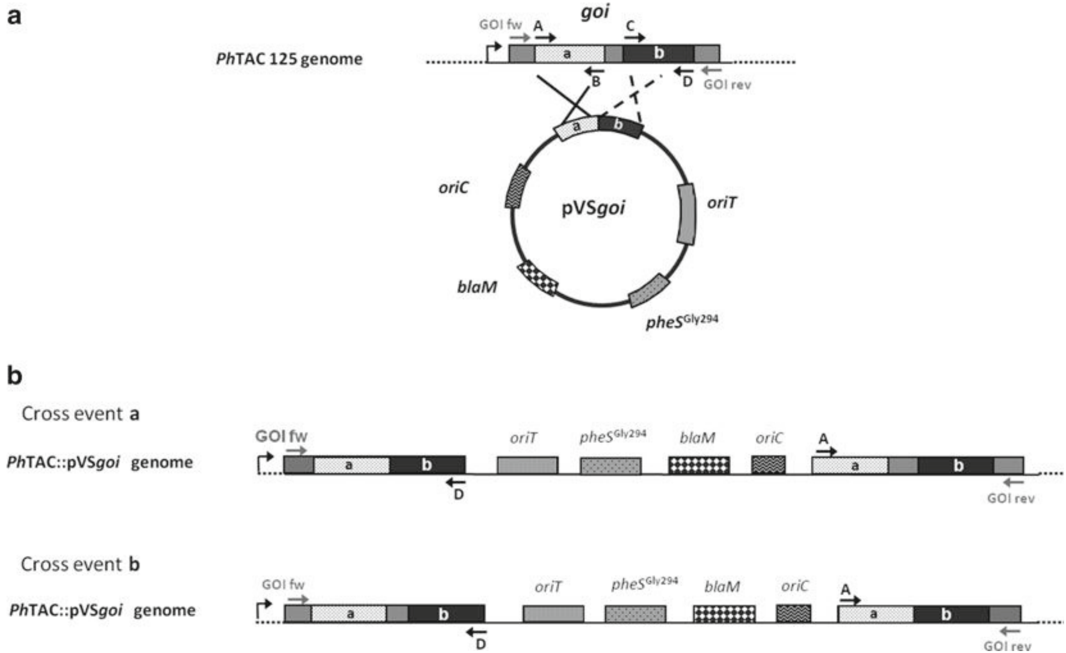


Fig. 1. Schematic representation of the first recombination event. Since the pVSgoi contains two goi fragments (a and b), two different crossing over events may occur, depending on which fragment undergoes recombination (**A**) resulting in two different orientations of pVSgoi insertion on the genome. (**B**) Whatever crossing-over event occurs, selected clones will contain two mutated copies of goi, both carrying a specific deletion, resulting in a *P. haloplanktis* TAC125::VSgoi mutant.

be carefully checked, in order to avoid in-frame translation of a mutated gene product.

4. The fragments are suitably digested and cloned in the pVS corresponding sites, generating pVS*goi*. Amplification and cloning reactions are carried out as described in step 1.

3.2. Construction of P. haloplanktis TAC125 Genomic Mutants

3.2.1. First Recombination Event: Integration of the pVSgoi in the Genome

1. The resulting vector (pVS*goi*) is mobilized into *P. haloplanktis* TAC125 by intergeneric conjugation (2). Cells are plated on TYP solid medium containing 30 μg/mL carbenicillin (see Note 9) and incubated at 4°C to select those clones in which a single recombination event – between a *goi* fragment on the pVS*goi* and the corresponding homologous sequence on *goi* genomic copy – has occurred.
2. The first recombination event allows the pVS*goi* insertion on the genome. Since the pVS*goi* contains two *goi* fragments (a and b), two different crossing over events may occur, depending on which fragment undergoes recombination (Fig. 1B). In order to characterize the occurred genomic insertion, genomic DNA from some carbenicillin-resistant clones is extracted and subjected to different PCR analyses (Fig. 1B). Amplification of *pheS* ^{Gly294} gene by PheSSNfw and PheSXrv primers is used to confirm plasmid insertion. Positive clones are further analyzed by PCR to distinguish between the occurred recombination

events. Two oligonucleotide sequences designed on the upstream (GOIfw) and downstream (GOIrev) *goi* sequences, respectively, are used in combination with primers A and D to characterize the genome organization of the mutants. Depending on the crossing over event (a or b), amplification with primer pairs GOIfw-D or GOIrev-A will give fragments of different sizes (Fig. 1B). PCR amplifications are performed in standard conditions, using Taq polymerase.

Whatever crossing-over event occurs, selected clones will contain two mutated copies of *goi* resulting in a *P. haloplanktis* TAC125::VS *goi* mutant.

3.2.2. Selection of the Second Recombination Event

For some applications, such as the construction of engineered strains lacking a single gene function or entire metabolic pathways, an integration mutation is not sufficient to guarantee the genetic stability of the mutant strain. Indeed, in the absence of selective condition, the merodiploid recombinants can resolve the duplication via a second recombination event between the duplicated chromosomal segments, yielding a strain carrying either the wild type allele or the mutated allele, depending on the site of recombination. The following protocol allows the specific selection of knockout mutants by using the counterselectable marker, present in the chromosome of transformants that have integrated the suicide vector.

1. Inoculate a single colony of the *P. haloplanktis* TAC125::VS*goi* mutant strain in 3 mL of liquid TYP medium.
2. Incubate at 15°C while shaking (250 rpm) for 18–24 h, until the cell density reaches OD_{600} 3–5 (see Note 10).
3. Perform serial dilution of the culture (1:10, 1:100, 1:1,000 v/v) in 100 μL of counterselection liquid medium in sterile polypropylene tubes.
4. Plate the dilutions on solid counterselection medium containing *p*-chlorophenylalanine (hereafter called counterselection plates).
5. Incubate the plates at 15°C until colonies development (10–15 days). Growth on plates containing *p*-chlorophenylalanine (*p*-Cl-Phe) indicates the occurrence of the second recombination event, which causes the excision of the suicide vector from the target *locus* of the bacterial genome. Indeed, insertion mutants carry the *phe*S^{Gly294} gene, which confers to the cells sensitivity to the phenylalanine toxic analogue, *p*-Cl-Phe. The counterselection medium maximizes the recombinant expression of mutated phenylalanine tRNA synthetase and therefore it does not allow the growth of insertion mutants.

3.2.3. Isolation of *P. haloplanktis* TAC125 Deletion Mutants

Clones able to grow on counterselection medium supplemented with *p*-Cl-Phe are bacteria in which the counterselectable marker (the *phe*S^{Gly294} gene) is not expressed. Apart from those clones in which the second recombination event has occurred and consequently

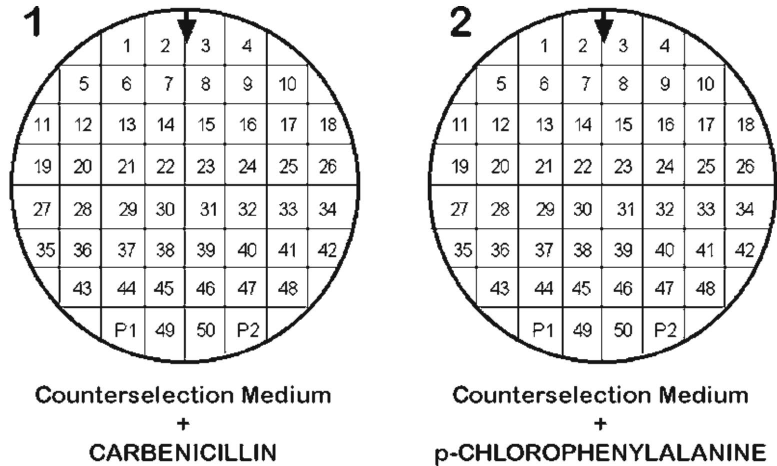


Fig. 2. Multi-gridded templates for replica plating.

lack the suicide vector insertion in the genome, other clones could have developed different mutations conferring the ability to overcome the toxicity of the counterselective compound despite the presence of *phe*S^{Gly294} gene in the chromosome. For this reason, a second screening is required for the selection of clones in which the second crossing over is occurred. It will be carried out by a replica plating strategy where a single clone is plated in the presence of carbenicillin (where only the cells carrying the suicide vector can grow) and replicated in the presence of *p*-Cl-Phe (where only cells lacking the suicide vector can grow).

1. Prepare two 150-mm multi-gridded plates with counterselection solid medium, one supplemented with carbenicillin and the other with *p*-Cl-Phe. A grid needn't be drawn onto each plate as long as the top is marked with an arrow for alignment purposes. Use a piece of card with Petri plate covers attached and copies of the grid taped inside (Fig. 2). Label the back of each plate with an alignment arrow, date and medium type.
2. Pick a single colony from one of the counterselection plates (choose the plate where the number of colonies allows the easiest selection of single clones) with a sterile needle and streak it in the grid number "1" first on the plate containing carbenicillin and then on the plate containing *p*-Cl-Phe. Repeat it for at least 50 colonies filling all the empty grids on the plates. The squares labelled P1 and P2 are for the wild type strain and the insertion mutant respectively, a control to check if they are behaving as expected, and eventually to aid in phenotype analysis.
3. Incubate the plates at 15°C for 5–10 days. Clones able to grow on medium containing *p*-Cl-Phe and NOT in the presence of carbenicillin are selected for the next molecular characterization.

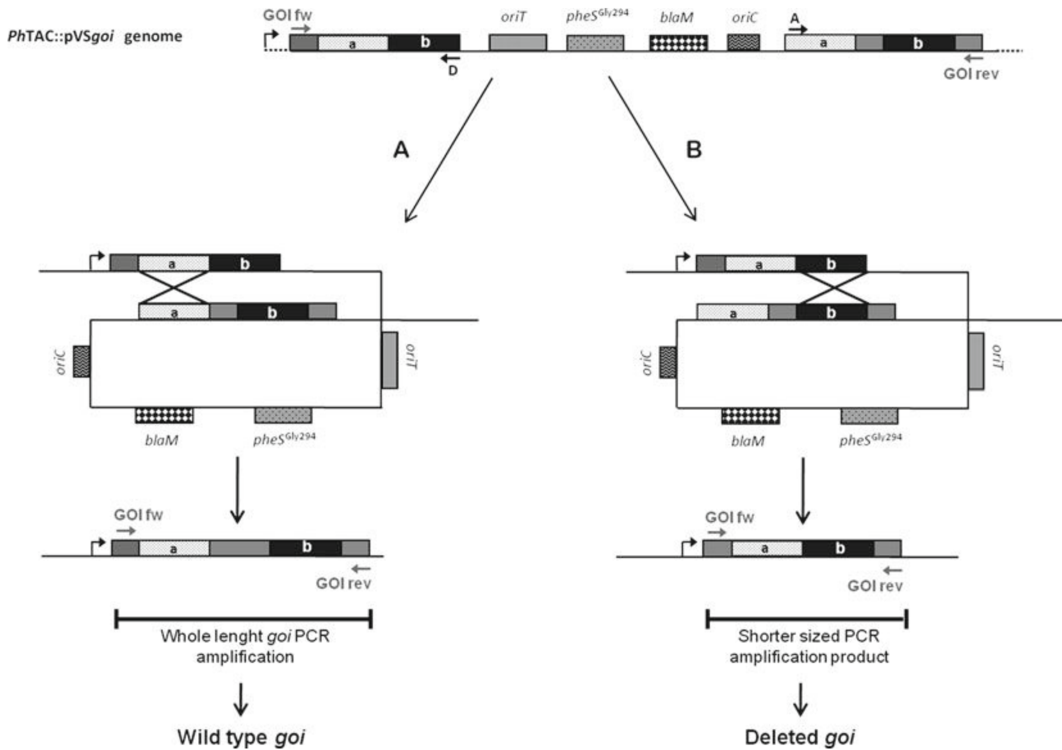


Fig. 3. Schematic representation of second recombination event. Depending on the site of recombination, the intramolecular crossing-over can yield either the wild type gene (**A**) or the deleted form of goi (**B**).

If none of the clones show the desired phenotype repeat the screening analyzing more colonies.

4. Inoculate the selected clones in 3 mL of liquid TYP medium.
5. Incubate at 15°C while shaking (250 rpm) for about 24 h.
6. Centrifuge 1.5 mL of each culture 10 min at 3,300 × *g* at 4°C, remove the supernatant and use the resulting pellets for genomic DNA extraction. The remaining culture can be used for strain storage (see Note 2).
7. Use 100 ng of genomic DNA for *pheS*^{Gly294} gene amplification by using PheSSNfw and PheSXrv primers. Use the same amount of genomic DNA extracted from *P. haloplanktis* TAC125::VS*goi* strain as positive control and from *P. haloplanktis* TAC125 wild type as negative control. Clones in which the suicide vector has excised from the chromosome will NOT show *pheS*^{Gly294} gene amplification.

3.2.4. Molecular Characterization of P. haloplanktis Deletion Mutants

The second recombination event causes the excision of the suicide vector from the target *locus* in the bacterial genome. Depending on the site of recombination (Fig. 3), the intramolecular crossing over can yield either the wild type gene (Fig. 3A) or the deleted form of *goi* (Fig. 3B). If the mutation does not confer a disadvantageous

phenotype the probability of merodiploids to revert to wild type or to mutant is of 50% (but it can be much different if the mutation confers an unfavorable phenotype). Hence, apart from a phenotypic analysis, a molecular characterization is required in order to isolate the clones in which the desired deletion has occurred.

1. Use 100 ng genomic DNA of each clone selected by the previous screening as template for a PCR reaction using the oligonucleotide pair GOIfw-GOIrv. Use the same amount of genomic DNA extracted from *P. haloplanktis* TAC125 wild type strain as negative control. Perform PCR amplifications in standard conditions, using Taq polymerase. The PCR amplification will give a fragment of the same size of that obtained for the negative control (whole *goi* length) if the clone has resolved the merodiploid state reverting to wild type, while a shorter sized fragment is expected for the deleted copy of *goi* (Fig. 3A, B). PCR products can be directly cloned using the pGEM®-T Easy system and sequenced.

3.3. Characterization of the Role of a 2-on-2 Haemoglobin in Oxidative and Nitrosative Stress Resistance of *P. haloplanktis* TAC125

The 2-on-2 haemoglobins, previously named truncated, are monomeric, low-molecular weight oxygen binding proteins that share the overall topology with vertebrate haemoglobins (16). Although several studies on 2-on-2 haemoglobins have been reported, their physiological and biochemical functions are not yet well defined, and various roles have been suggested. The genome of *P. haloplanktis* TAC125 is endowed with three genes encoding 2-on-2 haemoglobins (17). To investigate the function played by one of the three trHbs, *Ph*HbO, a *P. haloplanktis* TAC125 genomic mutant strain was constructed, in which the encoding gene was inactivated by an integration strategy (18). For the construction of a suitable suicide vector, a genomic fragment corresponding to a 241-bp long region of the *PHSAa0030* gene was PCR amplified by using specific oligonucleotides as primers (18).

The mutant strain was grown under controlled conditions and several aspects of bacterium physiology were compared with those of wild type cells when dissolved oxygen pressure in solution and growth temperature were changed. Interestingly, inactivation of the *Ph*HbO encoding gene makes the mutant bacterial strain sensitive to high solution oxygen pressure, to H_2O_2, and to a nitrosating agent, suggesting the involvement of *Ph*HbO in oxidative and nitrosative stress resistance (18).

3.4. Cell Engineering of *P. haloplanktis* TAC125: Construction of a Mutant Strain with Reduced Exo-Proteolytic Activity

A "cold" gene-expression system implemented for the secretion of recombinant proteins in *P. haloplanktis* TAC125 was previously set up (19). Such a system could effectively conjugate the positive effect of low temperature on the recombinant product solubility with the obvious advantages linked to extra-cellular protein targeting. This novel system makes use of the psychrophilic α-amylase

from *P. haloplanktis* TAB23 (3) as secretion carrier. Several chimerical proteins were produced and used to test the versatility and efficiency of the novel secretion system. All the chimerical proteins were efficiently produced and secreted (19).

However, bacteria belonging to *Pseudoalteromonas* genus are reported to secrete a wide range of exo-proteins, especially proteases. This feature could hamper both applicability and efficiency of the cold-adapted secretion system, due to the possible recombinant product degradation. The *P. haloplanktis* TAC125 genome in silico analysis highlighted the presence of a putative Type II secretion system (T2SS), while the extra-cellular targeting of the cold α-amylase depends on a novel and still uncharacterized secretion pathway (15). A simple and successful experimental approach was designed for the construction of an improved psychrophilic host for the cold α-amylase-dependent recombinant secretion system through the development of a *P. haloplanktis* TAC125 mutant strain where the T2SS was inactivated. For the construction of a suitable suicide vector, two genomic fragments were PCR amplified corresponding to the 5′ 360-bp portion of *gsp*C and 3′ 300-bp portion of *gsp*N respectively. The integration–segregation strategy led to the construction of a *P. haloplanktis* TAC125 where the T2SS cluster was almost totally deleted (8). While the growth behaviour and some physiological features of the mutant strain are indistinguishable from the wild type ones, the deleted strain displays a remarkable reduction in the protease content in the culture supernatant. This aspect makes it a truly improved host with a remarkable biotechnological potential in recombinant protein secretion at low temperature. Moreover, this work demonstrates that *P. haloplanktis* TAC125 is a versatile psychrophilic host for recombinant protein production since it that can be easily improved by a directed engineering approach. To the best of our knowledge this is the first described example of a strain improvement strategy applied to an Antarctic bacterium.

4. Notes

1. *Pseudoalteromonas haloplanktis* TAC125 was kindly provided by C. Gerday, University of Liege, Belgium. The strain was isolated from the sea water in the surrounding of the Dumont d'Urville Antarctic station (66°40′S, 40°01′E) during the 1988 summer campaign of the "Expeditions Polaires Française" in Terre Adélie.
2. Bacteria can be stored indefinitely in cultures containing glycerol at low temperature (either –20 to –70°C). To store bacteria, pick a single, well-isolated colony growing on the surface

of an agar plate, with a sterile inoculating needle, dip the needle into the liquid broth and grow the bacterial culture up to medium-late exponential phase. To 0.85 mL of bacterial culture, add 0.15 mL of sterile glycerol (sterilize by autoclaving for 20 min at 1 atm on liquid cycle). Vortex the culture to ensure that the glycerol is evenly dispersed. Transfer the culture to a labelled storage tube, freeze the culture in ethanol-dry ice or in liquid nitrogen, and then transfer the tube to –70°C for long-term storage.

3. After antibiotic addition the medium can be stored at 4°C up to 2 weeks.
4. The *p*-chlorophenylalanine is poorly soluble at the required concentration. Heat the solution up to 50°C while shaking to completely solubilize it.
5. Most of the salts contained in counterselection medium precipitate after autoclaving and therefore they must be prepared separately in higher concentrations, sterilized by filtration and finally added to the basal medium.
6. We recommend a DNA polymerase such as Phusion® High-Fidelity DNA Polymerase (Finnzymes) to minimize the occurrence of mutations during PCR.
7. When using enzymes that exhibit *star activity*, use the minimum reaction time required for complete digestion given that prolonged incubation may result in increased star activity.
8. To prevent re-circularization of the vector in cloning strategies, the DNA needs to be dephosphorylated by treatment with alkaline phosphatase.
9. Carbenicillin (an ampicillin analogue) is preferred to ampicillin as selecting agent, due to its higher stability and resistance to degradation.
10. The growth phase of the starter culture is a critical factor. Lag and stationary phase must be avoided; the suggested range of concentration will ensure the culture is in logarithmic growth phase.

References

1. Médigue, C., Krin, E., Pascal, G, *et al.* (2005). Coping with cold: the genome of the versatile marine Antarctica bacterium *Pseudoalteromonas haloplanktis* TAC125. *Genome Res.* **15**, 1325–35.
2. Duilio, A., Tutino, M.L., Marino, G. (2004). Recombinant protein production in Antarctic Gram negative bacteria. *Methods Mol. Biol.*, **267**, 225–237.
3. Parrilli, E., Duilio, A., Tutino, M.L. (2008). Heterologous protein expression in psychrophilic hosts. In *Psychrophiles: from Biodiversity to Biotechnology*. Edited by: RMea. *Berlin Heidelberg: Springer-Verlag*., 365–379.
4. Hensel, M., and Holden, D.W. (1996). Molecular genetic approaches for the study of virulence in both pathogenic bacteria and fungi. *Microbiology* **142**, 1049–1058.
5. Schwarzer, A. & PUhler, A. (1991). Manipulation of *Corynebacterium glutamicum* by gene disruption and replacement. *BioTechnology* **9**, 84–87.
6. McFadden, J. (1996). Recombination in mycobacteria. *Mol. Microbiol.* **21**, 205–211.

7. Stibitz, S. (1994). Use of conditionally counterselectable suicide vectors for allelic exchange. *Methods Enzymol.* **235**, 458–465.7
8. Parrilli, E., Cusano, A.M., Giuliani, M., (2006). Cell engineering of Pseudoalteromonas haloplanktis TAC125: construction of a mutant strain with reduced exo-proteolytic activity. *Microb. Cell Fact.* **5**(1), P36
9. Parrilli, E., De Vizio, D., Cirulli, C., (2008). Development of an improved Pseudoalteromonas haloplanktis TAC125 strain for recombinant protein secretion at low temperature. *Microb. Cell Fact.* **7**, 2
10. Kast, P. (1994). PKSS – a second-generation general purpose cloning vector for efficient positive selection of recombinant clones. *Gene* **138**, 109–114.
11. Li, M.Z. and Elledge, S.J. (2005). MAGIC, an in vivo genetic method for the rapid construction of recombinant DNA molecules. *Nat. Genet.* **37**, 311–319.
12. Kristich, C.J., Chandler, J.R., and Dunny, G.M. (2007). Development of a host-genotype-independent counterselectable marker and a high-frequency conjugative delivery system and their use in genetic analysis of *Enterococcus faecalis. Plasmid* **57**, 131–144.
13. Tascon, R. I., Rodriguez-Ferri, E. F., Gutierrez-Martin, C. B., (1993). Transposon mutagenesis in *Actinobacillus pleuropneumoniae* with a Tn10 derivative. *J. Bacteriol.* **175**, 5717–5722.
14. Tutino, M. L., Duilio, A., Moretti, M. A., (2000). A rolling-circle plasmid from Psychrobacter sp. TA144: evidence for a novel Rep subfamily. *Biochem. Biophys. Res. Commun.* **274**, 488–495.
15. Parrilli, E., Giuliani, M., Pezzella, C., (2010). PhPssA is required for alpha-amylase secretion in Antarctic *Pseudoalteromonas haloplanktis. Microbiol.*, **156**(1), 211–219.
16. Wittenberg, J.B., Bolognesi, M., Wittenberg, B.A., (2002). Truncated haemoglobins: a new family of haemoglobins widely distributed in bacteria, unicellular eukaryotes and plants. *J. Biol. Chem.* **277**, 871–874.
17. Giordano, D., Parrilli, E., Dettaï, A. (2007). The truncated haemoglobins in the Antarctic psychrophilic bacterium *Pseudoalteromonas haloplanktis* TAC125. *Gene* **398**, 69–77.
18. Parrilli, E., Giuliani, M., Giordano, D., (2010). The role of a 2-on-2 haemoglobin in oxidative and nitrosative stress resistance of Antarctic *Pseudoalteromonas haloplanktis* TAC125. *Biochimie* **92**, 1003–1009.
19. Cusano, A.M., Parrilli, E., Marino, G. (2006). A novel genetic system for recombinant protein secretion in the Antarctic *Pseudoalteromonas haloplanktis* TAC125. *Microb. Cell Fact.* **5**, 40.

Chapter 12

A New Bacterial Co-expression System for Over-expressing Soluble Protein and Validating Protein–Protein Interaction

Jumei Zeng and Zheng-Guo He

Abstract

Toxic, membrane, and hydrophobic proteins are usually difficult to individually over-express in *Escherichia coli* because they require a binding-partner protein for folding and stability. To obtain these types of soluble proteins or protein complexes, protein co-expression is used. Such co-expression systems are extremely suitable for the high-throughput validation of protein–protein interactions. In a previous study, we developed a novel co-expression vector, pHEX, which is compatible, and thus can be partnered, with many commercially available *E. coli* vectors, such as pGEX and pMAL. Either of the vectors allows proteins to be expressed individually as a tagged fusion protein and can be used directly for protein co-purification. This protocol presents the experimental procedure for the co-expression method.

Key words: Co-expression, pHEX, pGEX, Protein complexes

1. Introduction

Escherichia coli cells are superior host for heterologous gene expression because of their high levels of protein expression, rapid growth rate, and ease of genetic manipulation (1). However, for a number of protein types, such as toxic, membrane, or hydrophobic proteins, difficulties are usually experienced when individually over-producing them in *E. coli* (2, 3). The co-expression of multiple genes, such as members of a stable protein complex or a protein with a chaperone (4), can, in many cases, alleviate these problems (5).

Different techniques are used to co-express proteins in cells via more than two plasmids. An ideal replicon-compatible co-expression system is excellent for varying the genes of interest. Vectors employed in this system carry different antibiotic resistance and contain very convenient affinity tags, such as HA (an epitope derived from the Influenza protein haemagglutinin), MYC

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_12, © Springer Science+Business Media, LLC 2012

(an epitope derived from the human proto-oncoprotein MYC), SBP (streptavidin binding peptide), 6×His (polyhistidine-tag), GST (glutathione s-transferase), and MBP (maltose binding protein), among others, which are beneficial to co-purification experiments (6, 7). Duet co-expression vectors have been developed by Novagen (6), and Scheich and his colleagues developed the pQLink vector series (8). Dzivenu and his colleagues constructed pOKD vector series (9), Stols and his colleagues gained pMCSG derived vectors (10), and Wakamori and his colleagues developed a series of bacterial co-expression system-pWM vectors (8). In a previous study, we developed a novel co-expression vector, pHEX (11). Co-expression and co-purification could be more successful and feasible when the positive qualities of these vector systems are taken advantage of.

2. Materials

2.1. Amplification of the Target Oligonucleotide

1. 10 μM stocks of appropriate primers: Use 1 μL for a 50 μL PCR amplification reaction.
2. 100 ng/μL *Mycobacterium tuberculosis* H37Rv template DNA: Use 0.5 μL for a 50 μL PCR amplification reaction.
3. 10 mM Deoxynucleoside triphosphates (dNTPs): Use 1 μL for a 50 μL PCR amplification reaction.
4. Pyrobest DNA polymerase: Use 1 μL for a 50 μL PCR amplification reaction.
5. 10× amplification buffer (New England Biolabs): Use 5 μL for a 50 μL PCR amplification reaction.
6. Use ddH_2O, add to 50 μL.

2.2. Cloning of the Amplification Product into the Desired Vector

1. Vectors: pBT, pTRG, pHEX, and pGEX/pMAL (see Table 1).
2. Restriction enzymes.
3. T4 ligase.
4. DNA purification kits (Watson).
5. *E. coli* DH5α competent cells.
6. Plasmid extraction kits (Watson).
7. Deoxynucleoside triphosphates (dNTPs).

2.3. Co-transformation and Verification

1. *E. coli* BL21 (DE3) competent cells.
2. *E. coli* XR competent cells.
3. Plasmid extraction kits.
4. rTaq DNA polymerase.
5. Deoxynucleoside triphosphates (dNTPs).

Table 1
Vectors used in this protocol

Vector	Antibiotic resistance	Replicon	Tag	Application	Source
pBT	Chloramphenicol	p15A ori	–	Bacterial two-hybrid	Stratagene
pTRG	Tetracycline	ColE1 ori	–	Bacterial two-hybrid	Stratagene
pHEX	Kanamycin	p15A ori	6×His	Co-expression	(11)
pMAL	Ampicillin	ColE1 ori	MBP	Co-expression	(11)
pGEX	Ampicillin	ColE1 ori	GST	Co-expression	GE Healthcare

2.4. Bacterial Two-Hybrid Analysis of Protein–Protein Interactions

1. Basic medium: 1.5 g Bacto agar, 76 mL ddH_2O pH 7.0, autoclave at 121°C for 30 min.
2. 10× M9 salts: 67.8 g Na_2HPO_4, 30 g KH_2PO_4, 5 g NaCl, and 10 g NH_4Cl, add ddH_2O to 1 L.
3. Solution I: 2 mL 20% glucose (filter-sterilized), 1 mL 20 mM adenine–HCl (Sigma, filter-sterilized), and 10 mL 10× His drop-out amino acid (BD/Clontech); autoclave at 121°C for 15 min.
4. Solution II: 0.1 mL 1 M $MgSO_4$, 0.1 mL 1 M thiamine–HCl, 0.1 mL 10 mM $ZnSO_4$, 0.1 mL 100 mM $CaCl_2$, and 0.1 mL 50 mM IPTG (isopropyl-β-D-thiogalacto pyranoside).
5. Chloramphenicol stock solution: 34 mg/mL in ethanol; sterilize by filtration, store at −20°C, and use at a final concentration of 34 μg/mL.
6. Tetracycline stock solution: 12.5 mg/mL in 75% ethanol; sterilize by filtration, store at −20°C, use at a final concentration of 12.5 μg/mL.
7. Streptomycin stock solution: 8 mg/mL in double-distilled water; sterilize by filtration, store at −20°C, use at a final concentration of 8 μg/mL.
8. 1 M 3-AT (Sigma) stock solution: 8.408 g 3-AT (3-Amino-1,2,4-triazole) dissolved in 100 mL DMSO (dimethyl sulfoxide), sterilize by filtration, store at −20°C (see Note 1), and use at a final concentration of 0.5 mM.
9. CK^+: Co-transformant containing pBT-LGF2 and pTRG-Gal11P as a positive control.
10. CK^-: Co-transformant containing pBT and pTRG as a negative control.

2.5. Co-expression of the Heterodimeric Protein Complexes

1. LB medium: 10 g/L bacto tryptone, 5 g/L bacto yeast extract, and 5 g/L NaCl, autoclave at 121°C for 30 min.

2. Kanamycin stock solution: 10 mg/mL in double-distilled water; sterilize by filtration, store at −20°C, and use at a final concentration of 25 μg/mL.
3. Ampicillin stock solution: 100 mg/mL in double-distilled water; sterilize by filtration, store at −20°C, use at a final concentration of 100 μg/mL.
4. Isopropyl-β-D-thiogalactopyranoside (IPTG) stock solution: 1 M IPTG in double-distilled water; sterilize by filtration, store at −20°C, and use at a final concentration of 0.5 mM.

2.6. Co-purification Assays of a Pair of Recombinant Proteins

2.6.1. GST (Glutathione s-Transferase) Pull-Down

1. GST column: soaked in 20% ethanol, store at 4°C, pre-equilibrated with buffer A before use.
2. Buffer A (PBS, phosphate buffer solution): 150 mM NaCl, 3 mM KCl, 10 mM Na_2HPO_4, and 2 mM KH_2PO_4 pH 7.4; store at room temperature (RT), with cooling during use.
3. Buffer B: 16 mM Na_2HPO_4, 4 mM KH_2PO_4, and 20 mM GSH, store at 4°C (see Note 2).

2.6.2. His-Tag Pull-Down

1. Ni-NTA column (Qiagen): soaked in 20% ethanol, store at 4°C.
2. Binding buffer: 5 mM imidazole, 0.5 M NaCl, and 20 mM Tris–HCl pH 7.9.
3. Elution buffer: 1 M imidazole, 0.5 M NaCl, and 20 mM Tris–HCl pH 7.9.
4. 50 mM Imidazole elution buffer: 50 mL elution buffer and 950 ml ddH_2O.
5. 100 mM Imidazole elution buffer: 100 mL elution buffer and 900 ml ddH_2O.
6. 150 mM Imidazole elution buffer: 150 mL elution buffer and 850 mL ddH_2O.
7. 200 mM Imidazole elution buffer: 200 mL elution buffer and 800 mL ddH_2O.
8. 250 mM Imidazole elution buffer: 250 mL elution buffer and 750 mL ddH_2O.
9. Change buffer: 50 mM $NiSO_4$; do not adjust pH.
10. Strip buffer: 100 mM EDTA-2Na, 0.5 M NaCl, and 20 mM Tris–HCl pH 8.0.

2.6.3. MBP (Maltose Binding Protein) Pull-Down

1. Amylose resin column: soaked in 20% ethanol, store at 4°C.
2. Cell lysis buffer: 30 mM Tris–HCl pH 7.1, 1 mM EDTA, and 20% (w/v) sucrose.
3. Cell wash buffer: 20 mM Tris–HCl pH 7.4, 0.2 M NaCl, 10 mM β-mercaptoethanol, and 1 mM EDTA.
4. Column elute buffer: 20 mM Tris–HCl pH 7.4, 0.2 M NaCl, 10 mM β-mercaptoethanol, 1 mM EDTA, and 10 mM maltose.
5. Column wash buffer: 10 mM Tris–HCl pH 7.4, and 1 M NaCl.

2.7. SDS-Polyacrylamide Gel Electrophoresis

1. 2×SDS loading buffer: 100 mM Tris–HCl pH 6.8, 4% (w/v) SDS (sodium dodecyl sulfate), 0.2% (w/v) bromophenol blue, 20% (v/v) glycerol, and 200 mM DTT (DL-dithiothreitol) added before use.
2. 10×Running buffer: 250 mM Tris, 2.5 M glycine, and 1% (w/v) SDS; store at RT.
3. 30% Acrylamide/bis solution (37.5:1 with 2.6% C) and *N,N,N,N′*-tetramethyl-ethylenediamine (TEMED, BioRad).
4. 2 M Tris–HCl pH 8.8, store at RT.
5. 1 M Tris–HCl pH 6.8, store at RT.
6. 10% (w/v) SDS, store at RT.
7. Ammonium persulfate: 10% solution in double-distilled water, freeze immediately in single-use 1-mL aliquots at −20°C.
8. Isopropanol, store at RT.

2.8. Western Blotting

1. Transfer buffer: 25 mM Tris, 190 mM glycine, 20% (v/v) methanol, and 0.025% (w/v) SDS; do not adjust pH and store at RT with cooling during use.
2. Supported nitrocellulose membrane and filter paper.
3. Tris-buffered saline with Tween (TBS-T): 20 mM Tris–HCl pH 7.5, 150 mM NaCl, and 0.05% Tween 20; store at 4°C.
4. Blocking buffer: 5% (w/v) skimmed dry milk in TBS-T.
5. Primary antibody dilution buffer: TBS-T supplemented with 2% (w/v) bovine serum albumin (BSA).
6. Primary antibody (antisera to His-tag): Use at a 1:1,000 (v/v) dilution and store at −20°C.
7. Secondary antibody IgG-HRP (goat anti-rabbit): Use at a 1:10,000 (v/v) dilution and store at −20°C.
8. DAB detection reagents (Boster): Store at −20°C.

3. Methods

This co-expression system is demonstrated to efficiently co-produce and co-purify protein complexes through convenient affinity tags. In addition, it is also effectively used to characterize and confirm protein–protein interactions. The two-hybrid system is valuable for screening and determining protein–protein interactions (12, 13), but it may generate many false positives, requiring confirmation of results by another method. Protein co-purification may be an important tool for the validation of suspected protein–protein interactions or for discovering novel protein interactions (11). Association of the bacterial two-hybrid system with pull-down/co-purification

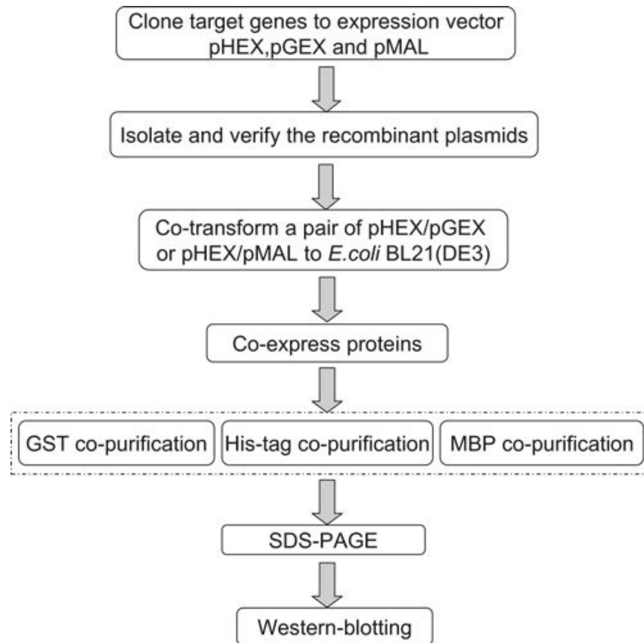


Fig. 1. Workflow of the co-expression system protocol.

assays brings more systematic and reliable results. An example of the results produced is shown in ref. 11. A summary of the method is presented in Fig. 1.

3.1. Design and Amplification of the Target Oligonucleotide

1. The appropriate PCR primers are designed and synthesized. The primers should include 17–30 nucleotides complementary to the sequence of interest and appropriate restriction enzyme sites for the chosen vector.
2. For each target to be cloned, 50 μL PCR amplification reactions are needed.
3. The target nucleic acids are amplified according to the standard PCR program. The adoption of appropriate melting temperatures is ensured to generate specific amplification.

3.2. Cloning of the Amplification Product into the Desired Vector

1. The chosen vector DNA (pBT, pTRG, pHEX, pGEX/pMAL) is prepared by digesting ~2 μg of the vector with restriction enzymes for 4 h at 37°C. The vector DNA is dephosphorylated (14) after gel-purification.
2. The dephosphorylated vector and PCR product are gel-purified using the DNA purification kit according to the manufacturer's instructions.
3. The insert DNA is prepared by digesting 50 μL of the purified PCR product with restriction enzymes for 4 h at 37°C.

4. The insert DNA is purified, and the ligation reaction is performed using 6 μL of the insert DNA, 2 μL of the vector DNA, 1 μL ligation buffer, and 1 μL ligase. The reaction is incubated at 16°C for 4 h.
5. About 5 μL of the ligation reaction are used to transform *E. coli* DH5α competent cells. The cells are plated on LB agar containing the appropriate antibiotic for the chosen vector and incubated at 37°C for 18 h.
6. The transformants are selected and 3 mL cultures of each are prepared in LB broth containing the appropriate antibiotics. The cultures are incubated with vigorous shaking at 37°C overnight.
7. About 600 μL of the cultures is removed for the preparation of a glycerol stocks (add glycerol to a final concentration of 20%). The mixture is stored at −80°C.
8. Another 1 mL of the culture is transferred to a microcentrifuge tube, and the cells are collected by centrifuging at 13,000 × *g* for 1 min. The supernatant is discarded. Plasmid DNA from each culture is isolated and bacterial clones are verified for each target by PCR and sequencing.

3.3. Co-transformation and Verification

3.3.1. For Bacterial Two-Hybrid Analysis

1. About 1 μL (20–50 ng) of each plasmid (recombined pBT, pTRG) is used to transform *E. coli* XR competent cells. These are resuscitated after 1 h.
2. The cells are plated on LB agar containing the chloramphenicol and tetracycline solutions for the chosen vector, after which they are incubated at 37°C overnight.
3. The transformants are selected and 3 mL cultures of each are set in LB broth containing chloramphenicol and tetracycline.
4. The cultures are incubated with vigorous shaking at 37°C overnight.
5. About 600 μL of the cultures is removed for the preparation of a glycerol stock (add glycerol to a final concentration of 20%) and stored at −80°C.
6. Another 1 mL of the culture is transferred to a microcentrifuge tube, and the cells are collected by centrifuging at 13,000 × *g* for 1 min. The supernatant is discarded. Plasmid DNA is isolated and bacterial clones are verified for each target by PCR.

3.3.2. For Co-expression

1. About 1 μL (20–50 ng) of each plasmid (recombined pHEX, pGEX/pMAL) is used to transform *E. coli* BL21 (DE3) competent cells. These are resuscitated after 1 h (see Note 3).
2. The cells are plated on LB agar containing the kanamycin and ampicillin solutions for the chosen vector, after which they are incubated at 37°C for 18 h overnight.

3. Several transformants are selected and 3 mL cultures of each are set in LB broth containing kanamycin and ampicillin.
4. The cultures are incubated to an OD_{600} of 0.8 with vigorous shaking at 37°C.
5. About 600 μL of the cultures are removed for the preparation of a glycerol stock (add glycerol to a final concentration of 20%) and stored at −80°C.
6. Another 1 mL of the culture is transferred to a microcentrifuge tube, and the cells are collected by centrifuging at 13,000×*g* for 1 min. The supernatant is discarded. Plasmid DNA is isolated and bacterial clones are verified for each target by PCR.

3.4. Bacterial Two-Hybrid Analysis of Protein–Protein Interactions

1. The basic medium is heated in a microwave. When the agar has cooled to 70°C, 10 mL of 10× M9 salts are added.
2. When the agar mixture has cooled to 50°C, Solutions I and II are immediately added. The mixtures are shaken gently (see Note 4).
3. Other additives are added: 100 μL chloramphenicol stock solution, 100 μL tetracycline stock solution, 100 μL streptomycin stock solution, and 500 μL 3-AT stock solution. The solution is mixed gently.
4. The solution is poured into petri dishes (~25 mL/100-mm plate). When the agar solidifies, the B2H plates are wrapped and stored at 4°C.
5. After verification of the co-transformants, 3 mL of the LB broth medium containing chloramphenicol and tetracycline are inoculated with the bacterial two-hybrid co-transformants. Bacteria are allowed to grow at 37°C overnight.
6. The culture (3 μL) is separately packed onto a B2H plate and an LB plate. CK^+ and CK^- are added.
7. The plates are incubated at 37°C for 24 h. If colonies are not apparent, the plates are transferred to RT and allowed to further incubate in a dark location (to preserve the tetracycline) for an additional 16 h.

3.5. Co-expression of the Heterodimeric Protein Complexes

1. About 3 mL of the LB medium containing kanamycin and ampicillin is inoculated with the co-transformants in glycerol stock. Bacteria are allowed to grow at 37°C overnight.
2. About 20 mL of the LB medium containing kanamycin and ampicillin is inoculated with 500 μL of the overnight culture. Bacteria are allowed to grow at 37°C with vigorous shaking until the OD_{600} reaches 0.6.
3. IPTG is added to a final concentration of 0.5 mM. The culture is grown at 37°C with vigorous shaking for 4 h (see Note 5).
4. The cells are harvested by centrifugation at 13,000×*g* for 1 min.

3.6. Growth of Cultures for Large-Scale Expression

1. Depending on the requirements of the amount of proteins to be expressed, for large-scale expression, the glycerol stock is used to inoculate a starter culture in 20 mL LB broth supplemented with kanamycin and ampicillin. Incubation at 37°C for 18 h follows.
2. About 2 mL of the starter culture is used to inoculate 100 mL of the LB broth supplemented with kanamycin and ampicillin in 250 mL flasks. Series flasks can be used for large-scale expression.
3. Growth is allowed at 37°C with vigorous shaking until OD_{600} reaches 0.6.
4. IPTG is added to a final concentration of 0.5 mM. The cultures are grown at 37°C with vigorous shaking for 4 h.
5. The cells are harvested by centrifugation at 13,000 × *g* for 1 min.

3.7. Co-purification Assays of Recombinant Proteins

3.7.1. GST Pull-Down

1. The cell pellets are resuspended in 10 mL ice-cold buffer A. About 10 μL are removed for SDS-PAGE.
2. The mixtures are then sonicated until the cloudy suspension becomes translucent (see Note 6). The cell debris is collected by centrifuging at 10,000 × *g* for 30 min at 4°C. The supernatants are collected, and the pellet and supernatant of each cleared lysate are prepared for SDS-PAGE analysis.
3. The GST column is equilibrated with 5 mL buffer A and centrifuged for 2 min at 500 × *g* (see Note 7).
4. The supernatants are loaded onto the GST column and gently mixed for 1 h (see Note 8).
5. The column is centrifuged at 500 × *g* for 2 min. The supernatant is carefully removed and discarded.
6. Each column is washed thrice with 3 mL buffer A and eluted with 500 μL Buffer B.
7. Fractions are pooled and concentrated. About 10 μL of each fraction is removed for SDS-PAGE assay. Remaining fractions are stored at –80°C for further assays.
8. The column is regenerated using the following steps: Add 3 mL Buffer A, centrifuge at 2,000 × *g* for 1 min, and discard the washing solution. Add 5 mL 2 M NaCl, centrifuge at 2,000 × *g* for 1 min, and discard it. Add 3 mL Buffer A, wash, and discard it. Finally, add 5 mL 1 M NaOH and 3 mL Buffer A. Store in 20% ethanol.

3.7.2. His-Tag Pull-Down

1. The cell pellets are resuspended in 10 mL ice-cold binding buffer. About 10 μL are removed for SDS-PAGE assay.
2. The mixtures are sonicated until the cloudy suspension becomes translucent. The cell debris is collected by centrifuging at

10,000 × *g* for 30 min at 4°C. The supernatants are collected, and the pellet and supernatant of each cleared lysate are prepared for SDS-PAGE analysis.

3. The Ni-NTA column is pre-equilibrate with 5 mL binding buffer and centrifuged for 2 min at 500 × *g*. About 5 mL of the change buffer is loaded onto the Ni-NTA column. The beads become light blue. They are centrifuged for 2 min at 500 × *g*, after which the column is washed with 5 mL binding buffer.
4. The supernatants are loaded onto the Ni-NTA column and gently mixed for 1 h.
5. The column is centrifuged at 500 × *g* for 2 min, after which the supernatant is carefully removed and discarded.
6. Each column is washed with an imidazole gradient (50 mM, 100 mM, 150 mM, and 200 mM) and eluted with 500 μL 250 mM imidazole.
7. Fractions are pooled and concentrated. About 10 μL of each fraction is removed for SDS-PAGE. The remaining fractions are stored at –80°C for further assays.
8. The column is regenerated using the following steps: Add 3 mL binding buffer, wash three times, and discard the washing solution. Add 5 mL strip buffer, centrifuge at 2,000 × *g* for 1 min, and discard it. Finally, add 5 mL binding buffer, wash, and discard it. Store in 20% ethanol.

3.7.3. MBP Pull-Down

1. The cell pellets are resuspended in 10 mL ice-cold cell lysis buffer and about 10 μL is removed for SDS-PAGE assay.
2. The mixtures are then sonicated until the cloudy suspension becomes translucent. The cell debris is collected by centrifuging at 10,000 × *g* for 30 min at 4°C. The supernatants are collected and the pellet and supernatant of each cleared lysate are prepared for SDS-PAGE analysis.
3. The amylose resin column is pre-equilibrated with 5 mL cell wash buffer.
4. The supernatants are loaded onto the amylose resin column and gently mixed for 1 h.
5. The column is centrifuged at 500 × *g* for 2 min. The supernatant is carefully removed and discarded.
6. Each column is washed with the cell wash buffer and eluted with 500 μL column elute buffer.
7. Fractions are pooled and concentrated. About 10 μL of each fraction is removed for SDS-PAGE. The remaining fractions are stored at –80°C for further assays.
8. The column is regenerated using the following steps: Add 3 mL ddH_2O, wash three times, and discard the washing solution.

Add 3 mL 0.1% SDS, centrifuge at 2,000 × *g* for 1 min, and discard it. Add 1 mL ddH_2O, wash, and discard it. Finally, add 3 mL column wash buffer, centrifuge, and discard it. Store in 20% ethanol.

3.8. SDS-Polyacrylamide Gel Electrophoresis

1. The analysis samples are prepared to include the induced cell, pellet, supernatant, and elution solution. About 10 μL of the sample are moved to a microcentrifuge tube, 10 μL 2 × SDS loading buffer is added, and the tube is heated at 95°C for 10 min.
2. SDS-PAGE is performed on a Bio-rad PowerPac Basic system. A 1-mm thick 13.5% gel is prepared by mixing of 1.71 mL water with 3 mL 30% acrylamide/bis solution, 0.94 mL 2 M Tris–HCl pH 8.8, 50 μL 10% SDS, 50 μL ammonium persulfate solution, and 5 μL TEMED. The gel is poured after mixing, leaving a space for the stacking gel, and overlaid with isopropanol. The gel is polymerized for about 30 min.
3. The isopropanol is poured off and the top of the gel is rinsed twice with water. Filter paper is used to absorb the remaining water. The stacking gel is prepared by mixing of 1.38 mL water, 0.335 mL 30% acrylamide/bis solution, 0.25 mL 1 M Tris–HCl pH 6.8, 20 μL 10% SDS, 25 μL ammonium persulfate solution, and 3 μL TEMED. About 1.5 mL of this is used to quickly rinse the top of the gel. It is also poured over the stack. The comb is inserted into the stack afterwards. The stacking gel is polymerized for 20 min.
4. The running buffer is prepared by diluting 100 mL of the 10× running buffer with 900 mL of water. The container is covered and inverted to allow mixing. Once the stacking gel curdles, the comb is carefully removed and the wells are washed with the running buffer (see Note 9). The running buffer is added to the inner and outer chambers of the gel unit, and 10 μL of each sample is added to each well.
5. Assembly of the gel unit is completed, and it is connected to a power supply. The gel always runs at 80 V through the stacking gel and 120 V through the separating gel. When the dye front runs to the bottom of gel, the run is stopped and the gel is stripped for Coomassie staining.

3.9. Western Blotting

1. The transmembrane experiment is performed on the BioRad Mini Trans-Blot cell system. Three trays are filled with the transfer buffer. The samples that had been separated by SDS-PAGE are stripped, the stacking gel is removed and discarded, and one corner is cut from the separating gel to allow its

orientation to be tracked. It is then submerged into the first tray. Two pieces of foam and four sheets of filter paper are submerged into the second tray, and a sheet of the nitrocellulose membrane just larger than the size of the separating gel is laid on the surface of the third tray. The membrane is allowed to become wet by capillary action.

2. A transfer cassette is laid out with a piece of its foam and two sheets of filter paper on one side. The separating gel is laid out on the top of the filter paper, after which the nitrocellulose membrane is laid on top of the gel, ensuring that no bubbles are trapped between the membrane and the gel. Another two sheets of filter paper are laid on top of the membrane. The second wet foam sheet is laid on top and the transfer cassette closed.
3. The cassette is placed into the transfer cell, ensuring that the nitrocellulose membrane is between the gel and the anode. It is very important to ensure this orientation (see Note 10). The lid is put on the cell and the power supply is activated. Transfers are performed at 4°C (see Note 11) and can be accomplished at 250 mA for 120 min.
4. After electrophoresis, the protein bands are transferred to a nitrocellulose membrane, and the cassette is taken out of the cell and carefully disassembled, with the top sponge and sheets of filter paper removed. The membrane is soaked in 20 mL TBS-T with 5% milk and blocked for 2 h at RT on a rocking platform.
5. The blocking buffer is discarded and the membrane is rinsed quickly prior to the addition of a 1:1,000 dilution of the primary antibody (antisera to His-tag) in primary antibody dilution buffer for 1 h at RT on a rocking platform.
6. The primary antibody is removed and the membrane is washed three times for 5 min each with 20 mL TBS-T.
7. The secondary antibody IgG-HRP (goat anti-rabbit) is freshly prepared for each experiment at 1:10,000 dilution in TBS-T and added to the membrane for 1 h at RT on a rocking platform.
8. The secondary antibody is discarded and the membrane is washed three times for 5 min each with TBS-T (see Note 12).
9. During the final wash, 500 μL aliquots of each of the DAB detection reagents are mixed in 10 mL ddH_2O. Once the final wash is removed from the blot, the blot is transferred into the DAB reagents mixture immediately. The signals are recorded by photography when they are stable.

4. Notes

1. 3-AT is toxic and it can only be stored for 1 month. Use it as soon as possible after preparation.
2. In the GSH redox cycle, reduced glutathione (GSH) is prone to forming oxidized glutathione (GSSG). Buffer B should be distributed into small volumes and not be stored for long periods of time.
3. When transforming multiple plasmids, it is essential to allow the cells to recover for at least 1 h before plating on the LB agar containing the appropriate antibiotic.
4. Moderate temperatures are low at this time and the mixture is prone to bubbles, so shake the medium gently.
5. Optimizing growth and induction conditions may increase the solubility from inclusion bodies. Lowering the induction temperature, altering the concentration of inducing agents, inducing the reaction for shorter periods of time at higher cell densities, and increasing aeration during growth are all options for either enhancing bait protein solubility or decreasing the formation of inclusion bodies. According to the characteristics of the protein, the induction temperature can be lowered to between 18 and 30°C, and the induction time can vary from 2 to 20 h.
6. Freeze-thawing treatments prior to sonication will increase lysis efficiency for certain strains, such as the BL21 series, and enhance the lysis of cultures grown to a high density. Freezing at −20°C for 15–20 min or on dry ice for 5–10 min is usually sufficient. The ultrasonic power used should be appropriate; 200 W is enough. Using powers that are too high will destroy the protein complex and lead to low or no translation efficiency. In addition, ensure that the cell lysate is at low temperature.
7. Control the amount of beads to avoid high background noise or the appearance of many contaminating bands. To ensure efficient binding, it is important not to exceed $500 \times g$ when centrifuging the beads.
8. Incubation for 1 h may help facilitate bait and prey protein complex affinity to the column. In addition, gently mixed during this incubation period may help remove non-specific adherent proteins and reduce background noise. A different incubation temperature and time may be required for your specific protein–protein interaction.
9. Washing the wells before loading the samples is beneficial to sample loading and ensuring the lanes' appearance.

10. It is vitally important to ensure this orientation. Otherwise, the proteins may be lost from the gel into the buffer rather than transferred to the nitrocellulose.
11. Adequate cooling to keep the buffer no warmer than RT by use of a refrigerated/circulating bath is essential to prevent heat-induced damage to the apparatus and the experiment.
12. Increasing the number and volume of washes may help reduce non-specific interactions. However, the stability of different protein–protein interactions is protein pair-specific and depends on the K_d (dissociation constant) of the interaction. Optimization of washing conditions may be required for less stable interactions. For example, the number of washes and the volume of each wash may need to be changed.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (30930003).

References

1. Makrides, S. C. (1996) Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol. Rev.* **60**, 512–538.
2. Wuu, J. J. and Swartz, J. R. (2008) High yield cell-free production of integral membrane proteins without refolding or detergents. *Biochim. Biophys. Acta*. **1778**, 1237–1250.
3. Zhang, J., Zhang, Y., and Inouye, M. (2003) Characterization of the interactions within the mazEF addiction module of *Escherichia coli*. *J. Biol. Chem.* **278**, 32300–32306.
4. Bross, P., Andresen, B. S., Winter, V., et al., (1993) Co-overexpression of bacterial GroESL chaperonins partly overcomes non-productive folding and tetramer assembly of *E. coli*-expressed human medium-chain acyl-CoA dehydrogenase (MCAD) carrying the prevalent disease-causing K304E mutation. *Biochim. Biophys. Acta*. **1182**, 264–274.
5. Romier, C., Ben Jelloul, M., Albeck, S., et al., (2006) Co-expression of protein complexes in prokaryotic and eukaryotic hosts: experimental procedures, database tracking and case studies. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 1232–1242.
6. Held, D., Yaeger, K., and Novy, R. (2003) New co-expression vectors for expanded compatibilities in *E.coli*. *inNovations* **18**, 4–6.
7. Wakamori, M., Umehara, T., and Yokoyama S. (2010) A series of bacterial co-expression vectors with rare-cutter recognition sequences. *Protein Expr. Purif.* **74**, 88–98.
8. Scheich, C., Kümmel, D., Soumailakakis, D., et al., (2007) Vectors for co-expression of an unrestricted number of proteins. *Nucleic Acids Res.* **35**, e43.
9. Dzivenu, O. K., Park, H. H., and Wu, H. (2004) General co-expression vectors for the overexpression of heterodimeric protein complexes in *Escherichia coli*. *Protein Expr. Purif.* **38**, 1–8.
10. Stols, L., Zhou, M., Eschenfeldt, W. H., et al., (2007) New vectors for co-expression

of proteins: structure of *Bacillus subtilis* ScoAB obtained by high-throughput protocols. *Protein Expr. Purif.* **53**, 396–403.

11. Zeng, J., Zhang, L., Li, Y., et al., (2010) Over-producing soluble protein complex and validating protein-protein interaction through a new bacterial co-expression system. *Protein Expr. Purif.* **69**, 47–53.
12. Wang, J., Jiang, P. X., Feng, H., et al., (2007) Three eukaryote-like Orc1/Cdc6 proteins functionally interact and mutually regulate their activities of binding to the replication origin in the hyperthermophilic archaeon *Sulfolobus solfataricus* P2. *Biochem. Biophys. Res. Commun.* **363**, 63–70.
13. Zhang, L., Zhang, L., Liu, Y., et al., (2009) Archaeal eukaryote-like Orc1/Cdc6 initiators physically interact with DNA polymerase B1 and regulate its functions. *Proc. Natl. Acad. Sci. USA* **106**, 7792–7797.
14. Sambrook. J., and Russell, D.W. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Chapter 13

Heterologous High-Level Gene Expression in the Photosynthetic Bacterium *Rhodobacter capsulatus*

Nadine Katzke, René Bergmann, Karl-Erich Jaeger, and Thomas Drepper

Abstract

The functional expression of heterologous genes in standard hosts such as *Escherichia coli* is often hampered by various limitations including improper folding, incomplete targeting, and missassembly of the corresponding enzymes. This observation led to the development of numerous expression systems that are based on alternative, metabolic versatile hosts. One such organism is the Gram-negative phototrophic nonsulfur purple bacterium *Rhodobacter capsulatus*. During photosynthetic growth, *R. capsulatus* exhibits several unique properties including the formation of an intracytoplasmic membrane system as well as the synthesis of various metal-containing cofactors. These properties make *R. capsulatus* a promising expression host particularly suited for difficult-to-express proteins such as membrane proteins. In this chapter, we describe a novel *R. capsulatus* expression system and its application.

Key words: *Rhodobacter capsulatus*, T7 expression system, Heterologous expression, Photosynthetic bacteria

1. Introduction

Over the past decade, various systems have been established for the functional expression of heterologous proteins in diverse hosts, including both prokaryotic and eukaryotic microorganisms (1–11). To date, bacterial systems are most widely used, predominantly due to their high productivity combined with cost efficiency and the potential for easy genetic modification.

Here, we describe a novel expression system recently developed in our group that is based on the photosynthetic Gram-negative bacterium *Rhodobacter capsulatus* (12). As a member of the group of photosynthetic nonsulfur purple α-proteobacteria, *R. capsulatus*

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_13, © Springer Science+Business Media, LLC 2012

is capable of growing either phototrophically in the light under exclusion of oxygen, acquiring energy by anoxygenic photosynthesis, or chemoheterotrophically in the dark performing aerobic respiration to generate energy. The genome of *R. capsulatus* is fully sequenced (13) and is freely available at the ERGO-Light database (http://www.ergo-light.com/ERGO/). Unique properties of *R. capsulatus* render this organism advantageous for heterologous expression of certain classes of "difficult-to-express" proteins (12, 14–17). For example, an extensive system of intracytoplasmic membranes synthesized by *R. capsulatus* during phototrophic growth allows accommodation of high levels of heterologous membrane proteins due to the highly enlarged membrane surface and the efficient membrane protein folding and translocation machinery. In addition, the capability of *R. capsulatus* to produce various types of metal-containing cofactors may be advantageous for functional heterologous expression of cofactor-dependent types of enzymes.

The *R. capsulatus* expression system is based on a set of expression plasmids conferring diverse modes of expression, selection, and purification of expression products. Plasmids of the pRhok series allow the constitutive expression under control of the promotor of the kanamycin resistance gene, P_{aphII}. By contrast, the pRhot series of expression plasmids allows the T7-RNA polymerase-dependent expression of target genes when used in combination with *R. capsulatus* strain B10S-T7 which carries a genomic T7 RNA polymerase gene under control of a fructose-inducible promoter. Plasmids of both series are available with either kanamycin and chloramphenicol or kanamycin and spectinomycin resistance and with either His_6 or StrepII as tagging options.

In this chapter, we describe experimental protocols for heterologous expression using *R. capsulatus* as an expression host.

2. Materials

2.1. Bacterial Strains, Media and Antibiotics

1. *R. capsulatus*: B10S (18) is used for constitutive expression with expression plasmids based on the pRhok series, while B10S-T7 (12) is employed for T_7-dependent expression.
2. *Escherichia coli*: DH5α (Invitrogen) is used for DNA cloning and plasmid amplification, S17-1 (19) is used for transfer of plasmids to *R. capsulatus* by conjugation.
3. Luria–Bertani (LB) medium: 10 g/L Tryptone (peptone from casein), 5 g/L yeast extract and 5 g/L NaCl is prepared as described by Sambrook and Russell (20).
4. Trace Element Solution: 0.4 g $MnSO_4 \cdot H_2O$, 0.7 g H_3BO_3, 0.01 g $Cu(NO_3)_2 \cdot 3H_2O$, 0.06 g $ZnSO_4 \cdot 7H_2O$, and 0.02 g $Na_2MoO_4 \cdot 2H_2O$. Dissolve in 250 mL H_2O.

5. RCV minimal medium: 2 mL 1% EDTA, 1 mL 20% $MgSO_4$, 1 mL Trace Element Solution, 1 mL 7.5% $CaCl_2$, 2.4 mL 0.5% $FeSO_4$, 1 mL 1% thiamine, and H_2O to 1 L. Autoclave the basic solution (121°C, 2 bar, 20 min) and allow to cool, then add 9.6 mL 1 M phosphate buffer (81.3 g KH_2PO_4 and 78.7 g K_2HPO_4 in 500 mL H_2O pH 6.8), 40 mL of a 10% DL-malate solution pH 6.8 and 10 mL of a 10% $(NH_4)_2SO_4$ solution (see Note 1).
6. PY medium: 10 g/L Bacto peptone (BD) and 0.5 g/L Bacto yeast extract (BD). After sterilization (121°C, 2 bar, 20 min) and cooling, add 2 mL/L of each 1 M $MgCl_2$ and 1 M $CaCl_2$, as well as 2.4 mL/L of a 0.5% $FeSO_4$ solution (supplemented with 2 mL/L 37% HCl) are added to the medium. For biparental mating experiments with *E. coli* S17-1, this medium is prepared without $FeSO_4$ (see Note 2).
7. PYG buffer: 3 g/L Bacto peptone and 3 g/L Bacto yeast extract. 2.5 mL/L of a 20% $MgCl_2$ solution and 4 mL/L of a 7.5% $CaCl_2$ solution are added after sterilization.
8. If necessary, antibiotics are added to culture media to a final concentration of 25 μg/mL kanamycin, 10 μg/mL spectinomycin, or 200 μg/mL streptomycin.
9. Agar plates: Liquid media supplemented with 1.5% (w/v) agar. For *E. coli*, add agar-agar (Carl Roth), or for *R. capsulatus* add select agar (Invitrogen), to the media prior sterilization.

2.2. Vectors

1. pRhokHi-2 (Fig. 1): The P*aphII* expression vector contains a broad-host-range origin of replication suitable for both *R. capsulatus* and *E. coli*, the *aphII* gene conferring resistance to kanamycin as well as a MOB site for plasmid mobilization to *R. capsulatus*. The expression of target genes cloned into the multiple cloning site is under control of the P*aphII* promoter, enabling constitutive expression at moderate levels. A His_6-tag downstream of the target gene allows for easy purification and immunodetection by using affinity chromatography and specific His_6-tag antibodies, respectively.
2. pRhotHi-2 (Fig. 1): The P_{T7} expression vector is identical to pRhokHi-2 with exception of the orientation of the *aphII* gene. A T_7 promoter upstream and a T_7 terminator downstream of the multiple cloning site allow for T_7 RNA polymerase dependent high-level expression of subcloned target genes.
3. pRhokHi-6 and pRhotHi-6, which are otherwise identical to pRhokHi-2 and pRhotHi-2, respectively, harbor a spectinomycin resistance gene instead of a chloramphenicol resistance gene.
4. pRhokS-2 and pRhotS-2, which are otherwise identical to pRhokHi-2 and pRhotHi-2, respectively, carry an alternative DNA sequence encoding the StrepII-tag instead of His_6-tag.
5. The properties of all pRho vectors are summarized in Table 1.

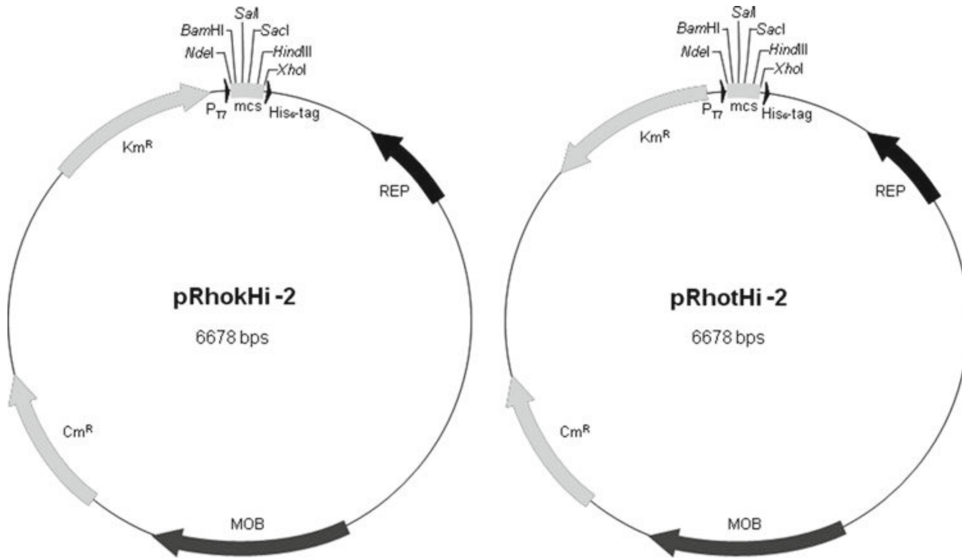


Fig. 1. The *Rhodobacter capsulatus* pRho expression vectors. The pRhokHi-2 vector harbors the constitutive promoter of the *aphII* kanamycin resistance gene (Km^R) that facilitates moderate heterologous gene expression. By contrast, the pRhotHi-2 vector allows the T7 RNA polymerase dependent high-level expression via the T7 promoter (P_{T7}). Both vectors carry an origin of replication (REP), an origin of transfer (MOB) and an additional chloramphenicol resistance gene (Cm^R). The target gene can be cloned into the multiple cloning site (MCS) and is fused to a His_6-tag.

Table 1
Properties of the pRho expression vectors

pRho vector	Promoter	Induction	Antibiotic resistance	Affinity tag
pRhokHi-2	P_{aphII}	Constitutive	Km, Cm	His_6
pRhokHi-6	P_{aphII}	Constitutive	Km, Sp	His_6
pRhokS-2	P_{aphII}	Constitutive	Km, Cm	StrepII
pRhotHi-2	P_{T7}	Inducible	Km, Cm	His_6
pRhotHi-6	P_{T7}	Inducible	Km, Sp	His_6
pRhotS-2	P_{T7}	Inducible	Km, Cm	StrepII

2.3. Solutions for Preparation of Competent E. coli Cells and Transformation

All solutions used for preparation of competent *E. coli* cells and for transformation are sterilized at 121°C, 200 kPa, and 21 min.

1. Mg^{2+} solution: 500 mM $MgCl_2$ and 500 mM $MgSO_4$.
2. TMF buffer: 100 mM $CaCl_2$, 50 mM $RbCl_2$, and 40 mM $MgCl_2$.

2.4. Solutions for Protein Isolation and Detection

1. KH_2PO_4 wash solution: 0.01 M KH_2PO_4 pH 7.0.
2. SP resuspension buffer: 22 mM KH_2PO_4, 40 mM K_2HPO_4, and 150 mM NaCl. If necessary, this buffer can be supplemented with protease inhibitor tablets (e.g., Complete, EDTA-free; Roche, 1 tablet per 50 mL).
3. SDS sample buffer: 1.25 mL of a 0.5 M Tris–HCl pH 6.8, 1.2 mL 87% glycerol, 3 mL 10% SDS, 0.5 mL β-mercaptoethanol, 0.6 mL 0.5% bromophenol blue solution, and 2.65 mL Millipore water (see Note 6).
4. 20× NUPAGE running buffer (Invitrogen).
5. 20× NUPAGE blotting buffer (Invitrogen).
6. Coomassie blue staining solution: 0.1% (w/v) Coomassie Brilliant Blue pigment (Serva) in 42% (v/v) ethanol (technical) and 16% (v/v) acetic acid in H_2O. The solution is reusable several times.
7. Coomassie blue destaining solution: 20% ethanol (technical) and 7% acetic acid in dH_2O. The solution may be used several times when filtered through active coal.
8. 20× TBS buffer: 60.5 g Tris, 87.6 g NaCl, and 2 g $MgCl_2$ in 500 mL H_2O. Adjust pH to 6.8 with HCl. Note that the components of the solution will not dissolve properly while at a pH above 7.
9. TBST buffer: Dilute 50 mL of TBS buffer in 950 mL of H_2O, supplemented with 2% Tween 20 (Carl Roth).
10. ECL solution A: 37% hydrogen peroxide solution.
11. ECL solution B: 11 mg *p*-coumaric acid (Sigma-Aldrich) in 10 mL dimethylsulfoxide (DMSO).
12. ECL solution C: 50 mg Luminol sodium salt (Sigma-Aldrich) in 500 mL 100 mM Tris–HCl pH 6.8. Store at 4°C.
13. Amidoblack staining solution: 0.1% (w/v) Amidoblack pigment (Merck) in 250 mL mixture of 45% ethanol (technical) and 10% acetic acid. The solution is reusable several times; store at room temperature (RT) in the dark.
14. PBS buffer: 4 mM KH_2PO_4, 16 mM Na_2HPO_4, and 115 mM NaCl pH 7.4.
15. Reaction buffer: 100 mM NaCl, 5 mM $MgCl_2$, and 100 mM Tris–HCl pH 8.8.
16. NTB solution: 7.5% (w/v) Nitrotetrazolium blue in 70% (v/v) dimethylformamide. Store at –20°C.
17. BCIP solution: 5% (w/v) 5-Bromo-4-chloro-3-indolyl-phosphate in dimethyl-formamide. Store at –20°C.

2.5. Solutions for Agarose Gel Electrophoresis

1. 5× TBE buffer: 89 mM Tris–HCl pH 8.3, 89 mM borate, and 2.5 mM EDTA.

2.6. Commercial Kits

1. Plasmid DNA preparation: innuPrep Plasmid Mini Kit (Analytik Jena).
2. DNA isolation from agarose gels: innuPrep DOUBLEpure Kit (Analytik Jena).
3. Anti-His (C-term)-HRP antibody (Invitrogen).
4. Strep-Tactin® AP conjugate detection kit (iBA GmbH).

2.7. Enzymes

Enzymes are obtained from Fermentas and applied with buffers at optimal reaction temperature:

1. Restriction enzymes: *Nde*I (10 U/μL), *Xho*I (10 U/μL).
2. T4 DNA Ligase: 1 U/μL.

2.8. DNA Ladder

1. GeneRuler 1 kb DNA ladder (Fermentas), 14 fragments (in bp): 10,000, 8,000, 6,000, 5,000, 4,000, 3,500, 3,000, 2,500, 2,000, 1,500, 1,000, 750, 500, 250.

2.9. Protein Ladder

1. Precision Plus Protein Standard Dual Color (Bio-Rad), 10 bands (in kDa): 250, 150, 100, 75, 50, 37, 25, 20, 15, 10.

2.10. Devices

The devices described here may be replaced by alternative devices with adequate specifications.

1. For anaerobic cultivation of *R. capsulatus* on solid media, the Microbiology Anaerocult A system (Merck), consisting of air-tight containers as well as gas packs to deplete atmospheric oxygen, is used.
2. French Press cell disruptor (Thermo Scientific).
3. French pressure cell: 40 K, 1″ piston diameter (Thermo Scientific).
4. Xcell SureLock™ electrophoresis cell with SDS PAGE and Western blot appliances (Invitrogen).
5. NUPAGE 4–12% Bis-Tris ready to use gels (Invitrogen).
6. STELLA Imaging System (Raytest).

3. Methods

3.1. Restriction of Vectors and Insert Fragments

1. Hydrolyze 5 μg DNA (vector or PCR product) with 1.5 U of each *Nde*I and *Xho*I in a total reaction volume of 50–100 μL containing the appropriate reaction buffer for double digestion (for Fermentas enzymes: red or orange) (see Note 3).
2. Incubate reaction mixture at 37°C overnight (see Note 4).

3. Inactivate the restriction enzymes by heat treatment (at 80°C for 20 min).
4. Add 10 μL DNA sample buffer and apply the whole mixture to agarose gel electrophoresis (Subheading 3.3).
5. Excise a gel slice with the DNA fragment of the appropriate size (~6.6 kbp for vectors pRhokHi-2 and pRhotHi-2). Purify the vector from the gel slice using the innuPrep DOUBLEpure kit.

3.2. Cloning of DNA Fragments into pRho Plasmids

1. A DNA fragment containing the gene to be expressed in *R. capsulatus* is obtained from plasmid or genomic DNA by PCR amplification. Primers may be designed with standard methods, but appropriate restriction sites have to be added to each primer to clone the fragment into plasmids pRhokHi-2 or pRhotHi-2: An *Nde*I-site has to be attached to the up-primer that specifically binds at the 5′-end of the target gene including the start codon (see Note 5). Add at least six further bases to the 5′ end of the up-primer for optimal restriction of the PCR product with *Nde*I. For the down-primer, design the primer without the stop codon to fuse the gene product to one of the short tag peptides and add an *Xho*I restriction site. For efficient *Xho*I-digestion of the PCR product, extend this primer at the 5′-end by at least three bases.
2. Amplify the desired gene by PCR.
3. Digest both the PCR fragment from step 2 and the target vector separately by restriction with *Nde*I and *Xho*I (Subheading 3.1). Purify the DNA fragments by elution from an agarose gel using the innuPrep DOUBLEpure kit.
4. Use 5 μL each of a solution containing vector and insert for agarose gel electrophoresis (Subheading 3.3). Estimate DNA concentrations by comparison with DNA standards. For the ligation reaction, calculate the ratio for fragment and vector DNA using the following formula:

$$\text{Mass fragment}[\text{ng}] = (5 \times \text{mass vector}[\text{ng}] \times \text{fragment length}[\text{bp}]) / (\text{vector length}[\text{bp}]).$$

   Usually, 1 μL of vector DNA with the calculated amount of insert DNA is sufficient for efficient ligation; however, at low DNA concentrations, those volumes may be increased.

5. Mix insert and vector DNA in the calculated amounts and dilute the mixture with Millipore H_2O to a volume of 17 μL (see Note 6). Add 2 μL 10× T4 DNA ligase buffer (Fermentas) and 1 μL T4 DNA ligase (1 U/μL) and mix gently by pipetting up and down.

6. Incubate at RT for 2 h or at 16°C overnight. After incubation, inactivate the ligase by heating to 65°C for 10 min.
7. Use the reaction mixture for heat shock transformation into *E. coli* DH5α cells (Subheading 3.4.2).
8. Pick ten clones from the transformation plate and grow them separately in 5 mL LB-medium containing kanamycin at 37°C overnight. Harvest cells and isolate plasmid DNA using the innuPrep Plasmid Mini Kit as indicated by the manufacturer. Analyze a 5 μL aliquot of each isolated DNA sample for correct insertion of the recombinant gene by restriction with *Nde*I and *Xho*I (Subheading 3.1) followed by separation on a 0.8% agarose gel (Subheading 3.3). When cloning of the insert was successful, two bands should be visible in the gel at the appropriate positions for vector (~6.6 Kb) and insert.

3.3. Agarose Gel Electrophoresis

1. Agarose gel electrophoresis is performed according to a standard protocol and gels are stained with ethidium bromide (see ref. 20).
2. Prepare a solution of 0.8% (w/v) agarose in TBE buffer. Boil until the agarose has been completely dissolved. Disperse 10 μL of ethidium bromide per 50 mL into the solution.
3. Cast the gel and transfer it into the electrophoresis equipment. Fill the electrophoresis chamber with 0.5× TBE buffer and apply samples to different wells.
4. Agarose gel electrophoresis is performed at 100 V for 20 min (time depends on gel size).
5. Ethidium bromide-stained DNA bands can be detected by subjecting the gel to UV light.

3.4. Transformation of Competent E. coli Cells

3.4.1. Preparation of Competent E. coli Cells

1. Transfer *E. coli* cells from a fresh LB master plate into 5 mL LB medium and incubate overnight at 37°C with constant shaking.
2. Prepare 100 mL LB medium supplemented with 2 mL Mg^{2+} solution in a 500 mL Erlenmeyer flask. Inoculate with 1 mL of the preculture from step 1. Grow this culture to an optical density (OD_{580} nm) of 0.4–0.6 and then keep it on ice during all following steps.
3. Harvest cells by centrifugation (3 min, 4,500×*g*, 4°C). Resuspend the pellet in 50 mL TMF buffer which has been precooled to 4°C. Incubate on ice for 30–60 min.
4. After a second centrifugation step, resuspend the cells in 10 mL of chilled TMF buffer containing 20% (v/v) glycerol.
5. Divide the cell solution into aliquots of 200 μL each. Competent cells may be stored at −80°C.

3.4.2. Heat Shock Transformation of Competent E. coli Cells

1. Thaw one aliquot (200 μL) of competent *E. coli* cells (from Subheading 3.4.1) on ice. Mix the *E. coli* cells with 100 μL chilled TMF buffer where either a previously prepared ligation mixture (Subheading 3.2) or 5 μL of isolated plasmid DNA (20 ng) was added. Mix gently by pipetting up and down.
2. Incubate the mixture on ice for 30 min, then perform a heat shock by transferring the reaction tube to a preheated thermo block at 42°C. Keep the tube at this temperature for 2 min, then remove the tube from the heating block and incubate it for 5 min on ice. Add 700 μL of LB medium without antibiotics.
3. Incubate the transformation tubes at 37°C for 2 h under constant shaking.
4. Centrifuge the cell suspension at 16,000 × *g* for 3 min at RT. Remove 800 μL of supernatant and resuspend the cells in the remaining liquid. Plate the cells on LB agar plates containing kanamycin.

3.5. Conjugational Transfer of Plasmid DNA into R. capsulatus

Plasmids are transferred into *R. capsulatus* by conjugation; transformation does not work.

1. Two days before conjugation starts, streak the *R. capsulatus* recipient strain on PY agar plates and incubate at 30°C for 48 h under phototrophic conditions (i.e., anaerobically in the light, Subheading 3.7.2). The day before conjugation, transform the mobilizable plasmid (e.g., pRho vectors) into the *E. coli* donor strain S17-1 (Subheading 3.4.2).
2. For conjugational transfer of mobilizable plasmids from *E. coli* S17-1 to *R. capsulatus*, first completely scrape *R. capsulatus* single colonies off the plate from step 1 and resuspend the cells in 5 mL RCV medium by pipetting up and down until no cell clumps remain. For a single conjugation approach, 1 mL of this cell suspension is transferred to a 2-mL reaction tube.
3. Resuspend 20–30 single colonies of donor *E. coli* cells in 1 mL iron-free PY medium. Handle the cells gently, since when treated too roughly the cells might lose their pili, thus reducing conjugation efficiency.
4. Add 0.5 mL of *E. coli* suspension to the *R. capsulatus* cell suspension.
5. Centrifuge for 10 min at 16,000 × *g* at RT. During centrifugation, prepare PY agar plates without antibiotic by placing a sterile cellulose acetate filter (0.2 μm pore size, 25 mm diameter, Whatman) in the center of the plate.
6. After centrifugation, remove supernatant except for a small volume to resuspend the cells. Again, resuspension has to be carried out gently (e.g., via agitation with a pipette tip) to prevent *E. coli* donor cells from shearing their pili. Transfer the whole

resuspension volume onto the filter on the prepared PY agar plate. Ensure to avoid spilling cell suspension from the filter to the agar plate.

7. Incubate the plate overnight at 30°C in the dark.
8. Fill a 2-mL reaction tube with 1 mL RCV medium without antibiotic. Transfer the filter with the cell suspension into the tube. Ideally, the filter should neatly line the tube wall after insertion, as any crumpling or folding could decrease resuspension efficiency. Vortex vigorously to remove the cell suspension from the filter (the vortex step immediately stops cell conjugation). Use a pipette to break down any remaining cell clumps.
9. Plate 250 μL of resuspension onto a PY agar plate selective for the transferred plasmid. To avoid growth of the *E. coli* donor strain on selective PY agar plates, 200 μL/mL streptomycin is also added to the medium (see Note 7). Centrifuge the remaining cell suspension (10 min at 16,000 × *g*), remove most of the supernatant and use the remaining supernatant to resuspend the pellet. Plate 250 μL of this resuspension on a second selective PY agar plate.
10. Incubate the plates under phototrophic growth conditions (Subheading 3.7.2) until red *R. capsulatus* colonies have reached a diameter of about 1 mm (usually after 2–3 days of inoculation).

3.6. Preparation of *R. capsulatus* Freeze Cultures

1. Recombinant *R. capsulatus* strains can be stored for several years at −20°C. First, prepare a phototrophically grown PY agar plate with the desired *R. capsulatus* strain (Subheading 3.7.2).
2. Harvest the cell material from whole plate and resuspend it in 3 mL RCV medium without antibiotic.
3. Centrifuge 1.5 mL of the cell suspension at 16,000 × *g* for 10 min. Remove supernatant with a pipette and resuspend the pellet in 1 mL PYG buffer.
4. Divide the cell suspension in two sterile 2 mL reaction tubes. Add 0.75 mL of 87% glycerol to each sample and mix by inverting the tube several times until the cell suspension is completely dispersed in the glycerol.
5. Store *R. capsulatus* freeze cultures at −20°C (see Note 8).

3.7. Cultivation of *R. capsulatus*

3.7.1. General Remarks

1. *R. capsulatus* is a facultative phototrophic bacterium. It can be cultivated either under chemoheterotrophic conditions (i.e. aerobically in the dark) or phototrophic conditions (i.e., anaerobically in the light) (see Note 9).
2. For growth of *R. capsulatus* on solid media, PY agar plates are used. Growth in liquid medium is usually carried out in RCV medium.

3. As mentioned above, both *R. capsulatus* B10S and B10S-T7 are streptomycin resistant. Thus, beside contraselection against the *E. coli* donor strain after conjugation, 200 μL/mL streptomycin can also be added to PY medium (solid or liquid) to avoid contamination. In RCV minimal medium, however, this antibiotic is omitted as it may pose too much stress on the cells.
4. The optimal cultivation temperature for *R. capsulatus* is 30°C. Deviations from this temperature may severely hamper cell growth.
5. To assess cell density, the optical density of a liquid culture is measured at a wavelength of 660 nm (OD_{660}) (see Note 10).

3.7.2. Cultivation on Solid Medium

1. As solid medium, PY agar plates are used.
2. Streak a small amount of *R. capsulatus* cell mass onto the plate.
3. Plates of *R. capsulatus* can be either incubated under aerobic conditions by placing them in an appropriate cell incubator in the dark or phototrophic conditions. For phototrophic growth, the plates are placed in a gas-tight jar (Fig. 2a) (e.g., Microbiology Anaerocult A system, Merck). Here, a disposable gas generating envelope (Gas Pack for Microbiology Anaerocult A system, Merck) can be used for the generation of an anaerobic environment. Therefore, douse the powder in the gas pack evenly with 15 mL of distilled water, put it into the jar and close the jar (see Note 11). Place the jar between six customary 60 W light bulbs for proper illumination.
4. In both cases, incubation is carried out at 30°C for 48–72 h.

3.7.3. Preparation and Growth of R. capsulatus Precultures

1. For inoculation of expression cultures (i.e., the main culture within an expression experiment), first prepare a 10 mL preculture.
2. Fill an appropriate volume of RCV medium into a Hun gate anaerobic reaction tube (Fig. 2b). If necessary add antibiotics for the stable maintenance of explicative plasmids (e.g., kanaymcin for pRho vectors).
3. Pick a single colony from freshly grown *R. capsulatus* master plates (Subheading 3.7.2). Resuspend the cells in the RCV medium by pipetting up and down.
4. Incubate the precultures as instructed for anaerobic cultivation (Subheading 3.7.5).
5. After incubation, measure the optical density (OD_{660}) of the preculture. Calculate the volume of preculture necessary to inoculate the subsequent expression culture with a cell density corresponding to an OD_{660} of 0.02 (see Note 12).

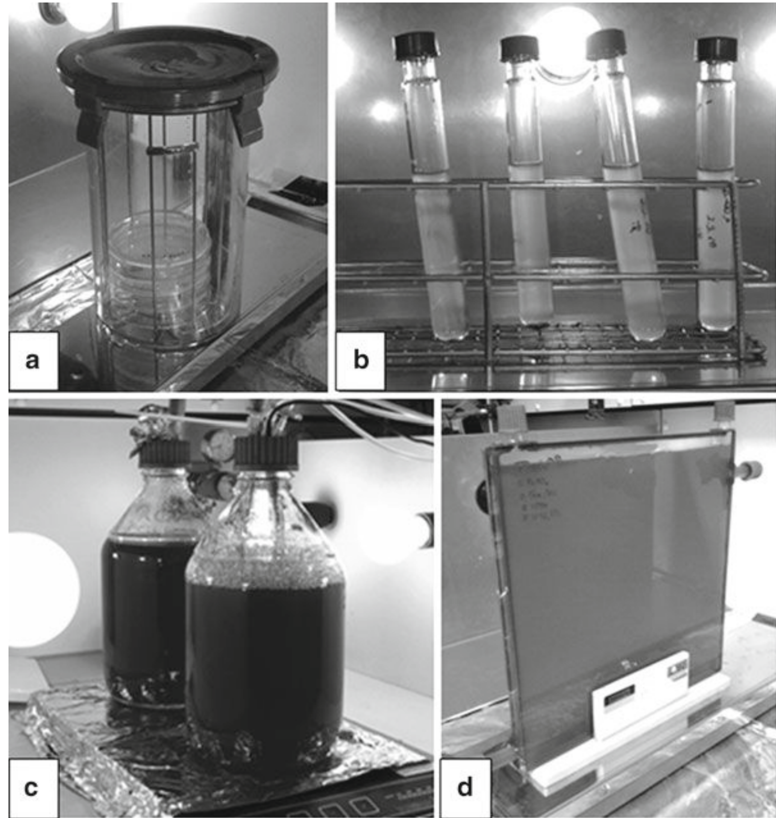


Fig. 2. Phototrophic growth of *Rhodobacter capsulatus*. A variety of different gas-tight containers may be used for anaerobic incubation of *R. capsulatus*. Cultivation on solid medium is carried out in gas-tight anaerobic jars (Merck) equipped with oxygen-depleting packs (**a**). For cultivation in liquid medium, diverse vessels such as Hungate tubes (**b**), bottles (**c**), or flat-panel reactors (custom-made, **d**) may be used depending on culture volume.

3.7.4. Cultivation of R. capsulatus Expression Cultures Under Aerobic Conditions

1. Prepare a sterile 1 L Erlenmeyer flask containing 100–200 mL RCV liquid medium supplemented with kanamycin.
2. Inoculate the expression culture from an anaerobic preculture (Subheading 3.7.3) to an OD_{660} of 0.02. For T_7-dependent expression, add 12 mM fructose to the medium as an inducer (see Note 13).
3. Grow the aerobic cultures at 30°C in a microbiological incubator (shaking frequency 100 rpm) in complete darkness to a maximum OD_{660} of 1. If necessary, wrap aluminum foil around the cultivation flask to avoid illumination of the cultures.
4. Harvest cells by centrifugation for 20 min, at 6,000 × *g* and RT.

3.7.5. Cultivation of R. capsulatus Expression Cultures Under Phototrophic Conditions

1. For anaerobic cultivation of *R. capsulatus*, various gas-tight vessels can be used (Fig. 2b–d). Small volumes (up to 15 mL) are cultivated in Hun gate anaerobic tubes that can be sealed with a screw cap and a rubber septum. Larger volumes (up to 2 L)

are cultivated in appropriate bottles or Erlenmeyer flasks sealed with rubber plugs with septum. To avoid self-shading, flat panel bioreactors can be used. We are using self-made 600-mL bioreactors with a width of 1 cm that are also closed by rubber-sealed screw caps (Fig. 2d). All kinds of containers may usually be filled to maximum volume as oxygen availability is not an issue under anaerobic conditions (see Note 14).

2. Inoculate the culture from an anaerobic preculture (Subheading 3.7.3) with a cell density according to an OD_{660} of 0.02. For T_7-dependent expression, add 12 mM fructose to the medium as an inducer (see Note 13).
3. Seal the vessel by a gas-tight screw cap with rubber septum (Hungate tubes and flat panel reactors) or a rubber plug with septum (Erlenmeyer flasks and bottles).
4. Attach a flexible tube to an argon source and add an injection needle to the free end of the tube. Use the injection needle to pierce the septum of the gas-tight vessel. Add a second injection needle to the septum to create an outlet for excess gas. Flush gas through the septum into the vessel for 5–10 min (depending on the size of dead volume) to exchange all oxygen in the gas phase by argon (see Note 15).
5. Incubate the expression cultures phototrophically by illuminating them with six 60 W light bulbs (~2,500 lux) on the bench. At constant RT, the heat emitted by the light bulbs might be sufficient to incubate the cultures at 30°C without further technical equipment. In that case, temperature within the cultivation vessel should be controlled frequently. If the diameter of the cultivation vessel exceeds 1 cm, the culture can be stirred slowly during incubation to reduce self-shading effects.
6. Harvest cells by centrifugation for 20 min, at 6,000 × *g* and RT (see Note 16).

3.8. Analysis of Protein Accumulation in *R. capsulatus* Expression Strains

3.8.1. Disruption of *R. capsulatus* Cells via French Press

1. After harvesting the *R. capsulatus* cells from expression cultures, wash the cells two times with 0.01 M KH_2PO_4 buffer pH 7.0.
2. Resuspend the cell pellet in a small volume of SP buffer supplemented with protease inhibitor. Dilute the cell suspension to a volume of maximum 30 mL depending on the optical density of the respective culture (the final cell density should correspond to an OD_{660} of 5). Higher cell densities (up to OD_{660} 10) will result in lower disruption efficiency, but will increase protein concentration in the final homogenate.
3. Assemble the French Press cell (Thermo Scientific) according to manufacturer's instructions. Insert the disruption cell into the French Press device and set up the device to a pressure of 550 bar ("high").

4. Pull sample into the French Press cell via the inlet/outlet tube.
5. Slowly press the sample back into the sample tube. Ensure that the pressure never drops below 550 bar by careful regulation via the adjusting screw.
6. Repeat steps 4 and 5 five times. The sample should appear as a transparent reddish solution.

3.8.2. Cell Fractionation

1. To separate soluble from insoluble and membrane bound proteins, first centrifuge cell lysates (Subheading 3.8.1) at 2,500 × *g* for 5 min, at 4°C. The resulting pellet comprises cell debris, undisrupted cells, and proteins that aggregated in insoluble inclusion bodies. The supernatant contains soluble and membrane proteins.
2. Centrifuge supernatant by ultracentrifugation at 135,000 × *g*, and 4°C, for 1 h. The supernatant comprises soluble proteins, while the pellet exclusively contains proteins bound to the cell membranes.

3.8.3. Protein Sample Preparation for SDS-PAGE

1. Prepare a 100 µL aliquot from either whole cells (washed twice with KH_2PO_4-solution) or cell lysate (Subheading 3.8.1) and cell fractions (Subheading 3.8.2) containing a cell density corresponding to an OD_{660} of 2–3 (see Note 17).
2. Add 100 µL of SDS sample buffer and mix by pipetting.
3. Boil the sample at 100°C for 12 min while shaking vigorously. Sample tubes should be tightly closed at all times to prevent vaporization of the samples.
4. Let the sample cool to RT for 5 min, then spin down the condensed water collected under the lid. Vortex the sample briefly.
5. Samples can be used instantly for SDS-PAGE (Subheading 3.8.4) or may be stored at −20°C (when stored, repeat this protocol from step 3 prior to electrophoresis). Sometimes, a centrifugation step (5 min at 14,000 × *g* and RT) is necessary before loading the samples on a gel, as nonprotein material may interfere with clean separation.

3.8.4. SDS Polyacrylamide Gel Electrophoresis

1. Prepare a NUPAGE 4–12% Bis-Tris ready to use gel (Invitrogen) in a Xcell SureLock™ electrophoresis cell and fill the buffer chamber of the cell with 500 mL NUPAGE running buffer (Invitrogen).
2. Apply 10 µL samples and protein marker on the gel and run the protein gel electrophoresis at 200 V until the blue band reaches the bottom of the gel (see Note 18).

3. Remove the gel from the plastic casing and wash it briefly in dH_2O. The gel may be used for Western Blot (Subheading 3.8.5) and subsequent immunodetection of tagged proteins (Subheading 3.8.6). Alternatively, the gel may be stained by Coomassie Blue (proceed with step 4).
4. Incubate the gel in Coomassie Blue staining solution for 20 min. Shake gently during staining. Rinse the gel with distaining solution once, then incubate the gel in distaining solution overnight on a shaking platform at RT and photograph for documentation.

3.8.5. Transfer of Proteins onto PVDF Membranes (Western Blot)

1. Equilibrate the PVDF membrane (Bio-Rad) in methanol for 1 min and then transfer the membrane into 20 mL of NUPAGE transfer buffer (Invitrogen).
2. Assemble the blotting sandwich and chamber (Xcell SureLock™ electrophoresis cell with blotting appliance) as instructed by the manufacturer. Fill the chamber with 800 mL of transfer buffer and blot for 1 h at 25 V.
3. Disassemble the blotting chamber and retrieve the membrane. Proceed with the immunodetection procedure (Subheading 3.8.6) or stain the transferred proteins by incubating the membrane in Amidoblack solution for 10 min, followed by air-drying on the bench.

3.8.6. Immunodetection of Heterologous Proteins

3.8.6.1. Immunodetection of His_6-Tagged Proteins

1. After Western blotting incubate a PVDF membrane with transferred proteins (Subheading 3.8.5) in TBST solution containing 1.5% skim milk powder for 2 h at RT to block free binding sites of the membrane. Alternatively, incubation may be carried out at 4°C overnight (see Note 19).
2. Wash membrane briefly in TBST buffer.
3. Incubate membrane in 25 mL TBST buffer containing Anti-His (C-term)-HRP antibody (dilution 1:5,000, v/v) for 1 h under constant shaking. Ensure that the membrane is completely covered by the antibody solution during the procedure.
4. Wash the membrane by shaking it in 25 mL of TBST buffer several times. Do first wash for 15 min and subsequent ones for 5 min to remove free antibodies.
5. Place the membrane in a clear petri dish. Mix 0.3 µL of ECL solution A, 100 µL of ECL solution B and 1 mL of ECL solution C in a reaction tube. Disperse the mixture over the membrane several times; ensure that the membrane is completely and evenly covered.
6. Chemiluminescence of the antibody-labeled recombinant protein can be detected by a chemiluminograph (e.g., STELLA system, Raytest).

3.8.6.2. Immunodetection of StrepII-Tagged Proteins

For the immunodetection of StrepII-tagged proteins the Strep-Tactin® AP conjugate detection kit (iBA) is used.

1. After Western blotting incubate a PVDF membrane with transferred proteins (Subheading 3.8.5) in PBS solution containing 3% BSA and 0.5% Tween 20 (v/v) for 1 h at RT to block free binding sites of the membrane. Alternatively, incubation may be carried out at 4°C overnight.
2. Wash membrane three times in PBS buffer containing 0.1% (v/v) Tween 20.
3. Incubate membrane in 10 mL PBS buffer with 2.5 μL Strep-Tactin alkaline phosphatase conjugate antibody (dilution 1:4,000, v/v) for 1 h under constant shaking. Ensure that the membrane is completely covered by the antibody solution during the procedure.
4. Wash the membrane twice for 1 min in PBS buffer containing 0.1% (v/v) Tween 20 to remove free antibodies.
5. Wash the membrane twice for 1 min in PBS buffer to remove the detergent.
6. Place the membrane in a clear petri dish. Mix 20 mL reaction buffer with 10 μL NBT solution and 60 μL BCIP solution in a reaction tube.
7. Incubate the membrane in the reaction solution until purple colored signals appear. Proceed with the chromogenic reaction until optimal signal-to-background ratio is achieved (see Note 20).
8. Stop the reaction by washing the membrane either several times with distilled water or rinse the membrane in distilled water until all reaction solution is removed.
9. Air-dry the membrane and store it in the dark.

4. Notes

1. Addition of the listed components before cooling-down leads to irreversible precipitation. $(NH_4)_2SO_4$ is the standard nitrogen source used in most experiments. In the literature, serine is described as a substitute nitrogen source (applied at the same concentration); note however, that the use of serine results in slower growth and activation of the enzyme nitrogenase, resulting in production of high amounts of hydrogen gas that might lead to explosion of gastight glass vessels (risk of injury!). In this case, a dead volume of about one third of the culture vessel should be maintained.
2. This is important since *E. coli* is not able to survive on high iron concentrations.

3. While the *Nde*I site has to be used for cloning to ensure the optimal distance between gene and ribosome binding site, the *Xho*I site may be exchanged by another site for a single cutting enzyme present within the multiple cloning site (*Bam*HI, *Sac*I, *Sal*I, or *Hind*III). In this case, modify the primer design and restriction procedures according to the new enzyme combination.
4. Some enzymes may exhibit star activity at high concentrations and long incubation times. Check information of manufacturer for the chosen enzymes. If star activity is an issue, reduced enzyme concentrations and incubation times may prevent its occurrence. However, this might require to separate restriction reactions for both enzymes involved.
5. The ATG start codon of the target coding region has to be part of the 6-bp recognition site of *Nde*I: CATATG.
6. Millipore water is a registered trademark for water purification systems manufactured by Millipore providing water with a conductance of 18.2 mΩ at 25°C. After sterilization (121°C, 2 bar, 20 min) store at 4°C.
7. We generally use *R. capsulatus* strains that base on strain B10S, a spontaneous streptomycin resistant mutant of *R. capsulatus* wild type strain B10 (18).
8. In contrast to *E. coli*, *R. capsulatus* does not survive storage at −80°C.
9. Illumination of *R. capsulatus* in the presence of oxygen leads to photooxidative processes that might cause cell death.
10. This is an important issue because *R. capsulatus* produces photopigments under phototrophic growth conditions that absorb light at the standard wavelength for measuring bacterial cell densities ($\lambda = 580 - 600$ nm), which affects correct determination of cell density.
11. Activation of the gas pack by addition of water, placing in the jar and closing of the container has to be carried out quickly to avoid exhaustion of the gas pack before the jar is closed. For the same reason, do not reopen the jar once closed, as the generated anaerobic environment will be lost.
12. Within the stationary growth phase, dead cells of the preculture are forming a dark layer at the bottom of the Hungate reaction tube. Thus, do not mix precultures after incubation. Use cell suspension from the upper phase of the culture for cell density measurements and inoculation.
13. In most cases, best expression efficiencies were observed when the inducer fructose was directly added to the medium. However, in some cases, it might be beneficial to allow cells to grow to the logarithmic growth phase prior induction of

heterologous gene expression. In this case, add fructose when the cultures have reached a cell density corresponding to an OD_{660} of 0.4–0.6.

14. Although maximal filling of cultivation vessels is generally sufficient to exclude oxygen from cultures, it might be advantageous to flush the cultures with argon or nitrogen to remove all traces of oxygen and to facilitate the sampling process.
15. When argon is not available, nitrogen may be used as an alternative gas. However, in contrast to nitrogen argon gas exhibits a higher density than air and will form a layer above the culture surface preventing contact with oxygen. Therefore, when using nitrogen gas, ensure that the containers are completely airtight as otherwise *R. capsulatus* will not grow.
16. Some heterologously expressed proteins may be subjected to proteolytic degradation, resulting in low protein yields when the cells are grown to stationary phase. In that case, higher protein yields may be achieved within the logarithmic growth phase ($OD_{660} \leq 1$).
17. Low-level expression may require aliquots with higher protein concentrations.
18. If protein concentration is low in the samples, higher sample volumes can be used.
19. When levels of heterologous proteins are low, decrease skim milk concentration in the TBST buffer to 1% or 0.75%.
20. Signal and background generally bleach after air-drying. As the background bleaches far more than the signal, it is recommended to stop the reaction when the signal is distinct, regardless of the background being stained as well.

References

1. Baneyx F (1999) Recombinant protein expression in *Escherichia coli*. *Curr. Opin. Biotechnol.* **10**, 411–421.
2. Terpe K (2006) Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.* **72**, 211–222.
3. Wong SL (1995) Advances in the use of *Bacillus subtilis* for the expression and secretion of heterologous proteins. *Curr. Opin. Biotechnol.* **6**, 517–522.
4. Weickert MJ, Doherty DH, Best EA, et al., (1996) Optimization of heterologous protein production in *Escherichia coli*. *Curr. Opin. Biotechnol.* **7**, 494–499.
5. Li W, Zhou X, Lu P (2004) Bottlenecks in the expression and secretion of heterologous proteins in *Bacillus subtilis*. *Res. Microbiol.* **155**, 605–610.
6. Cereghino JL, Cregg JM (2000) Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol. Rev.* **24**, 45–66.
7. Cregg JM, Cereghino JL, Shi J, et al., (2000) Recombinant protein expression in *Pichia pastoris*. *Mol. Biotechnol.* **16**, 23–52.
8. León-Banares R, González-Ballester D, Galván A, et al., (2004) Transgenic microalgae as green cell-factories. *Trends. Biotechnol.* **22**, 45–52.
9. De Schutter K, Lin Y-C, Tiels P, et al., (2009) Genome sequence of the recombinant protein

production host *Pichia pastoris*. *Nat. Biotechnol.* **27**, 561–566.

10. Nevalainen KM, Te'o VS, Bergquist PL (2005) Heterologous protein expression in filamentous fungi. *Trends Biotechnol.* **23**, 468–474.
11. Punt PJ, van Biezen N, Conesa A, et al., (2002) Filamentous fungi as cell factories for heterologous protein production. *Trends Biotechnol.* **20**, 200–206.
12. Katzke N, Arvani S, Bergmann R, et al., (2010) A novel T7 RNA polymerase dependent expression system for high-level protein production in the phototrophic bacterium *Rhodobacter capsulatus*. *Protein. Expr. Purif.* **69**(2), 137.
13. Strnad H, Lapidus A, Paces J, et al., (2010) Complete genome sequence of the photosynthetic purple nonsulfur bacterium *Rhodobacter capsulatus* SB 1003. *J. Bacteriol.* doi:10.1128/JB.00366-10.
14. De Smet L, Kostanjevecki V, Guisez Y, et al., (2001) A novel system for heterologous expression of flavocytochrome c in phototrophic bacteria using the *Allochromatium vinosum* rbcA promoter. *Arch. Microbiol.* **176**, 19–28.
15. Drepper T, Arvani S, Rosenau F, et al., (2005) High-level transcription of large gene regions: a novel T(7) RNA-polymerase-based system for expression of functional hydrogenases in the phototrophic bacterium *Rhodobacter capsulatus*. *Biochem. Soc. Trans.* **33**, 56–58.
16. Kappler U, McEwan AG (2002) A system for the heterologous expression of complex redox proteins in *Rhodobacter capsulatus*: Characterization of recombinant sulphite: cytochrome c oxidoreductase from *Starkeya novella*. *FEBS Lett.* **529**, 208–214.
17. Laible PD, Scott HN, Henry L, et al., (2004) Towards higher-throughput membrane protein production for structural genomics initiatives. *J. Struct. Funct. Genomics.* **5**, 167–172.
18. Klipp W, Masepohl B, Pühler A (1988) Identification and mapping of nitrogen fixation genes of *Rhodobacter capsulatus*: Duplication of a nifA-nifB region. *J. Bacteriol.* **170**, 693–699.
19. Simon R, Priefer U, Pühler A (1983) A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in Gram negative bacteria. *Biotechnology* **1**, 784–791.
20. Sambrook J, Russell DW (ed.) (2001) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Chapter 14

Plasmid DNA Production for Therapeutic Applications

Alvaro R. Lara and Octavio T. Ramírez

Abstract

Plasmid DNA (pDNA) is the base for promising DNA vaccines and gene therapies against many infectious, acquired, and genetic diseases, including HIV-AIDS, Ebola, Malaria, and different types of cancer, enteric pathogens, and influenza. Compared to conventional vaccines, DNA vaccines have many advantages such as high stability, not being infectious, focusing the immune response to only those antigens desired for immunization and long-term persistence of the vaccine protection. Especially in developing countries, where conventional effective vaccines are often unavailable or too expensive, there is a need for both new and improved vaccines. Therefore the demand of pDNA is expected to rise significantly in the near future. Since the injection of pDNA usually only leads to a weak immune response, several milligrams of DNA vaccine are necessary for immunization protection. Hence, there is a special interest to raise the product yield in order to reduce manufacturing costs. In this chapter, the different stages of plasmid DNA production are reviewed, from the vector design to downstream operation options. In particular, recent advances on cell engineering for improving plasmid DNA production are discussed.

Key words: DNA vaccines, Plasmid DNA, pDNA, Cell engineering, Minicircles

1. Introduction

Plasmids are extrachromosomal DNA present in archaea, bacteria, and some yeast with at least one origin of replication (*ori*), at which the plasmid replication starts independently of the chromosomal DNA. Plasmid DNA codes for additional genetic information such as resistances against antibiotics, heavy metals, or UV light. They also can code for enzymes participating in the catabolism of unusual substrates or can confer the possibility to form surface antigens, antibiotics, or toxins. Plasmid sequences like the fertility factor allow passing plasmids from one cell to another via conjugation. Hence, microorganisms receive advantages from the additional

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_14, © Springer Science+Business Media, LLC 2012

capacities encoded in the acquired plasmids. During the last 40 years, plasmids have played a crucial role in the development of biotechnology, being probably the molecular tool most widely used for DNA manipulation, transfer, and gene expression in a variety of microorganisms and animal cells (1). In the field of pharmaceutical biotechnology, plasmids are essential to produce heterologous proteins. Such proteins should substitute defective proteins present in the patient, or provide a lost function due to lack of the natural active protein. An alternative to treat diseases related to defective production of a protein, or missing genetic information, is the administration of the gene of interest to the patient, a strategy that is known as gene therapy or genetic medicine. The underlying concepts of gene therapy date back to the 1960s. In the early 1990s the transfer of genes to humans was reported (2). Since then, gene therapy and genetic vaccination have attracted much attention.

Gene therapy and DNA vaccination require the identification of the gene or genes related to a particular disease (inherited or acquired), the fabrication of a therapeutic gene, the design of a molecular vector (and its formulation), and the introduction of the gene into the patient (Fig. 1). When the gene is expressed in the patient, the correct protein is expected to be formed and function. Hence, problems related to recombinant protein production, such as complex glycosylation, are eliminated. There exist several possible vectors to introduce the genetic information into human cells. The most relevant are virus (adenovirus or retrovirus) and plasmid DNA (pDNA), which can be used in aqueous solution (naked) or included in lipids or other formulations.

Until now, over 1,500 clinical trials of human gene therapy with more than 220 genes in almost 30 countries have been carried out since the first gene therapy trial was conducted (3). More than 60% of the trials have been performed in the USA and almost 30% in Europe. As shown in Table 1, pDNA plays a very important role as a vector for gene therapy and DNA vaccination. Almost 20% of the trials for human gene therapy have been based on naked pDNA, whereas lipofection (which also requires pDNA production) counts for 6.6% of the trials. Together, both approaches represent nearly 25% of the techniques used in clinical trials.

The highly efficient transfection machinery of viruses makes them a very important tool for gene therapy. Indeed, they represent about 50% of the vectors used in current trials (Table 1). Although expression of the therapeutic gene following transfection with viral vectors is significantly higher than that using pDNA, the direct use of plasmid has important advantages over viruses. For instance, compared to retrovirus vectors, pDNA presents decreased risk of integration of the foreign DNA into the human chromosomes (4).

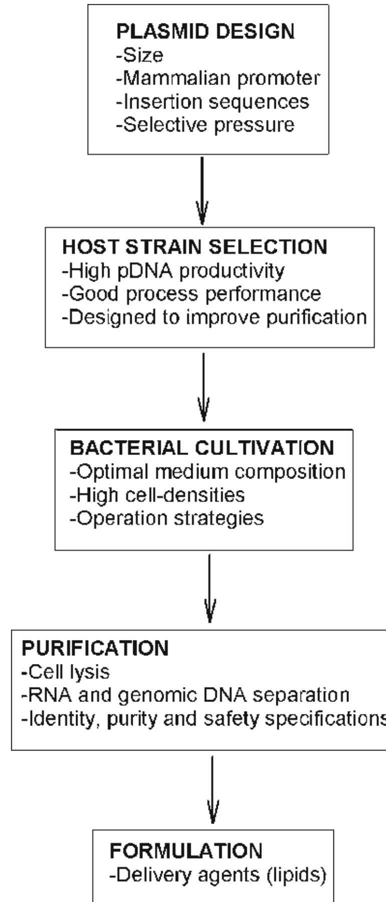


Fig. 1. Overview of the main steps for pDNA vaccine design and production.

Table 1
Vectors used for gene therapy trials worldwide

Gene delivery vector	Number of trials	% of trials
Adenovirus	392	23.8
Retrovirus	341	20.7
Naked/plasmid DNA	301	18.3
Lipofection	109	6.6
Vaccinia virus	102	6.2
Others	399	24.4

Data source: The Journal of Gene Medicine (John Wiley & Sons) (3)

Isolated cases of safety issues were reported for early clinical trials based on retroviral vectors, such as severe immune and inflammatory response that led to the death of one patient (5). Another advantage of pDNA is that the gene size is not limited to the size of the viral capsid. The immune response of pDNA when used for vaccination is usually lower than that of viral vectors. Hence, large amounts of pDNA are necessary for a successful immunization. However, such drawback can be overcome by using different adjuvants (6, 7) and/or gene delivery methods such as needle-free jet injection, electroporation/electro pulses or a gene gun (8–10). Since pDNA is produced in bacterial cultures, all the accumulated knowledge with such systems can be readily applied to manufacture processes. In comparison to animal cell cultivations, needed for viral vectors production, microbial processes are faster, easier and more information exists on scaling-up issues.

The Food and Drug Administration of the USA (FDA) defines DNA vaccines as "purified plasmid preparations containing one or more DNA sequences capable of inducing and/or promoting an immune response against a pathogen" (11). For a plasmid DNA vaccine, the target DNA sequence is cloned into a plasmid under the control of a eukaryotic promoter. After purification of the plasmid DNA, the final vaccine is injected into the recipient, where the target DNA sequence is expressed by the eukaryotic promoter. The recipient should then generate a symptom-free immune response to the plasmid-encoded immunogen.

In current clinical trials based on pDNA, the immunization against or the treatment of some severe illnesses, especially of cancer and cardiovascular diseases, is investigated, including malignant melanoma, pancreatic carcinoma, renal cell carcinoma, critical limb ischemia, and coronary heart disease. In addition, pDNA as a vector also holds great potential against infectious, monogenetic, and neurological diseases such as HIV (1), hepatitis B, influenza, Huntington's disease, hemophilia, multiple sclerosis, and diabetic neuropathy (Table 2). In 2005, the first DNA vaccine was released for immunization of horses against the West Nile Virus (12). Today, in the veterinary field, five medical applications (four of them for vaccination purposes) based on pDNA have already been approved for clinical use (13).

The discussed information allows estimating a growing demand for pDNA in the near future. In this chapter, details about pDNA vectors construction, bacterial strain selection, and modification will be analyzed, with an emphasis on the interaction of host–vector, bacterial physiology, and culture methods. Finally, developments in pDNA purification and delivery will be shortly reviewed.

Table 2
Clinical trials for pDNA-based gene therapy for selected diseases

Disease type	Examples of applications
Cancer:	Malignant melanoma, pancreatic carcinoma, renal cell carcinoma, head-and neck-squamous cell carcinoma, metastatic carcinoma, breast cancer, lymphoma, leukemia, prostate carcinomas, lung cancer, multiple myeloma, malignant glioma, colorectal cancer
Cardiovascular diseases:	(Critical) limb ischemia, coronary heart disease, peripheral artery occlusive disease, angina
Infectious diseases:	HIV(1), hepatitis B, tetanus, influenza
Monogenetic diseases:	Huntington's disease, hemophilia
Neurological diseases:	Multiple sclerosis, diabetic neuropathy
Others:	Retinitis pigmentosa (progressive retinal dystrophy), type I diabetes, erectile dysfunction, hip fracture, pulmonary hypertension

Data source: The Journal of Gene Medicine (John Wiley & Sons) (3)

2. Considerations for the Design of Plasmids for Gene Therapy and Vaccination

2.1. Basic Concepts on Plasmid Replication

After a target gene has been defined, it is inserted into a plasmid (Fig. 1). The plasmid is then used as vector to introduce the genetic information into the patient's cells. Plasmids are autonomously replicating, circular DNA that often code additional (metabolic) capabilities in some unicellular organisms, such as resistance to antibiotics, heavy metals, and toxic anions. Plasmids for therapeutic applications are commonly produced in bacterial cultivations; therefore, it is necessary to preserve all the needed elements for their replication in the microorganism. A constant concentration of intact plasmids per cell is maintained actively by the plasmids themselves, by means of a plasmid-encoded repressor (negative active control). Replication control takes place at the level of the initiation of the leading strand synthesis by auto-repression: An unstable repressor protein (Cop), which is constantly expressed from the plasmid, inhibits the formation of a part of the initiation complex (such as initiator proteins -Rep- or primers) in case the plasmid concentration is too high (14).

The inhibitor dilution model describes that at a certain plasmid concentration equilibrium is established between the degradation and synthesis of the Cop protein. Since the Cop protein is unstable, its concentration rapidly decreases in case the plasmid concentration decreases, e.g., when the cell grows. As a result the Rep protein expression is no longer inhibited and plasmid replication is reestablished. The increasing plasmid number again leads to a

higher Cop protein concentration that upon reaching a critical concentration will again inhibit plasmid replication (14).

The most commonly used plasmids for gene therapy and vaccination are derived from ColE1 plasmid from *Escherichia coli*. Initially, the selection of such plasmid backbones was based on the positive results for recombinant protein production (15). The origin of replication of ColE1 plasmids confers a relatively low copy number (20–40 copies per cell). Regulation of ColE1 plasmid involves several molecules coded by about 600 bases of its own sequence. Plasmid ColE1 replicates via the theta mechanism (name derived from the structural similarity to the Greek letter θ of the two DNA strands during the replication). Bidirectional replication begins at the site *oriV* (16), and is not dependent on a plasmid-encoded Rep protein. The primer that leads strand replication is formed by RNAII, a transcript from the plasmid. ColE1-type plasmid replication is initiated by stable, host-encoded proteins, which allow plasmid replication to continue for a significant time even without synthesizing new proteins (unlike other plasmids which need permanent synthesis of the unstable Rep protein). The primer for plasmid replication, RNAII, needs to fold in its active form and bind to its complementary DNA strand in order to initiate plasmid replication. Here replication control takes place by the molecule RNAI, which binds to RNAII and inhibits firstly the correct folding of RNAII and consequently the binding of RNAII to the complementary DNA strand (16). A further part of the replication control system is the RNA one modulator (Rom) protein, which increases the binding speed of RNAI. RNAI is unstable with a half-value period ($t_{1/2}$) of 2 min and hence is synthesized very fast (14).

Replication of ColE1-type plasmids normally requires RNase H for cutting out the DNA primer. The elongation of the new DNA strand is carried out by DNA polymerase I by connecting complementary nucleotides with a transcription rate of about 60 nucleotides per second at 37°C (14). Additionally, there is an alternative way for replication without the need of RNase H and DNA polymerase I: RNAII hybridizes to the DNA and then displaces the non-transcript strand while extending. Further on, a helicase unwinds the plasmid DNA (14, 17, 18).

A preferred origin of replication, derived from ColE1 is present in the pUC family, in which the *ori* gene has a mutation in a single base, affecting RNA sequences that control the replication of the plasmid (19). Additionally, the pUC plasmids do not contain the *rop* gene, which exerts control mechanisms to keep the copy number low. With such modifications, pUC plasmids can reach copy numbers of more than 500 copies per cell (20). An important exception to the ColE1 origin of replication for pharmaceutical use is the design reported by Soubrier et al. (21), who reported the use of a conditional origin of replication that also requires molecular modifications in the host. This particular example will be discussed below.

2.2. Selectable Markers

As the plasmid is inserted into a bacterium, a means to selectively cultivate plasmid-bearing population is needed. Traditionally, the selectable marker for this purpose has been a gene coded in the plasmid that confers resistance to an antibiotic. Although β-lactamase is a commonly found marker in a variety of plasmids (providing resistance to bacterial lytic penicillin-based antibiotics), the use of this antibiotic is undesirable in technical cultivations, since the enzyme degrades the antibiotic and hence the effectiveness of the marker decreases with cultivation time. The presence of penicillin-based antibiotics is also undesirable, since this antibiotics are used in humans to treat a variety of infections, and if the antibiotic is not completely removed during plasmid purification, antibiotic resistance could be favored and in other cases, allergic reactions could occur (22). This is also the case of bacteriostatic resistance marker such as *tet*^{R} (20). The tetracycline selection marker can also be toxic to *E. coli* (the most common host used for pDNA production) under certain conditions (23).

For plasmids used in gene therapy and vaccination, the preferred antibiotic resistance marker is the TN903 gene, coding for an aminoglycoside enzyme that confers resistance to kanamycin. As kanamycin is not widely used in humans, a lower risk of allergic responses than other antibiotics is expected. However, in general, the introduction of antibiotic resistance into the plasmids is an undesirable characteristic. Current developments on the design of safer resistance marker are discussed below.

2.3. Elements for Eukaryotic Expression

Once the plasmid has reached the nucleus of the target cell, the therapeutic gene should be expressed. Therefore, the plasmid must contain the necessary elements for expression in human cells. The therapeutic gene sequence should be optimized in terms of the codons preferably used by the host organism. A proper promoter should be used to express the therapeutic gene, being the viral promoters CMV, RSV, and SV40 the most widely used (20). However, these promoters may not be the best option in a near future. For instance, it has been demonstrated that the use of the promoters from the human polyubiquitin C (UbC) and the elongation factor 1α (EF1α) genes result in persistent gene expression in the mouse lung, lasting up to four times more than using viral promoters (24). In the case where more than one therapeutic gene is included in a vector, bidirectional promoters can be introduced. For a review, see Müller et al. (25).

The plasmids taken into clinical trials also contain polyadenylation/termination sequences coupled to termination and processing of the therapeutic gene. These sequences are usually derived from the rabbit β-globin or bovine growth hormone genes (13, 20). The use of synthetic or SV40 signals increase the half-life of plasmids in cell lysates (26), which is desirable for extending the potency of a plasmid-based drug. Recent advances also include

extra sequences to improve the efficacy of the therapeutic gene, such as antigen-targeting sequences (promoting protein transport to the extracellular space) (27). It is assumed that no extra-chromosomal elements for expression of the therapeutic gene are necessary and thus no extra information is contained in the plasmid. Common plasmid backbones used for gene therapy and vaccination include (but are not limited to) pVAC, pDNAVACultra, pcDNA, and pCMV (28).

2.4. Structural Characteristics of Plasmids

The structural characteristics of plasmids are important from a therapeutic point of view, since stability and efficacy depend in part of the topology of the plasmid. The topology of the produced plasmid is also dependent on several physiological characteristics of the production host, which will be reviewed in Subheading 3 of this review chapter.

Double-stranded DNA can exist in several conformations, know as A, B, and Z, depending on the hydration level, sequence, chemical environment, and coiling (29, 30). The A-DNA form is a wider right-handed spiral, with a shallow, wide minor groove and a narrower, deeper major groove, compared to B-DNA. If there are methylated segments, a major change in conformation can occur, in which the strands turn left-handed around the helical axis, opposite to the spiral direction of the B-DNA conformation. This form, named Z-DNA can be recognized by specific binding proteins and is probably involved in transcriptional regulation (31). From the described forms, B and Z are the most commonly found in the cell (32).

Due to the circular nature of plasmids, other structures such as knots, and links with other plasmid molecules (known as catenanes), can be present (33). Concatamers can also exist, consisting in multimeric concatenations of several plasmid molecules, formed normally after replication by homologous recombination (14). Therefore, the formation of multimers depends on the sequence of the plasmid.

Most of the plasmids produced by bacteria are packed in a very compact form. This form is known as form I or closed circular, since each DNA strand is closed on itself, thus linking the two strands of the DNA double helix. The breakage of one DNA strand by nucleases or mechanical stress results in an open circular form by loss of the molecular coiling (34).

The *linking number* (Lk) describes the number of intersections between one strand and a spanning surface, defined by the other strand. By convention, Lk is positive for right-handed double helix (33). The value of Lk (which should be an integer, positive or negative) is usually very close to the number of base pairs in the molecule (N) divided by the number of base pairs per double-helix turn in linear DNA under given conditions (γ), a factor also known

as Lk_0 ($Lk_0 = \gamma/N$) (33). The properties of closed circular pDNA can be better described by the linking number difference (ΔLk), defined as:

$$\Delta Lk = Lk - \frac{\gamma}{N}. \quad (1)$$

If $\Delta Lk \neq 0$, closed circular pDNA is assumed to be supercoiled. This name arose from the fact that the stress resulting from the supercoiled condition leads to a spatial deformation of the helix axis, forming a helix of higher order, also known as "superhelicity" (33).

Alternatively, the value of superhelical density (σ) is also used to describe the topology of the plasmid. This is a normalized value of ΔLk, and is defined by (2):

$$\sigma = \frac{\Delta Lk}{Lk_0} = \frac{\gamma \cdot \Delta Lk}{N}. \quad (2)$$

Supercoiled pDNA has the advantage of displaying a reduced size. Supercoiling is an active, energy consuming process controlled by DNA gyrase, a type II topoisomerase (35). Energy is supplied by ATP hydrolysis. Intracellular relaxation of supercoiled pDNA can also occur, mediated by topoisomerase I, without the need for ATP. With both enzymes acting in an antagonistic way, a constant level of pDNA supercoiling is usually maintained. Except for archebacteria, the majority of pDNA supercoiling is negative, i.e., the sense of rotation is against that of the double helix (14), with values of σ between −0.03 and −0.09, usually near to −0.06 (36).

Current efforts tend to the production of the supercoiled form of pDNA, since a content of at least 80% of supercoiled pDNA in the final product has been recommended (11). It is believed that the supercoiled form is better protected against enzymatic degradation and thus more plasmid can reach the nucleus of the target cell. Several studies have shown that supercoiled pDNA is more efficient for vaccination (for instance refs. 37, 38). Other authors have shown that concatamers can elicit a higher immune response compared to monomers (39). This can be due to the several copies of the therapeutic gene contained in the concatamers. Concatamers can also show supercoiled and open circular forms (34). The analysis and quantification of pDNA is crucial for process development and product regulation. The next section addresses some issues on this regard.

2.4.1. Plasmid Analysis and Quantification

A traditional method to quantify pDNA is by measuring light absorbance at 260 nm. At this wavelength, there is a peak of maximum absorption of light of the cyclic nucleic acids bases. This method requires only a UV spectrophotometer. In most cases, with the use of standard cuvettes, there is no need for constructing a calibration curve, and a constant ratio of DNA concentration per absorbance

unit is assumed. Although straightforward and fast, absorbance data can be deviated by interference with proteins and RNA. In principle, all the nucleic acid bases will absorb light, regardless of the content of DNA or RNA, or whether the DNA is single or double stranded, supercoiled, open circular, or linear. Therefore, no information about the topology of the plasmid can be obtained, and highly pure samples are required for robust quantification.

The use of intercalating, fluorescent molecules has also been reported for quantifying pDNA. For instance, PicoGreen is an ultrasensitive fluorescent nucleic acid stain that intercalates specifically between the double stranded DNA (40, 41). When measured with fluorescence, DNA concentrations of the order of ng/mL can be determined. RNA is not detected with PicoGreen. It has been demonstrated that single stranded DNA does not interfere during the quantification of pDNA (42). However, pDNA topology causes different fluorescence of the dye. For instance, in Fig. 2 it is shown that linear fragments of DNA (lambda phage DNA digested with the enzyme *Hind*III), supercoiled pDNA and linearized pDNA yield very different slopes in calibration curves. This issue should be taken into account when using such method to quantify pDNA. This fact can, however, be used to monitor DNA damaged during production processes (43) and to detect the presence of topoisomers (44).

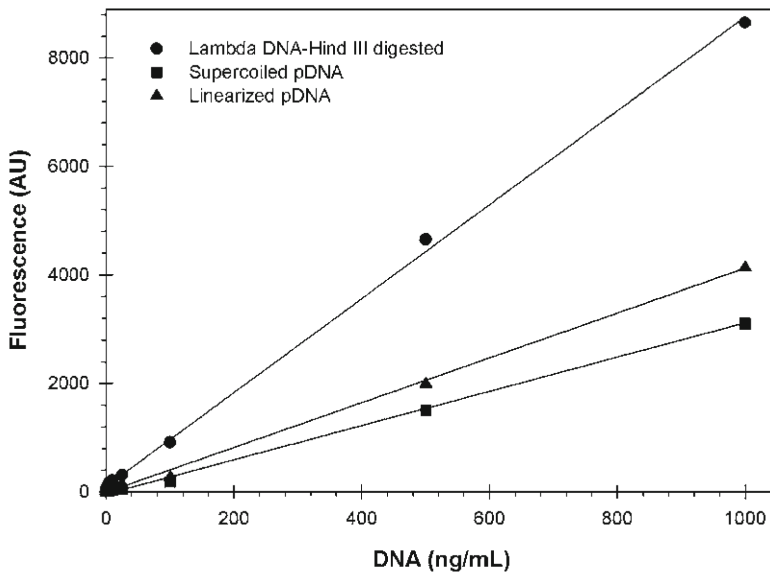


Fig. 2. Visualization of pDNA topoisomers in agarose gels. (**a**) A 0.8% agarose gel stained with ethidium bromide. Numbers at left correspond to the size of the bands (bp) in *lane 1*, where a linear DNA mass ladder was loaded. In *lane 2*, three bands of a 6-kbp plasmid can be seen. The faster isoform corresponds to supercoiled DNA, which migrated more than the linear DNA of the same size. (**b**) Lane 2 was image analyzed and two different areas can be detected. *Area 1* corresponds to isoforms with low electrophoretic mobility, while *area 2* corresponds to supercoiled pDNA, whose proportion in the sample can be calculated by integration of the two areas.

Detailed studies of plasmid topology can be carried out using advanced microscopy techniques (for instance ref. 45). This is, however, a technique that could rarely be implemented for monitoring industrial production. There are faster and simpler techniques that have been traditionally used in laboratory, such as the agarose gel electrophoresis. For this method, pDNA is separated under an electric field force and the different topological (different ΔLk) forms migrate with different mobilities in the polymer matrix. For instance, supercoiled pDNA migrates faster than the linear and open circular forms (which correspond to $\Delta Lk = 0$). The gel is then stained with an intercalating molecule, such as ethidium bromide. Upon exposure to UV radiation, the bands corresponding to the different pDNA forms can be visualized and identified (in principle) when compared to a linear or supercoiled ladder run in the same gel. This method has the advantage that several dozens of samples can be analyzed in the same gel. The gel can be further analyzed by image analysis methods. A simple example is seen in Fig. 3.

Image analysis can also be useful to quantify pDNA, for example, by comparing the intensity of color and size of the bands of a sample to those of the standard (provided that a quantitative standard is used). This method is usually very precise. However, it should be taken into account that the different pDNA topoisomers can absorb the intercalating and UV radiation at different intensities, which may yield wrong results (Fig. 3). The mobility of each topoisomer can vary with several factors, such as gel running conditions and the presence or absence of the intercalating agent during the electrophoresis, which may lead to mistakes during the identification of topoisomers. It has been suggested that untreated, UV irradiated and linear samples should be run in parallel for a precise identification of the topoisomers (46). Capillary gel electrophoresis allows a better

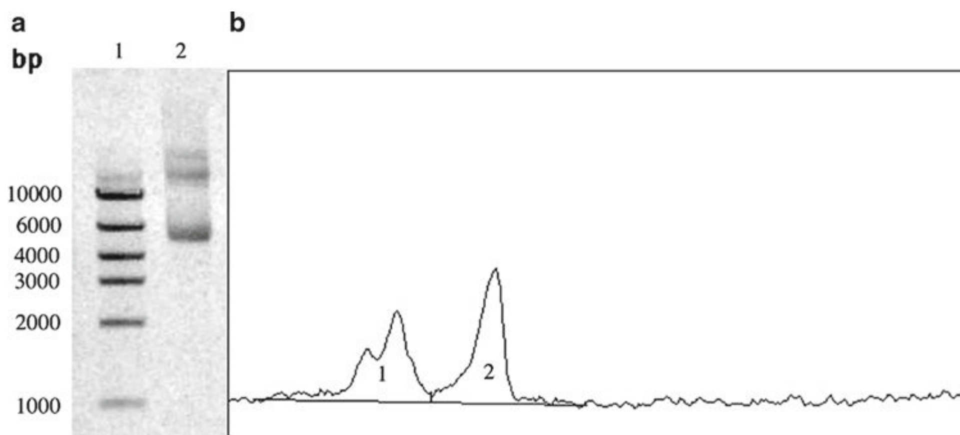


Fig. 3. Influence of the DNA form on the fluorescence signal using PicoGreen.

separation and quantification of plasmid structure (46), although up to now it cannot be applied for several samples at a time, which may limit its use for high-throughput research and development. Topoisomer mobility is determined by the absolute value of ΔLk, hence, a single band can correspond to topoisomers with positive and negative values of ΔLk. If a more detailed resolution of the topoisomers is needed, a second dimension electrophoresis of the sample can be run. This technique can also be useful to distinguish between Z-DNA and B-DNA (47–49).

Chromatographic methods have also been developed for monitoring, quantification, separation, and analysis of pDNA (50–54). A common method is anion exchange high performance liquid chromatography (HPLC), which allows the quantification of pDNA in the presence of impurities such as RNA or genomic DNA. Chromatographic methods present the advantage of having been scaled to several industrial biotechnology applications. Other characteristics like sequence, restriction patterns, radius of gyration, size, and shape can be monitored for industrial production using Raman spectroscopy, infrared spectroscopy, and light scattering (55, 56).

2.5. Engineering of Plasmid Vectors

2.5.1. Eliminating the Use of Antibiotic Resistance Genes

There exist several drawbacks in the basic plasmid characteristics described in previous sections. Due to the fact that plasmid backbones have been initially taken from bacterial plasmids, several bacterial elements remain in the plasmid. Such sections of pDNA are useless for therapeutic applications and are undesirable, as they can cause activation of cryptic expression signals (11). Therefore, a logical strategy to improve vector design is to eliminate those sections of the plasmid that are not necessary for the therapeutic application, but that will not affect its replication in the bacterial cell. This has the additional advantage of reducing the size of the plasmid, which can result in higher amounts of pDNA produced by bacterial cultivations. A shorter plasmid size is also desirable due to recent technologies that include larger sequences, such as genomic loci (57–59) or more than one therapeutic gene in the same vector.

There exists the possibility of dissemination of selectable marker genes, such as those that confer antibiotic resistance (60, 61) or other undesirable consequences. For instance, it has been shown that the expression of plasmid-coded neomycin resistance affected the expression of the therapeutic gene and caused negative metabolic effects (62). The expression of antibiotic resistance genes is generally strong, imposing a metabolic burden to the cell that impacts the metabolic fluxes and results in a general reduction of growth rate and biomass yields on glucose (63–65). It has been demonstrated that the expression of the kanamycin resistance gene encoded in a pCDNA3 (Invitrogen) derived plasmid was so strong that inclusion bodies were accumulated (66). A general approach to eliminate the use of antibiotic resistance genes is the construction of mutant

E. coli strains lacking a specific metabolic activity (auxotrophy). The metabolic activity is then restored by some elements present in the plasmid and the native phenotype is recovered. Hence, under the cultivation conditions, only those cells bearing the plasmid will grow and extra selective pressure is not needed. For example, deletion of *dap* gene resulted in *E. coli* cells with lysine auxotrophy. A copy of the *dap* gene was integrated into the chromosome, under transcriptional control of the *lac* promoter (67, 68). A high copy number plasmid containing the *lac* operator, relieves the repression of the *lac* promoter in the chromosome (69) and allows the expression of *dap*. Therefore, only cells containing the plasmid can successfully grow in minimal medium. This approach has also been applied to the production of a HIV DNA vaccine candidate (70). Another group reported the complementation of a *glyA* mutant strain (unable to synthesize glycine) with the insertion of *glyA* in the plasmid (71). The deletion of *infA* gene, coding for an essential translation initiation factor, and its complementation with a copy in the plasmid has also been reported (72). Thymidine auxotrophy has also been used for eliminating the need for antibiotic resistance in a cholera vaccine (73). Marie and coworkers used a thymidine auxotrophic *E. coli* strain and introduced an antibiotic marker-free plasmid containing the *thyA*, restoring the capacity of growth in minimal medium (74). The plasmid contained a tRNA amber suppressor gene. They found that the expression level of luciferase (reporter gene) was higher and more prolonged than when the original plasmid was used for transfection. An interesting design was presented by Dong and coworkers (75), who knocked out the chromosomal quinolinic acid phosphoribosyltransferase (QAPRTase), which resulted in a defective de novo biosynthesis of NAD in the *E. coli* mutant strain. The QAPRTase gene of *E. coli* or mouse was inserted into a model plasmid, and the antibiotic resistance gene was removed. Growth was restored in minimal medium, although the strain expressing the QAPRTase gene from *E. coli* accumulated more biomass and grew faster than that containing the same gene but from mammalian origin.

A more specific therapeutic-oriented plasmid design contains a conditional origin of replication (*ori* γ), which requires the expression of an initiator protein (π protein, encoded by the *pir* gene) produced by the host (76). With this system, named pCOR, the plasmid is only replicated by π-expressing bacteria and thus the host range and the possibility of plasmid dissemination is limited. The plasmid pCOR also contains a gene coding for a tRNA selectable marker that suppresses an *arg* gene mutation (arginine auxotrophy), which allows the selection of plasmid-bearing cells without the need for antibiotics. This system allowed higher luciferase activity levels (reporter gene) than the unmodified plasmid. Additional mutagenesis of *pir* resulted in a fourfold increase of pCOR production level in small-scale bioreactors (76).

Another option to eliminate the use of antibiotic resistance genes is the use of antisense RNA. An example is the system developed by Grabherr and coworkers (77, 78). The *murA* essential gene (coding for UDP-*N*-acetylglucosamine enolpyruvyl transferase) was placed under the transcriptional control of the pL*tet*O operator, so that its expression was hampered by the *tet* repressor (*tetR*) of *E. coli*. This repressor was modified to contain a sequence complimentary to RNA I. With this design, cells containing a plasmid with the ColE1 origin of replication and the antisense RNA sequence deactivated *tetR*. Consequently, the expression of *murA* was de-repressed and growth was restored only in cells bearing the plasmid. A more general system (not depending on ColE1-type plasmids) was proposed by Luke and coworkers (79), who inserted a sequence for 150 bp RNA-OUT antisense RNA in different plasmids. The levansucrase gene (*sacB*) from *Bacillus subtilis* was inserted in *E. coli*, where it was expressed constitutively, with the result that the mutant strain was killed in the presence of sucrose. When the antisense RNA coded in the plasmid was present, it repressed the *sacB* gene and growth was restored.

The examples explained above show that new and more sophisticated approaches are emerging to avoid the use of antibiotics for DNA vaccine production. Advances on the size reduction of plasmids and other considerations will be discussed below.

2.5.2. Minimizing the Plasmid Size

Reducing the size of the plasmid is an important issue for the success of pDNA-based drugs. A shorter plasmid would reduce the metabolic burden on the host (63), improve transfection efficiency (80) and would even be easier to purify, especially at large scale. Efforts to eliminate antibiotic resistance genes have been discussed, but other issues arise for pDNA-based drugs to reach safer (as explained below) and broader applications. The current approach is to eliminate all prokaryotic sequences of the plasmid and use only those that are necessary for its replication and expression in the patient. The designs of minimalistic plasmids include the MIDGE and the minicircle plasmids.

The "Minimalistic Immunogenically Defined Gene Expression" (MIDGE) concepts aim at the design of vectors of superior safety, strongly reduced size, and enhanced transfection and expression efficiency. MIDGE-vectors are plasmid-derived, linear double stranded covalently closed DNA molecules, with stem-loop structures that protect against exonuclease degradation and provide specific sites for targeting signals (81, 82). MIDGE vectors can be as small as 1,200 bp, whereas common plasmid size is well above 4,000 bp. Although molecular excisions and ligations are needed for obtaining the MIDGE from pDNA, Junhghaus and coworkers argue that the elevated transfection and expression efficiency balance the additional process cost at pilot level compared to the production of conventional pDNA (81). Some reported

applications of MIDGE vectors include experimental vaccines against *Leishmania* infection in mice (83), lymphoblastic leukemia (84), and hepatitis B (85).

Minicircles are supercoiled DNA that lack a bacterial origin of replication and antibiotic resistance marker. To produce the minicircles, pDNA is produced in *E. coli*, and should contain attP and attB specific sites. The plasmid is excised and ligated by att specific-recombination mediated by the phage λ integrase. Two products are obtained: the minicircle DNA, containing the therapeutic gene and eukaryotic promoter, and the miniplasmid, which contains the undesired prokaryotic sequences (86, 87). The size of the minicircle can be a half of that of the original plasmid. The minicircle can be separated from the miniplasmid using cesium chloride density gradients. However, undesirable dimeric, concatemeric and relaxed minicircles are frequently copurified along with the targeted supercoiled minicircles (88). The yield of the recombined products is low, around 60% of the starting material (86, 87). An alternative way to produce minicircles is using a Cre recombinase (89). Another technique utilizes phage ϕ31 integrase to catalyze recombination and produce minicircles, not requiring digestion enzymes. The resulting product can be purified on commercially available affinity columns, which is an advantage over cesium chloride gradients. Expression of the therapeutic gene using minicircles is 5–50 times higher than with the original or undigested plasmid (86, 87, 90, 91).

There is a current concern about the sequences within the plasmid backbone that may cause undesirable effects during pDNA production or administration. For instance, repetitive DNA sequences in the plasmid can cause mutations and recombination (90). Unmethylated bacterial CG dinucleotides (CpGs) could play a role in the decrease of gene expression and inflammation observed during the application of pDNA in lungs; effects that have been reduced when such sequences were eliminated form pDNA (91).

Although inflammatory response may be desirable for vaccination, it is not a desired effect for gene therapy protocols (61). Therefore, reducing the presence of CpGs in pDNA is a current trend. It has been reported that insertion of sequences IS1 in pDNA can occur during production cultures, affecting the productivity of the cells (92). A review on the consequences of plasmid sequences on stability and safety has been recently published (93).

3. Host Characteristics

The usual host for pDNA production is the gram negative bacterium *E. coli*. Ample knowledge has been accumulated on basic physiology, culture medium design, cultivation techniques, genetic

Table 3
Common *Escherichia coli* strains used for pDNA production

Strain	Genotype
DH1	F⁻ *endA1 hsdR17*(r_k^- m_k^+) *supE44 thi-1 recA1 gyr96* (*Nal*^{r}) *relA1*
DH5	F⁻ *deoR endA1 hsdR17*(r_k^- m_k^+) *supE44 thi-1 recA1 gyr96* (*Nal*^{r}) *relA1*
DH5α	F⁻ *endA1 hsdR17*(r_k^- m_k^+) *supE44 thi-1 recA1 gyr96* (*Nal*^{r}) *relA1* Δ(*lacZYA-argF*) *U196* Φ80*lacZ*Δ*M15*
JM105	F⁻ Δ(*lac-proAB*) *lacI*^{q} *thi repsL endA1slcB15 hadR4 traD36 proAB* Δ(*ZM15*)
BL21	F⁻ *ompT hsdS*$_B$ ($r_B^-m_B^-$) *gal dcm*
W3110	F⁻ λ⁻ *rph-1 INV* (*rrnD, rrnE*)

engineering, and industrial bioprocessing with this microorganism. From the reports published up to now, the most common *E. coli* strains for pDNA production are listed on Table 3, including their genotypes. Several of the surveyed reports come from industrial practitioners.

The most popular strain for pDNA production is DH5α. This may be due to its widespread use for pDNA replication in different laboratory-scale protocols. In the next section, the influence of the genotype on the amount and characteristics of the produced pDNA will be reviewed.

The first three *E. coli* strains described in Table 3 (by far the preferred hosts for pDNA production) share several genotypic characteristics. For instance, endonuclease I (coded by *endA*) is deleted to reduce the degradation of double stranded DNA. Endonuclease I is a magnesium-dependent periplasmic enzyme with a size of 12 kD. Although its function is not completely understood, it was found to cleave foreign double stranded DNA and cause degradation of plasmid DNA upon cell lysis (94). A deactivated *endA* gene is designated *endA1* and produces an inactive form of endonuclease I, improving pDNA stability (94, 95).

The gene *recA* produces a DNA strand exchange and recombination protein with a variety of functions, such as ATP-dependent recombination by catalyzing DNA strand exchange and insertions (96). Furthermore, it has a regulatory function involved in repairing DNA damages. *E. coli* strains lacking *recA* have shown increased plasmid segregational stability (97). A mutation in the gene for the subunit A of DNA gyrase (topoisomerase II), designated *gyrA96*, provides resistance against the gyrase inhibitor nalidixic acid formerly used as an antibiotic. This mutation is included in order

to increase pDNA supercoiling. Other authors have used *E. coli* JM105 to select clones with high production of multimeric pDNA, which may be advantageous for therapeutic use, as the multimers contain several copies of the therapeutic gene (98).

Table 3 shows that several other mutations have been incorporated in the described strains. The highly mutated strains tend to be unstable, i.e., they have a relatively high mutation rate leading to changes of the chromosomal as well as the pDNA (15). This can be a problem for safety, quality, and yield in regard to the production of pDNA as a drug. For example, the *E. coli* strain DH10β was found to have a 13.5-fold higher mutation rate than its wild-type MG1655, which is mainly attributed to higher rates of insertion sequence transposition (15). Since most of the studies regarding pDNA production focus on a few strains, there is still room for significant improvement in this field. Hence a re-evaluation of *E. coli* strains is necessary for optimized pDNA formation (15, 99, 100). For instance, Yau et al. (99) investigated the influence of the strain/plasmid combination on pDNA formation with eight commercial and nine non-commercial *E. coli* strains and three different plasmids of 5.8, 6.9, and 20 kb size. The strains were cultivated in shake flasks with complex medium. It was shown that plasmid formation depends to a huge extent on the strain and the strain/plasmid combination, wherein commercial strains were not necessarily better pDNA producers than non-commercial ones. In comparison to the other 16 strains, W3110 (a wild type strain, Table 3) showed a superior plasmid DNA yield and good growth rate when hosting a large plasmid (20 kb), and moderate values when hosting smaller ones (5.8 and 6.9 kb). In another study, Singer et al. (100) investigated 13 strains, most of them commercial, for their ability to produce supercoiled pDNA in shake flasks and fed-batch culture in defined medium containing glycerol or glucose. In contrast with Yau et al. (99), in this study the wild type MG1655 (wild type K-12 derived strain) appeared as a weak pDNA producer. Singer et al. (100) found that (with few exceptions) glycerol as the energy and carbon source lead to higher pDNA yields than glucose, but lower specific growth rates. Strain BL21 (commercial strain by Strategene), designed for the expression of non-toxic proteins, is noteworthy, as it has been shown to be a superior pDNA producer (95). This strain has the advantage of accumulating low amounts of acetate, which is a major issue for *E. coli* cultivations, as will be discussed later.

In general, no correlation between genotype and pDNA productivity or supercoiling degree has been found, as shown by the above mentioned studies. It should be noted that the discussed information is based on small-scale cultivations. Large-scale conditions could favor or affect the performance of one strain over others; however, information at this level is very scarce.

4. Cultivation Processes

Cultivation processes are a key step for the production of any biomolecule. In the case of pDNA, the well-established techniques for bacterial cultivations have been adopted. An important approach to maximize the productivity of a cultivation process is to achieve high cell densities, for example, to attain biomass concentrations higher than 30 g/L (measured as dry cell weight) in the bioreactor. Another important factor to improve the productivity is to increase the amount of pDNA that can be produced per cell (specific productivity). This parameter, known as specific yield or selectivity (some authors also count the number of plasmid molecules per cell, and report it as the *copy number*), is strongly dependent on the cultivation conditions, such as growth rate, and carbon and nitrogen source among others. The effect of some of these variables is discussed below.

4.1. Effect of Growth Rate on pDNA Production

The influence of the specific growth rate on pDNA formation has been actively studied. Understanding the effect of growth rate on specific pDNA yield is of particular importance for the design of cultivation strategies. Though there have been reports indicating a positive correlation or an optimal growth rate for pDNA formation, in general, low specific growth rates are considered to lead to higher pDNA synthesis.

4.1.1. Positive Correlation and Optimal Growth Rate

Reinikainen and Virkajärvi (101) studied the effect of dilution rate in chemostats using a complex medium (LB broth) on the copy number of two plasmids in *E. coli* K-12 HB101. They reported that copy number increased two to three times when the growth rate (μ, equivalent to the dilution rate) was increased from 0 to 1 h^{-1}. With further increase of the dilution rate, the copy number decreased sharply. The critical dilution rate (dilution for bioreactor wash out) was 1.4 h^{-1} when the cells were bearing the pBR322 plasmid (4.4 kb) and 1.2 h^{-1} with the recombinant pBR322-derivate pKTHI220 (14 kb). However, due to the complex nature of the LB broth, it is difficult to define the limiting substrate in those chemostats. As the stability and concentration of plasmids is also dependent on the nutritional situation (102), the mentioned results are difficult to interpret.

In a different study, the *E. coli* strain HB101 with plasmid pDM246 (a high copy number plasmid derived from pMB1) showed a maximum plasmid number per cell at a growth rate of 0.3 h^{-1} in minimal M9 medium chemostats. The same strain with plasmid RSF1050 in complex LB or minimal M9 medium showed a threefold increase of pDNA copy number per cell when lowering the growth rate from 1.25 to 0.35 h^{-1}, apparently having a maximum at the latter growth rate (103, 104).

Singer et al. (100) indicated that when using pUC-based plasmids and heat treatment, higher dilution rates might be advantageous for pDNA production, as the pDNA is formed faster (higher productivity) and the percentage of supercoiling is improved. On the other hand, the copy number of ColE1-type plasmids was found to have a negative correlation to the specific growth rate. However, Yau et al. (99) and Reinikainen et al. (101, 105) did not conclusively find a correlation between growth rate and the supercoiling level or pDNA yield.

4.1.2. Negative Correlation

During the feed stage of a feedback controlled fed-batch process, where the growth rate of *E. coli* DH10β in complex medium was controlled at 0.13 h^{-1}, Chen et al. (106) observed that the specific pDNA yield increased ten times compared to unlimited growth (0.69 h^{-1} attained without feedback control). Higher cell densities and the carbon/nitrogen limitations were also suggested as other possible causes of the improved pDNA yield.

A clearly negative correlation between specific growth rate and pDNA yield was found for *E. coli* RR1 with pBR329 plasmid in M9 minimal medium, M9CA minimal medium plus amino acids or LB complex medium (107), *E. coli* K-12 EC1005 with R1drd-19 in E minimal medium (108), *E. coli* M72 with pPLc-RP in LB complex medium (109), *E. coli* RL331T with pBR322 in LB complex medium or C minimal medium (110), *E. coli* JM103 with pUC8 plasmid in M9P minimal medium (111), *E. coli* M72 with pPLa23ll (3.8 kb), pPLa2311trpAl (8.1 kb), or pPLx23trpAl (6.5 kb) plasmid in M-56 minimal medium or M-56 with yeast extract (112), and several others (113, 114). Wunderlich (115) also found a negative correlation of the production of a 6-kb pDNA vaccine and growth rate in an *E. coli* W3110 in minimal medium. However, whereas the plasmid content of cells was maximum at a growth rate of 0.25 h^{-1}, a second local maximum was found at a growth rate of 0.15 h^{-1}.

In general, the negative correlation is attributed to a higher plasmid stability as well as plasmid replication being privileged over other biochemical pathways in reduced growth rates (116), which lead to a considerable pDNA accumulation since the replication of the plasmid continues at the same rate independently from the host cell division. Yet, the best growth rate should be determined for each particular pDNA vaccine.

4.2. Influence of Heat Treatment and the Chemical Environment

Since the replication of ColE1-like plasmids has *relaxed control*, which allows replication without the expression of new proteins, there is the possibility to increase the pDNA yield by supplying inhibitors of protein synthesis. For example, the supplement of chloramphenicol in the late exponential phase leads to an increase of up to 50-fold in plasmid copy number (14). However, applying this method in the production of active pharmaceutical ingredients may be difficult since it must be shown that all of the chloramphenicol

has been removed from the product. Accordingly, it is preferred to use a temperature increase to 42°C (105, 117), which is also an effective strategy.

Plasmids with a ColE1-type origin may be able to have an increased copy number as a result of a temperature elevation within the cultivation. In this way, for example in a fed-batch process, biomass can be grown in optimal conditions first with a later change of the cultivation temperature in order to induce the increased pDNA formation. The temperature-inducible copy number control can replace more complex copy number control systems such as a dual origin system or an inducible primer promoter. A prerequisite is a point mutation of guanine to adenine (G→A) immediately before the RNAI sequence. Depending on the temperature such a mutation prevents RNAII from folding normally, so RNAI cannot bind correctly, and consequently cannot efficiently control plasmid replication. However, this effect can be suppressed by lowering the cultivation temperature to 30°C or by the Rom protein. On the contrary, a temperature rise to 42–45°C in the absence of the Rom protein leads to highly increased plasmid copy numbers (from 20–50 to 400–500 copies per cell) (117–120). Singer et al. (100) reported that temperature treatment seems to have greater impact when bacteria are grown at a high growth rate. By applying the G→A point mutation, any ColE1-type plasmid can be converted into a temperature-inducible high copy number plasmid (117).

An alternative to improve pDNA yield is through adequate media composition and amino acid starvation. For example, Hofmann et al. (121) found in *arg⁻* strains of *E. coli* (CP79 and CP143) an increase of pDNA concentration of up to tenfold after full consumption of the amino acid arginine in both batch and fed-batch cultivation. In these strains the gene *relA* was deleted, meaning that upon arginine starvation the resulting strains were not able to form guanosine tetraphosphate (ppGpp), an alarmone of the stringent response, which was identified as an inhibitor of ColE1 plasmid replication (122). Other mutations preventing the stringent response, i.e., the change of metabolism caused by signal molecules such as ppGpp formed as a response to amino acid or carbon starvation, can provide similar results in pDNA production (123). Amino acid starvation is easy to apply and easy to scale-up. On the other hand it requires the addition of the special amino acid to the medium, which also implies additional costs and that time of amino acid depletion cannot be controlled precisely.

Another significant influence on the pDNA yield is the carbon-to-nitrogen ratio (C:N) of the cultivation medium. O'Kennedy et al. (124), using a complex nitrogen source, found that the maximum pDNA formation was reached at a C:N ratio of 2.78:1 mol_C/mol_N, whereas this ratio had no influence on the characteristics of cell lysis or breakage. Also the nitrogen source can be important. For instance, it has been reported that a combination of glutamate

and ammonium chloride at optimized concentrations can result in an important improvement of pDNA production (125). Carbon source can also have some influence on pDNA production (126).

Cultivation processes comprising medium optimization, fed-batch schemes, and temperature induction, have resulted in specific pDNA yields of up to 24.9 mg/L/OD_{600} and pDNA concentrations as high as 2 g/L (127).

4.3. High Cell-Density Processes

Attaining high cell densities in the bioreactor is a key technology for the commercialization of bioproducts. In the case of *E. coli*, several process limitations have influenced the operation mode of the cultivation. A prevalent problem in *E. coli* cultivation is the production of acetate, a by-product that is synthesized when the carbon uptake rate surpasses a certain threshold value even if fully aerobic conditions are maintained (128). This phenomenon is known as overflow metabolism, and has been attributed to an imbalance between the carbon fluxes in glycolysis and the tricarboxilic acid cycle (129). The accumulation of acetate provokes several undesirable effects on the physiology of *E. coli*. Also, it can affect the plasmid content, as shown in Fig. 4.

The traditional way to avoid overflow metabolism is to limit the carbon uptake rate below a threshold value, which can be done using fed-batch schemes. Usually, the threshold value corresponds to reduced specific growth rates of around 0.12 h^{-1}, being strongly

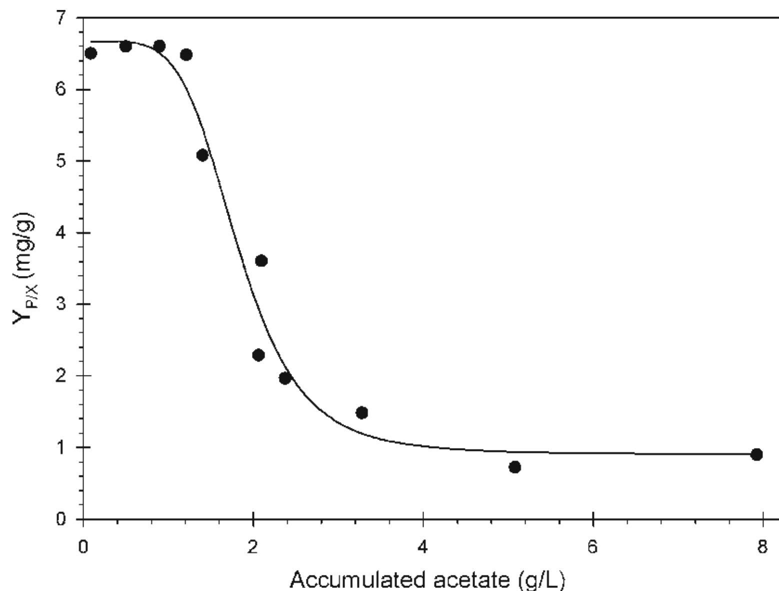


Fig. 4. Specific content of plasmid ($Y_{P/X}$) in *Escherichia coli* DH5α cells during a fed-batch cultivation at high growth rate. Acetate accumulated as a result of the elevated glucose uptake rate, causing a strong decrease in $Y_{P/X}$. Fully aerobic conditions were maintained and pH was controlled at 6.8.

dependent on the strain and cultivation conditions used. However, if the optimum growth rate is above the threshold value, acetate production will be unavoidable and the process will be operated under non-optimal conditions and the concomitant loss of productivity.

With a careful medium design it is possible to reach cell densities of more than 120 g/L (measured as dry cell weight). However, the bioreactor capacity for transferring oxygen can be limiting, especially as the cultivation scale increases. Although several strategies to increase oxygen transfer have been proposed, in general they are difficult to implement at large scale. Recently, the use of pressurized bioreactors has been proposed to supply enough oxygen to high cell-density cultures (130). Increasing the head-space pressure within the bioreactors increases the driving-force for oxygen transfer. This technique has been applied to *E. coli* DH5α high cell-density cultivations at a scale of 50 L and compared to the conventional strategy of sparging oxygen-enriched air to 1 L bioreactors, for the production of a 5.6-kb pDNA vaccine against hepatitis (131). In both cases, biomass concentrations of nearly 80 g/L were achieved, and the oxygen transfer rate (OTR) reached values of around 0.5 mol O_2/L h at both scales. Although the overpressure in the 50 L bioreactor was close to 8 bar at the end of the process, neither pDNA production nor its topology showed differences with that produced in small-scale cultivations. Therefore, pressurization can be potentially applied for the scale-up of pDNA production processes using high cell-density cultivations.

As the process scale increases, the existence of environmental gradients within the bioreactor becomes a prevalent problem (132). The effects of environmental gradients on *E. coli* have been the subject of detailed studies (133, 134); however, the effects of heterogeneities on pDNA production and topology are nearly inexistent. Cells can be subjected to additional stresses during pDNA production. For instance, the increase of temperature triggers a variety of physiological responses in *E. coli* known as the heat-shock response, as well as stringent response and SOS response (135, 136). This generalized stress response can limit the achievable biomass, and represent drainage of resources for pDNA synthesis. The particular effects of heat-shock on pDNA synthesis by *E. coli* have not been studied in detail. Induction by temperature has also been largely used for the production of recombinant proteins by *E. coli* using the pL and/or pR phage lambda promoters regulated by the thermolabile cI857 repressor. In this case, the heat-shock response represents a compromise between the adaptation of the cell to the temperature upshift and the production of the heterologous protein. It has been shown that at a relatively slow increase of temperature (from 30 to 42°C) at a rate of 0.4°C/min, the production of recombinant human pro-insulin was higher than under a fast heating rate (6°C/min) (137) and more biomass was accumulated. Similar strategies could also be evaluated for pDNA production.

5. Downstream Processes

Purification of pDNA has been a very active research area during the last few years. The initial pDNA purification step consists on disrupting the bacterial cells. However, traditional methods such as mechanical lysis are not suitable, since pDNA and chromosomal DNA can be sensitive to shear forces (138, 139) and DNA breakage can lead to product degradation or contamination. The method most commonly applied for cell disruption for initial pDNA recovery is based on the alkaline lysis, as originally described by Birnboim and Doly (140). The use of alkaline solutions and the increase in viscosity when pDNA is being isolated makes the scaling-up of this method an important challenge, since shear forces can be higher and cause a loss of supercoiled pDNA (141) or local regions of very high pH values could arise as an effect of reduction in mixing efficiency. Consequently, lytic processes should be carefully designed. For instance, Chamsart et al. (142), evaluated the effect of shear stress on the alkaline lysis of *E. coli* DH1, finding that at a shear rate of 662 s^{-1}, less than 2% of chromosomal DNA contaminated the pDNA produced. With this information, the process was scaled-up to 60 L scale with good results.

It has been reported that the amount of pDNA constitutes only 3% of the total *E. coli* lysates, whereas proteins represent 55% and RNA 21% (143). The separation of pDNA from relatively similar molecules, such as chromosomal DNA and RNA, is a key step. Usually, RNase A prepared from bovine pancreas is added to facilitate this step, with the consequent difficulties to totally eliminate such animal-derived protein in the final product (144). Tangential flow filtration has been proposed as an alternative method for RNA removal (145). Chromosomal DNA can also be removed by filtration through nitrocellulose membranes (146, 147). Darby et al. (148) reported the purification of pDNA by combining a crude lysate containing pDNA that was previously modified to include *lacO* sequences, with a second lysate containing the Lac repressor protein. Since the protein binds stably to the specific sequences, pDNA could be further separated by affinity chromatography without need for RNase.

The final purification of pDNA should ensure that all the contaminants are removed to acceptable levels and that supercoiled pDNA is selectively purified. It has already been mentioned that chromatography can be applied for pDNA purification. Hydrophobic interaction chromatography (149–151), thiophilic interaction chromatography (152), and anion exchange chromatography (153–155) have been applied for this purpose. Other downstream operations have been proposed for pDNA purification, such as thiophilic/aromatic adsorption (156), aqueous two-phase extraction (157, 158), and affinity interaction (159). Detailed

discussion of the different purification methods is beyond the scope of this chapter, but has been the theme of several published reviews (160–163).

6. Cell Engineering to Improve pDNA Synthesis and Purification

Modification of the cellular metabolism and physiology has been successfully applied to increase the synthesis and secretion of biomolecules, or to improve their purification. However, there are few reports on the application of metabolic engineering for improving pDNA production. Modified strains lacking *recA* and *endA* genes have been used for producing pDNA, as described in Subheading 3. Nevertheless, strains that can better cope with industrial conditions are also needed. For example, it has been previously described the undesirable effects of overflow metabolism. Commercial strains such as BL21 have a reduced overflow metabolism due to a more active glyoxylate shunt (164), which makes it possible to cultivate them in batch mode using up to 50 g/L of glycerol as carbon source with minimal acetate production. By knocking out the genes *recA* and *endA*, it was demonstrated that such a strain was able to produce even more pDNA than DH5α strain, which is generally a very efficient pDNA producer (95).

It has been demonstrated that the substitution of the PTS glucose transport system of *E. coli* by the galactose permease allows a more efficient glucose utilization, much lower acetate production (due to a lower glucose consumption rate) and higher biomass and recombinant protein production (129). The modified strain (PTS^- $GalP^+$) has been successfully grown up to 52 g/L of biomass in batch mode using 100 g/L of initial glucose concentration, with a very low amount of acetate accumulated and more than 8 g/L of recombinant protein produced (165). This strain has also been applied for the production of a 6-kb experimental pDNA vaccine. Not only higher cell densities were attained using 100 g/L of glucose in batch cultivation of the engineered strain but also the final titer of pDNA was more than double compared to the parent strain (166). This improvement in pDNA production has been explained not only by the reduction of overflow metabolism in the engineered strain but also because carbon flux through the pentose phospate pathway (PPP) is higher in PTS^- strains than in wild type strains (167), and this can result in a higher production of nucleotide precursors. The modified strain was also cultivated in exponential fed-batch mode at a growth rate of 0.35 h^{-1} (90% of its maximum growth rate in mineral medium). The productivity and final pDNA concentration showed an additional improvement of around 20% with respect to batch cultivation, whereas acetate production was negligible (167). Although such kind of cultivation strategies require very

high oxygen transfer rates, it has also been reported that pressurized bioreactors can supply enough oxygen (oxygen transfer rate of up to 0.45 mol O_2/L h) to ensure aerobic growth of the engineered strain using up to 130 g/L of glucose in batch mode (168).

Cunningham et al. (169) showed that a reduced activity of the pyruvate kinase resulted in higher pDNA copy number and cyclic AMP levels. By knocking out the *pykA* gene in the PTS^- $GalP^+$ strain, the pDNA production was tripled whereas acetate was undetectable in batch cultivation with 10 g/L of glucose (166). The reduction of carbon flux through the pyruvate kinase increases the flux through the PPP and also reduces the glycolytic flux, which explains the observed results.

It has been previously mentioned that introduction of a plasmid into *E. coli* results in a metabolic burden, generally reflected as a reduction in growth rate and biomass, and an increase in acetate production. Understanding the mechanism of the impact of plasmid production at a global physiological level is a challenge. Such knowledge can provide information useful for further cell engineering to improved pDNA production (170). For instance, based on metabolic flux analysis and enzyme activity studies, Wang et al. (65) over-expressed the gene *rpiA* (coding for ribose-5-phosphate isomerase A), leading to a more active PPP. As a result, the plasmid copy number increased nearly three times. Ow et al. (171, 172) inactivated the global metabolic regulator *fruR*. Apparently, this resulted in a reduced metabolic burden, as growth rate was importantly improved and pDNA production was modestly higher in the modified strain.

The described examples show the potential that metabolic engineering has to improve pDNA production. Based on a theoretical analysis, it has been suggested that the maximum achievable pDNA content in *E. coli* is 592 mg/g (173). The best published results up to now are far from such value. Hence, there is ample room for cellular modifications to increase pDNA yields in *E. coli*.

Cell engineering has also been applied to facilitate the purification of pDNA. For instance, the expression of ribonuclease A at the end of the cultivation of *E. coli* bearing pDNA efficiently reduced contaminating bacterial RNA. The enzyme translocated in the periplasm was released after alkaline lysis and was active enough to degrade most of the RNA (174). Likewise, by expressing endolysin, autolytical *E. coli* strains have eliminated the use of alkaline lysis with the consequent simplification of the process (175).

7. Perspectives

The application of pDNA as a therapeutic agent has reached an important degree of efficacy, and in a near future a variety of products will probably reach the market. This will demand the development

of better production strains, the design of more efficient and scalable downstream processes, as well as a parallel advance on the design of more potent, smaller, and safer vectors. The latter point should also include the development of ad hoc strain genotypes so that production can be improved. The trend to eliminate antibiotics as well as advances in cell engineering to facilitate downstream operations will result in more efficient processes. Metabolic engineering represents a very valuable tool to increase pDNA synthesis in *E. coli*, to obtain strains better prepared for process conditions, as well as to develop new cultivation techniques based on engineered strains, for safer and faster production of DNA vaccines.

Acknowledgments

This work was supported by CONACyT grants 84447 and 101847, PROMEP 47410089, and PAPIIT-UNAM IN-223210.

References

1. Palomares, L. A., Estrada-Mondaca, S., Ramírez, O. T. (2004). Production of recombinant proteins: Challenges and solutions. *Methods Molec. Biol.* **267**, 15–52.
2. Rossenberg, S.A., Aebersold, P., Cornetta, K., et al. (1990) Gene transfer into humans-immunotherapy of patients with advanced melanoma, suing tumor-infiltrating lymphocites modified by retroviral gene transduction. *N. Engl. J. Med.* **323**, 570–578.
3. Journal of Gene Medicine web presence 2010. Gene Therapy Clinical Trials Worldwide. http://www.wiley.co.uk/genmed/clinical/.
4. Nichols, W. W., Ledwith, B. J., Manam, S. V., et al., (1995) Potential DNA vaccine integration into host cell genome. *Ann. NY Acad. Sci.* **772**, 30–39.
5. Raper, S. E., Chirmule, N., Lee, F. S., et al., (2003) Fatal systemic inflammatory response syndrome in a ornithine transcarbamylase deficient patient following adenoviral gene transfer. *Mol. Genet. Metabol.* **80**, 148–158.
6. Locher, C., Putnam, D., Langer, R., et al., (2003) Enhancement of a human immunodeficiency virus env DNA vaccine using a novel polycationic nanoparticle formulation. *Immunol. Lett.* **90**, 67–70.
7. Wang, S., Liu, X., Fisher, K., et al., (2000) Enhanced type I immune response to a hepatitis B DNA vaccine by formulation with calcium- or aluminium phosphate. *Vaccine* **18**, 1227–1235.
8. Aguiar, J., Hedstrom, R., Rogers, W., et al., (2001) Enhancement of the immune response in rabbits to a malaria DNA vaccine by immunization with a needle-free jet device. *Vaccine* **20**, 275–280.
9. Babiuk, S., Baca-Estrada, M., Foldvari, M., et al., (2004) Increased gene expression and inflammatory cell infiltration caused by electroporation are both important for improving the efficacy of DNA vaccines. *J. Biotechnol.* **110**, 1–10.
10. Bachy, M., Boudet, F., Bureau, M., et al., (2001) Electric pulses increase the immunogenicity of an influenza DNA vaccine injected intramuscularly in the mouse. *Vaccine* **19**, 1688–1693.
11. FDA (2007) Guidance for industry: Considerations for plasmid DNA vaccines for infectious disease indications. US Department of Health and Human Services, Food and Drug Administration.
12. Han, Y., Liu, S., Ho, J., et al., (2009) Using DNA as a drug—Bioprocessing and delivery strategies. *Chem. Eng. Res. Des.* **87**, 343–348.
13. Williams, J. A., Carnes, A. E., Hodgson, C. P. (2009) Plasmid DNA vaccine vector design: Impact on efficacy, safety and upstream production. *Biotechnol. Adv.* **27**, 353–370.
14. Summers, D. K. (1996) The biology of plasmids. Blackwell Science, Oxford, UK.
15. Bower, D.M., Prather, K.L.J. (2009) Engineering of bacterial strains and vectors for the production of plasmid DNA. *Appl. Microbio. Biotechnol.* **82**, 805–813.

16. Schumann, W. (2001) The biology of plasmids, in *Plasmids for therapy and vaccination* (Schleef, M., Ed.), pp. 1–28 Wiley –VCH Verlag GmbH, Weinheim.
17. Cesareni, G., Muesing, M. A., Polisky, B. (1982) Control of ColE1 DNA replication: the rop gene product negatively affects transcription from the replication primer promoter. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6313–6317.
18. Lin-Chao, S., Chen, W., and Wong, T. (1992) High copy number of the pUC plasmid results from a Rom/Rop-suppressible point mutation in RNAII. *Molec. Microbiol.* **6**, 3385–3393.
19. Yanisch-Perron, C., Vieira, J., Messing, J. (1995) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mpl8 and pUC19 vectors. *Gene* **33**, 103–119.
20. Manthorpe, M., Hobart, P., Hermanson, G., et al., (2005) Plasmid Vaccines and Therapeutics: From Design to Applications. *Adv. Biochem. Eng. Biotechnol.* **99**, 41–92.
21. Soubrier, F., Cameron, B., Manse, B., et al. (1999) pCOR: A new design of plasmid vectors for nonviral gene therapy. *Gene Ther.* **6**, 1482–1488.
22. Butler, V. A. (1996) Points to consider on plasmid DNA vaccines for preventive infectious disease indications. Center for Biologics Evaluation and Research, Food and Drug Administration. Docket 96 N-0400.
23. Valenzuela, M. S., Siddiqui, K. A., Sarkar, B. L. (1996) High expression of plasmid-encoded tetracycline resistance gene in *E. coli* causes a decrease in membrane-bound ATPase activity. *Plasmid* **36**, 19–25.
24. Gill, D. R., Smyth, S. E., Goddard, C. A., et al., (2001) Increased persistence of lung gene expression using plasmids containing the ubiquitin C or elongation factor 1α promoter. *Gene Ther.* **8**, 1539–1546.
25. Müller, P. P., Oumard, A., Wirth, D., et al., (2001) Polyvalent vector for coexpression of multiple genes, in *Plasmids for therapy and vaccination* (Schleef, M., Ed.), pp. 119–137, Wiley –VCH Verlag GmbH, Weinheim.
26. Azzoni, A. R., Ribeiro, S. C., Monteiro G. A., et al., (2007) The impact of polyadenylation signals on plasmid nuclease-resistance and transgene expression. *J. Gene Med.* **9**, 392–402.
27. Carvalho, J. A., Azzoni, A. R., Prazeres, D. M., et al., (2010). Comparative analysis of antigen-targeting sequences used in DNA vaccines. *Molec. Biotechnol.* **44**, 204–212.
28. Ow. D. S. W., Lee, D. Y., Tung, H. H., Chao, S. L. (2009) Plasmid regulation at systems level effects on *Escherichia coli* metabolism, in *Systems Biology and Biotechnology of Escherichia coli*, (Lee, S. Y., Ed), pp. 239–294. Springer Science, Heidelberg.
29. Lee, C. H., Mizusawa, H., Kakefuda, T. (1981) Unwinding of double-stranded DNA helix by dehydration. *Proc. Natl. Acad. Sci. U.S.A.* **78**, 2838–2842.
30. Dayn, A., Malkhosyan, S., Duzhy, D., et al., (1991). Formation of $(dA\text{-}dT)_n$ cruciforms in *Escherichia coli* cells under different environmental conditions. *J. Bact.* **173**, 2658–2664.
31. Rothenburg, S., Koch-Nolte, F., Haag, F. (2001) DNA methylation and Z-DNA formation as mediators of quantitative differences in the expression of alleles. *Immunol. Rev.* **184**, 286–298.
32. Ghosh, A., Bansal, M. (2003) A glossary of DNA structures from A to Z. *Acta Crystallogr. D Biol. Crystallogr.* **59**, 620–626.
33. Higgins, N. P., Vologodskii, A. (2004). Topological behavior of plasmid DNA, in *Plasmids Biology* (Funnel, B. E., Philips, G., Eds), pp. 181–201. ASM Press, Washington, DC.
34. Schmidt, T., Friehs, K., Flaschel, E. (2001) Structures of plasmid DNA, in Plasmids for therapy and vaccination (Schleef, M., Ed.), pp. 119–137, Wiley –VCH Verlag GmbH, Weinheim.
35. Reece, R. J., Maxwell, A. (1991) DNA gyrase: Structure and function. *Crit. Rev. Biochem. Mol. Biol.* **26**, 335–75.
36. Bauer, W. R., Crick, F. H. C., White, J. H. (1980) Supercoiled DNA. *Sci Am.* **243**, 100–113.
37. Cupillard, L., Juillard, V., Latour, S., et al.,(2005). Impact of plasmid supercoiling on the efficacy of a rabies DNA vaccine to protect cats. *Vaccine* **23**, 1910–1916.
38. Pillai, V. B., Hellerstein, M., Yu, T., et al.,(2008). Comparative studies on in vitro expression and in vivo immunogenicity of supercoiled and open circular forms of plasmid DNA vaccines. *Vaccine* **26**, 1136–1141.
39. Leahy, P., Carmichael, G. G., Rossomando, E. F. (1997) Transcription from plasmid expression vectors is increased up to 14-fold when plasmids are transfected as concatemers. *Nucl. Acids Res.* **25**, 449–450.
40. Haugland, R. P. (1996) Handbook of fluorescence probes and research chemicals. Leiden, The Netherlands: Molecular Probes, Inc.
41. Singer, V. L., Jones, L. J., Yue, S.T., et al., (1997). Characterization of Picogreen reagent and development of a fluorescence-based solution assay for double stranded DNA quantitation. *Anal. Biochem.* **249**, 228–238.

42. Noites, I. S., O'Kennedy, R. D., Levy, M. S., et al., (1999) Rapid quantitation and monitoring of plasmid DNA using an ultrasensitive DNA-binding dye. *Biotechnol. Bioeng.* **66**, 195–201.
43. Rock, C., Shamlou P. A., Levy, M. S. (2003) An automated microplate-based method for monitoring DNA strand breaks in plasmids and bacterial artificial chromosomes. *Nucl. Acids Res.* **31**, e65.
44. Levy, M. S., Loftian, P., O'Kennedy, R., et al., (2000) Quantitation of supercoiled circular content in plasmid DNA solutions using a fluorescence-based method. *Nucl. Acids Res.* **28**, e57.
45. Tanaka, H., Mielke, S. P., Benham, C, J., et al., (2008) Visualization of the detailed structure of plasmid DNA. *J. Phys. Chem. B.* **112**, 16788–16792.
46. Schmidt, T., Friehs, K., Flaschel, E. (1999) Quantitative analysis of plasmid forms by agarose and capillary gel electrophoresis. *Anal. Biochem.* **274**, 235–240.
47. Keller, W. (1975) Determination of the number of superhelical turns in simian virus 40 DNA by gel electrophoresis. *Proc. Natl. Acad. Sci. U.S.A.* **72**, 4876–4880.
48. Kang, D. S., Wells, R. D. (1985) B-Z DNA junctions containing few, if any, nonpaired bases at physiological superhelical densities. *J. Biol. Chem.* **260**, 7783–7790.
49. Vetcher, A. A., McEwen, A. E., Abujarour, R., et al., (2010). Gel mobilities of linking-number topoisomers and their dependence on DNA helical repeat and elasticity. *Biophys. Chem.* **148**, 104–111.
50. Ferreira, G. N. M., Cabral, J. M. S., Prazeres, D. F. M. (1999) Monitoring of process streams in the large-scale purification of plasmid DNA for gene therapy applications. *Pharm. Pharmacol. Commun.* **5**, 57–59.
51. Kendall, D., Booth, A. J., Levy, M. S., et al., (2001) Separation of supercoiled and open-circular plasmid DNA by liquid-liquid counter-current chromatography. *Biotechnol. Lett.* **23**, 613–619.
52. Diogo M. M., Queiroz, J. A., Prazeres, D. M. F. (2005) Chromatography of plasmid DNA. *J Chromatogr A.* **1069**, 3–22.
53. Weigl, D., Molloy, M. J., Clayton, T. M., et al., (2006) Characterization of a topologically aberrant plasmid population from pilot-scale production of clinical-grade DNA. *J. Biotechnol.* **121**, 1–12.
54. Smith, C. R., DePrince, R. B., Dackor, J., et al., (2007) Separation of topological forms of plasmid DNA by anion-exchange HPLC: shifts in elution order of linear DNA. *J. Chromatogr. B.* **854**, 121–127.
55. Middaugh C, R., Evans, R. K., Montgomery, D. L., et al., (1998) Analysis of Plasmid DNA from a Pharmaceutical Perspective. *J. Pharm. Sci.* **87**, 130–146.
56. Latulippe, D. R., Zydney, A. L. (2010) Radius of gyration of plasmid DNA isoforms from static light scattering. *Biotechnol. Bioeng.* **107**, 134–412.
57. Gao, X., Kim, K.S., Liu, D. (2007) Nonviral gene delivery: what we know and what is next. *AAPS J.* **9**, E92–E104.
58. Lamb, B. T., Sisodia, S. S., Lawler, A. M., et al., (1993) Introduction and expression of the 400 kilobase amyloid precursor protein gene in transgenic mice. *Nat. Genet.* **5**, 22–30.
59. Lufino, M. M., Edser, P. A., Wade-Martins, R. (2008) Advances in high-capacity extrachromosomal vector technology: episomal maintenance, vector delivery, and transgene expression. *Molec. Ther.* **16**, 1525–1538.
60. Robertson, J. S., Griffiths, E. (2006) Assuring the quality, safety, and efficacy of DNA vaccines. *Methods Mol. Med.* **127**, 363–374.
61. Gill, D. R., Pringle, I. A., Hyde, S. C. (2009) Progress and Prospects: The design and production of plasmid vectors. *Gene Ther.* **16**, 165–171.
62. Valera, A., Perales, J. C., Hatzoglou, M., et al., (1994) Expression of the neomycin resistance (*neo*) gene induces alterations in gene expression and metabolism. *Hum. Gene Ther.* **5**, 449–456.
63. Rozkov, A., Avignone-Rossa, C., Ertl, P., et al., (2004) Characterization of the metabolic burden on *Escherichia coli* DH1 cells imposed by the presence of a plasmid containing a gene therapy sequence. *Biotechnol. Bioeng.* **88**, 909–915.
64. Ow, D. S. W., Nissom P. M., Philp R., et al., (2006) Global transcriptional analysis of metabolic burden due to plasmid maintenance in *Escherichia coli* DH5α during batch fermentation. *Enzyme Microb. Technol.* **39**, 391–398.
65. Wang, Z., Xiang, L., Shao, J., et al.,(2006) Effects of the presence of ColE1 plasmid DNA in *Escherichia coli* on the host cell metabolism. *Microb. Cell Fact.* **5**, 34.
66. Mairhofer, J., Cserjan-Puschmann, M., Striedner, G., et al., (2010) Marker-free plasmids for gene therapeutic applications—Lack of antibiotic resistance gene substantially improves the manufacturing process. *J. Biotechnol.* **146**, 130–137.
67. Cranenburgh, R.M., Hanak, J.A., Williams, S.G., et al., (2001) *Escherichia coli* strains that

allow antibiotic-free plasmid selection and maintenance by repressor titration. *Nucl. Acids Res.* **29**, E26.

68. Cranenburgh, R. M., Lewis, K. S., Hanak J. A. (2004) Effect of plasmid copy number and *lac* operator sequence on antibiotic-free plasmid selection by operator-repressor titration in *Escherichia coli. J. Mol. Microbiol. Biotechnol.* **7**, 197–203.
69. Williams, S. G., R. M. Cranenburgh, A. M. Weiss, C. J. et al., (1998). Repressor titration: a novel system for selection and stable maintenance of recombinant plasmids. *Nucl. Acids Res.* **26**, 2120–2124.
70. Hanke, T., McMichael, A.J. (2000) Design and construction of an experimental HIV-1 vaccine for a year-2000 clinical trial in Kenya. *Nat. Med.* **6**, 951–955.
71. Vidal, L., Pinsach, J., Striedner, G., et al., (2008) Development of an antibiotic-free plasmid selection system based on glycine auxotrophy for recombinant protein overproduction in *Escherichia coli. J. Biotechnol.* **134**, 127–136.
72. Hägg, P., de Pohl, J. W., Abdulkarim, F., et al., (2004) A host/plasmid system that is not dependent on antibiotics and antibiotic resistance genes for stable plasmid maintenance in *Escherichia coli. J. Biotechnol.* **111**, 17–30.
73. Morona, R., Yeadon, J., Considine, A., et al., (1991) Construction of plasmid vectors with a non-antibiotic selection system based on the *Escherichia coli thyA* gene: application to cholera vaccine development. *Gene* **107**, 139–144.
74. Marie, C., Vandermeulen, G., Quiviger, M., et al., (2010) pFARs, Plasmids free of antibiotic resistance markers, display high-level transgene expression in muscle, skin and tumour cells. *J. Gene Med.* **12**, 323–332.
75. Dong, W. R., Xiang, L. X., Shao, J. Z. (2010) Novel Antibiotic-Free Plasmid Selection System Based on Complementation of Host Auxotrophy in the NAD De Novo Synthesis Pathway. *Appl. Environ. Microbiol.* **76**, 2295–2303.
76. Soubrier, F., Laborderie, B., Cameron, B. (2005) Improvement of pCOR plasmid copy number for pharmaceutical applications. *Appl. Microbiol. Biotechnol.* **66**, 683–688.
77. Pfaffenzeller, I., Mairhofer, J., Striedner, G., Bayer, K., Grabherr, R. (2006) Using ColE1-derived RNA I for suppression of a bacterially encoded gene: implication for a novel plasmid addiction system. *Biotechnol. J.* **1**, 675–681.
78. Mairhofer, J., Pfaffenzeller, I., Merz, D., et al., (2008) A novel antibiotic free plasmid selection system: Advances in safe and efficient DNA therapy. *Biotechnol. J.* **3**, 83–89.
79. Luke, J., Carnes A. E., Hodgson, C. P., et al., (2010) Improved antibiotic-free DNA vaccine vectors utilizing a novel RNA based plasmid selection system. *Vaccine* **27**, 6454–6459.
80. Kreiss, P., Cameron, B., Rangara, R., et al., (1999) Plasmid DNA size does not affect the physicochemical properties of lipoplexes but modulates gene transfer efficiency. *Nucl. Acids Res.* **27**, 3792–3798.
81. Junhghaus, C., Schroff, M., Koening-Merediz, S. A., et al., (2001) Form follows function: the design of minimalistic immunogenically defined gene expression (MIDGE®) constructs, in *Plasmids for therapy and vaccination* (Schleef, M., Ed.), pp. 119–137, Wiley –VCH Verlag GmbH, Weinheim.
82. Moreno, S., López-Fuertes, L., Vila-Coro, A. J., et al., (2004) DNA immunization with minimalistic expression constructs. *Vaccine* **22**, 1709–1716.
83. López-Fuertes, L., Pérez-Jiménez, E., Vila-Coro, A.J., et al., (2002) DNA vaccination with linear minimalistic (MIDGE) vectors confers protection against *Leishmania major* infection in mice. *Vaccine* **21**, 247–257.
84. Köchling, J., Prada, J., Bahrami, M., et al., (2008) Anti-tumor effect of DNA-based vaccination and dSLIM immunomodulatory molecules in mice with Ph^+ acute lymphoblastic leukaemia. *Vaccine* **26**, 4669–4675.
85. Schirmbeck, R., Konig-Merediz, S. A., Riedl, P., et al., (2001) Priming of immune responses to hepatitis B surface antigen with minimal DNA expression constructs modified with a nuclear localization signal peptide. *J. Molec. Med.* **79**, 343–50.
86. Darquet, A. M., Cameron, B., Wils, P., et al., (1997) A new DNA vehicle for nonviral gene delivery: supercoiled minicircle. *Gene Ther.* **4**, 1341–1349.
87. Darquet, A. M., Rangara, R., Kreiss, P., et al., (2009) Minicircle: an improved DNA molecule for in vitro and in vivo gene transfer. *Gene Ther.* **6**, 209–218.
88. Faurez, F., Dory, D., Le Moigne, V., et al., (2010) Biosafety of DNA vaccines: New generation of DNA vectors and current knowledge on the fate of plasmids after injection. *Vaccine* **28**, 3888–3895.
89. Bigger, B.W., Tolmachov, O., Collombet, J. M., et al., (2001) An araC-controlled bacterial cre expression system to produce DNA minicircle vectors for nuclear and mitochondrial gene therapy. *J. Biol. Chem.* **276**, 23018–23027.

90. Vaysse, L., Gregory, L. G., Harbottle, R. P., et al., (2006) Nuclear-targeted minicircle to enhance gene transfer with non-viral vectors in vitro and in vivo. *J. Gene Med.* **8**, 754–763.
91. Zhang, X., Epperly, M. W., Kay, M. A., et al., (2008) Radioprotection in vitro and in vivo by minicircle plasmid carrying the human manganese superoxide dismutase transgene. *Hum. Gene Ther.* **19**, 820–826.
92. Prather, K. J., Edmonds, M. C., Herods, J. W. (2006) Identification and characterization of IS1 transposition in plasmid amplification mutants of *E. coli* clones producing DNA vaccines. *Appl. Microbiol. Biotechnol.* **73**, 815–826.
93. Oliveira, P. H., Prather, K. J., Prazeres, D.M.F., et al., (2009) Structural instability of plasmid biopharmaceuticals: challenges and implications. *Trends Biotechnol.* **27**, 503–511.
94. Schoenfeld, T., Mendez, J., Storts, D. R., et al., (1995) Effects of Bacterial Strains Carrying the *endA1* Genotype on DNA Quality Isolated with Wizard™ Plasmid Purification Systems. *Promega Notes* **53**, 12–19.
95. Phue, J. N., Lee, S. J., Trinh, L., et al., (2008) Modified *Escherichia coli* B (BL21), a superior producer of plasmid DNA compared with *Escherichia coli* K (DH5alpha). *Biotechnol. Bioeng.* **101,** 831–836.
96. Bell, C. E. (2005) Structure and mechanism of *Escherichia coli* RecA ATPase. *Mol. Microbiol.* **58**, 358–366.
97. Kuzminov, A. (1999) Recombinational Repair of DNA Damage in *Escherichia coli* and Bacteriophage l. *Microbiol. Mol. Biol. Rev.* **63**, 751–813.
98. Voss, C., Schmidt, T., Schleef, M., et al., (2003) Production of supercoiled multimeric plasmid DNA for biopharmaceutical application. *J. Biotechnol.* **105**, 205–213.
99. Yau, S. Y., Keshavarz-Moore, E., Ward, J. (2008) Host strain influences on supercoiled plasmid DNA production in *Escherichia coli*: Implications for efficient design of large-scale processes. *Biotechnol. Bioeng.* **101,** 529–544.
100. Singer, A., Eiteman, M. A., Altman E. (2009) DNA plasmid production in different host strains of *Escherichia coli*. *J. Ind. Microbiol. Biotechnol.* **36,** 521–530.
101. Reinikainen, P., Virkajärvi, I. (1989) *Escherichia coli* growth and plasmid copy numbers in continuous cultures. *Biotechnol. Lett.* **11**, 225–230.
102. Zabriskie, D.W., Arcuri, E.J. (1986) Factors influencing productivity of fermentations employing recombinant microorganisms. *Enzyme Microb. Technol.* **8**, 706–717.
103. Seo, J., Bailey, J. E. (1986) Continuous cultivation of recombinant *Escherichia coli*: existence of an optimum dilution rate for maximum plasmid and gene product concentration. *Biotechnol. Bioeng.* **28**, 1590–1594.
104. Seo, J., Bailey, J.E. (1985) Effects of recombinant plasmid content on growth properties and cloned gene product formation in *Escherichia coli*. *Biotechnol. Bioeng.* **27**, 1668–1674.
105. Reinikainen, P., Korpela, K., Nissinen, V., et al., (1989) *Escherichia coli* plasmid production in a fermenter. *Biotechnol. Bioeng.* **33**, 386–393.
106. Chen, W., Graham, C., Ciccarelli, R. B. (1997) Automated fed-batch fermentation with feed-back controls based on dissolved oxygen (DO) and pH for production of DNA vaccines. *J. Ind. Microbiol. Biotechnol.* **18**, 43–48.
107. Bentley, W. E., Mirjalili, N., Andersen, D. C., et al., (1990) Plasmid-encoded protein: The principal factor in the "metabolic burden" associated with recombinant bacteria. *Biotechnol. Bioeng.* **35**, 668–681.
108. Engberg, B., Nordström, K. (1975) Replication of R-factor R1 in *Escherichia coli* K-12 at different growth rates. *J. Bact.* **123**, 179–186.
109. Kim, J.Y., Ryu, D.D.Y. (1991) The effects of plasmid content, transcription efficiency, and translation efficiency on the productivity of a cloned gene protein in *Escherichia coli*. *Biotechnol. Bioeng.* **38**, 1271–1279.
110. Lin-Chao, S., Bremer, H. (1986) Effect of the bacterial growth rate on replication control of plasmid pBR322 in *Escherichia coli*. *Mol. Gen. Genet.* **203**, 143–149.
111. Ryan, W., Parulekar, S.J. (1991) Recombinant protein synthesis and plasmid instability in continuous cultures of *Escherichia coli* JM103 harbouring a high copy number plasmid. *Biotechnol. Bioeng.* **37**, 415–429.
112. Siegel, R., Ryu, D. D. (1985) Kinetic study of instability of recombinant plasmid pPLc23trpAI in *E. coli* using two-stage continuous culture system. *Biotechnol. Bioeng.* **27**, 28–33.
113. O'Kennedy, R. D., Ward, J.M., Keshavarz-Moore, E. (2003) Effects of fermentation strategy on the characteristics of plasmid DNA production. *Biotechnol. Appl. Biochem.* **37**, 83–90.
114. Rozkov, A., Avignone-Rossa, C. A., Ertl, P. F., et al., (2006) Fed batch culture with declining specific growth rate for high-yielding

production of a plasmid containing a gene therapy sequence in *Escherichia coli* DH1. *Enzyme Microb. Technol.* **39**, 47–50.

115. Wunderlich, M. (2010) Diploma Thesis. Technische Universität Dresden, Germany.
116. Prather, K. J., Sagar, S., Murphy, J., et al., (2003) Industrial scale production of plasmid DNA for vaccine and gene therapy: plasmid design, production, and purification. *Enzyme Microbial Technol.* **33**, 865–883.
117. Lahijani, R., Hulley, G., Soriano, G., et al., (1996) High-yield production of pBR322-derived plasmids intended for human gene therapy by employing a temperature-controllable point mutation. *Hum. Gene Ther.* **7**, 1971–1980.
118. Lin-Chao, S., Chen, W., Wong, T. (1992) High copy number of the pUC plasmid results from a Rom/Rop-suppressible point mutation in RNAII. *Mol. Microbiol.* **6**, 3385–3393.
119. Miki, T., Yasukochi, T., Nagatani, H., et al., (1987) Construction of a plasmid vector for the regulatable high level expression of eukaryotic genes in *Escherichia coli*: an application to overproduction of chicken lysozyme. *Protein Eng.* **1**, 327–332.
120. Wong, E. M., Muesing, M. A., Polisky, B. (1982) Temperature-sensitive copy number mutants of ColE1 are located in an untranslated region of the plasmid genome. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 3570–3574.
121. Hofmann, K. H., Neubauer, P., Reithdorf, S., et al., (1990) Amplification of pBR322 plasmid DNA in *Escherichia coli relA* strains during batch and fed-batch fermentation. *J. Basic Microbiol.* **30**, 37–41.
122. Hecker, M., Schroeter, A., Mach, F. (1985) *Escherichia coli relA* strains as hosts for amplification of pBR322 plasmid DNA. *FEMS Microbiol. Lett.* **29**, 331–334.
123. Williams, J. A., Luke, J., Langtry, S., et al., (2009b) Generic plasmid DNA production platform incorporating low metabolic burden seed-stock and fed-batch fermentation processes. *Biotechnol. Bioeng.* **103**, 1129–1243.
124. O'Kennedy, R. D., Baldwin, C., Keshavarz-Moore, E. (2000) Effects of growth medium selection on plasmid DNA production and initial processing steps. *J. Biotechnol.* **76**, 175–183.
125. Voss, C., Schmidt, T., Schleef, M., et al., (2004) Effect of ammonium chloride on plasmid DNA production in high cell density batch culture for biopharmaceutical use. *J. Chem. Technol. Biotechnol.* **79**, 57–62.
126. Zheng, S., Friehs, K., He, N., et al., (2007) Optimization of medium components for plasmid production by recombinant *E. coli* DH5α pUK21CMVβ1.2. *Biotechnol. Bioproc. Eng.* **12**, 213–221.
127. Carnes, A. E., Luke, J. M., Vincent, J. M., et al., (2010) Plasmid DNA fermentation strain and process-specific effects on vector yield, quality and transgene expression. *Biotechnol. Bioeng. In press.*
128. Eiteman, M. A., Altman, E. (2006) Overcoming acetate in *Escherichia coli* recombinant protein fermentations. *Trends Biotechnol.* **24**, 530–536.
129. De Anda, R., Lara, A. R., Hernández, V., et al., (2006) Replacement of the glucose phosphotransferase transport system by galactose permease reduces acetate accumulation and improves process performance of *Escherichia coli* for recombinant protein production without impairment of growth rate. *Metab. Eng.* **8**, 281–290.
130. Knoll, A., Bartsch, S., Husemann, B., et al., (2007) High cell density cultivation of recombinant yeasts and bacteria under non-pressurized and pressurized conditions in stirred tank bioreactors. *J. Biotechnol.* **132**, 167–179.
131. Lara, A. R., Knabben, I., Caspeta, L., et al., (2011) Comparison of oxygen enriched air vs pressurized cultivations to increase oxygen transfer and to scale-up plasmid DNA production fermentations. *Eng. Life Sci.* **11**, 382–386.
132. Lara, A. R., Galindo, E., Ramírez, O. T., et al., (2006) Living with heterogeneities in bioreactors: Understanding the effects of environmental gradients on cells. *Mol. Biotechnol.* **34**, 355–381.
133. Lara, A. R., Leal, L. I., Flores, N., et al., (2006) Transcriptional and metabolic response of recombinant *Escherichia coli* to spatial dissolved oxygen tension gradients simulated in a scale-down system. *Biotechnol. Bioeng.* **93**, 373–385.
134. Lara A. R., Taymaz-Nikerel, H., van Gulik, W., et al., (2009) Fast dynamic response of *Escherichia coli* fermentative metabolism to aerobic and anaerobic glucose pulses. *Biotechnol. Bioeng.* **104**, 1153–1161.
135. Wittmann, C., Weber, J., Betiku, E., et al., (2007). Response of fluxome and metabolome to temperature-induced recombinant protein synthesis in *Escherichia coli*. *J. Biotechnol.* **132**, 375–384.
136. Valdéz-Cruz, N. A., Caspeta, L., Pérez, N. O., et al., (2009) Production of recombinant proteins in *E. coli* by the heat inducible expression system based on the phage lambda pL and/or pR promoters. *Microb. Cell Fact.* **9**:18.

137. Caspeta, L., Flores N., Pérez, N. O., et al., (2009) The effect of heating rate on *Escherichia coli* metabolism, physiological stress, transcriptional response, and production of temperature-induced recombinant protein: A scale-down study. *Biotechnol. Bioeng.* **102**, 468–482.
138. Levy, M. S., O'Kennedy R. O., Ayazi-Shamlou, P., et al., (2000) Biochemical engineering approaches to the challenge of producing pure plasmid DNA. *Trends Biotechnol.* **18**, 296–305.
139. Zhang, H., Kong, S., Booth, A., et al., (2008) Prediction of shear damage of plasmid DNA in pump and centrifuge operations using an ultra scale-down device. *Biotechnol. Prog.* **23**, 858–865.
140. Birnboim, H. C., Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acids Res.* **7**, 1513–1523.
141. Clemson, M., Kelly, W. J. (2003) Optimizing alkaline lysis for plasmid DNA recovery. *Biotechnol. Appl. Biochem.* **37**, 235–244.
142. Chamsart, S., Patel, H., Hanak, J. A. J., et al., (2001) The impact of fluid-dynamic-generated stresses on chDNA and pDNA stability during alkaline cell lysis for gene therapy products. *Biotechnol. Bioeng.* **75**, 387–392.
143. Stadler, J., Lemmens, R., Nyhammar, T. (2004). Plasmid DNA purification. *J. Gene Med.* **6**, S54-S66.
144. Voss, C. (2007). Production of plasmid DNA for pharmaceutical use. *Biotechnol. Ann. Rev.* **13**, 201–221.
145. Eon-Duval, A., MacDuff, R. H., Fisher, C. A., et al., (2003) Removal of RNA impurities by tangential flow filtration in an RNase-free plasmid DNA purification process. *Anal. Biochem.* **316**, 66–73.
146. Levy, M. S., Collins, I. J., Tsai, J. T., et al., (2000) Removal of contaminant nucleic acids by nitrocellulose filtration during pharmaceutical-grade plasmid DNA processing. *J. Biotechnol.* **76**, 197–205.
147. Kendall, D., Lye, G. J., Levy, M. S. (2002) Purification of plasmid DNA by an integrated operation comprising tangential flow filtration and nitrocellulose adsorption. *Biotechnol. Bioeng.* **79**, 816–822.
148. Darby, R. A. J., Forde, G. M., Slater N. K. H., et al., (2007) Affinity purification of plasmid DNA directly from crude bacterial cell lysates. *Biotechnol. Bioeng.* **98**, 1103–1108.
149. Diogo, M. M., Queiroz, J. A., Monteiro, G. A., et al., (2000) Purification of a cystic fibrosis plasmid vector for gene therapy using hydrophobic interaction chromatography. *Biotechnol. Bioeng.* **68**, 576–583.
150. Diogo, M. M., Ribeiro, S. C., Queiroz, J. A., et al., (2001) Production, purification and analysis of an experimental DNA vaccine against rabies. *J. Gene Med.* **3**, 577–584.
151. Eon-Duval, A., Burke, G. (2004) Purification of pharmaceutical-grade plasmid DNA by anion-exchange chromatography in an RNase-free process. *J. Chrom. B.* **804**, 327–335.
152. Sandberg, L. M., Bjurling, A., Busson, P., et al., (2004) Thiophilic interaction chromatography for supercoiled plasmid DNA purification. *J. Biotechnol.* **109**, 193–199.
153. Horn, N. A., Meek, J. A., Budahazi, G., et al., (1995) Cancer gene therapy using plasmid DNA: purification of DNA for human clinical trials. *Hum. Gene Ther.* **6**, 565–573.
154. Ferreira, G. N. M., Cabral, J. M. S., Prazeres, D. M. F. (1999) Development of process flow sheets for the purification of supercoiled plasmids for gene therapy applications. *Biotechnol. Prog.* **15**, 725–731.
155. Guerrero-Germán, P., Prazeres, D. M. F., Guzmán, R., et al., (2009) Purification of plasmid DNA using tangential flow filtration and tandem anion-exchange membrane chromatography. *Bioproc. Biosyst. Eng.* **32**, 615–623.
156. Lemmens, R., Olsson, U., Nyhammar, T., et al., (2003) Supercoiled plasmid DNA: selective purification by thiophilic/aromatic adsorption. *J. Chrom. B.* **784**, 291–300.
157. Trindade, I. P., Diogo, M. M., Prazeres, D. M. F., et al., (2005) Purification of plasmid DNA vectors by aqueous two-phase extraction and hydrophobic interaction chromatography. *J. Chrom A.* **1082**, 176–184.
158. Luechau, F., Ling, T. C., Lyddiatt, A. (2010) Two-step process for initial capture of plasmid DNA and partial removal of RNA using aqueous two-phase systems. *Proc. Biochem.* **45**, 1432–1436.
159. Schluep, T., Cooney, C. L. (1998) Purification of plasmids by triplex affinity interaction. *Nucl. Acids Res.* **26**, 4524–4528.
160. Ferreira, G. M. N., Monteiro, G. A., Prazeres, D. M. F., et al., (2000) Downstream processing of plasmid DNA for gene therapy and DNA vaccine applications. *Trends Biotechnol.* **18**, 380–388.
161. Prazeres, D. M. F., Monteiro, G. A., Ferreira, G. N. M., et al., (2001) Purification of plasmids for gene therapy and DNA vaccination. *Biotechnol. Ann. Rev.* **7**, 1–30.
162. Ferreira, G.N.M. (2005) Chromatographic approaches in the purification of plasmid DNA for therapy and vaccination. *Chem. Eng. Technol.* **28**, 1285–1294.

163. Diogo, M. M., Queiroz, J. A., Prazeres, D. M. F. (2005) Chromatography of plasmid DNA. *J. Chrom. A.* **1069**, 3–22.
164. Phue, J., Noronha, S. B., Hattacharyya, R., et al., (2005) Glucose metabolism at high cell density growth of *E. coli* B and *E. coli* K: differences in metabolic pathways are responsible for efficient utilization in *E. coli* B as determined by microarrays and northern blot analyses. *Biotechnol. Bioeng.* **90**, 805–820.
165. Lara, A.R., Caspeta, L., Gosset, G., et al., (2008) Utility of an *Escherichia coli* strain engineered in the substrate uptake system for improved culture performance at high glucose and cell concentrations: an alternative to fed-batch cultures. *Biotechnol. Bioeng.* **99**, 893–901.
166. Soto, R., Caspeta, L., Barrón, B. L., et al., (2011) High cell-density cultivation in batch mode for plasmid DNA vaccine production by a metabolically engineered E. coli strain with minimized overflow metabolism. *Biochem. Eng. J.* **56**, 165–171.
167. Flores, S., Gosset, G., Flores, N., et al., (2002) Analysis of Carbon Metabolism in *Escherichia coli* Strains with an Inactive Phosphotransferase System by ^{13}C Labeling and NMR Spectroscopy. *Metab. Eng.* **4**, 124–137.
168. Knabben, I., Regestein, L., Marquering, F., et al., (2010) High cell-density processes in batch mode of a genetically engineered *Escherichia coli* strain with minimized overflow metabolism using a pressurized bioreactor. *J. Biotechnol.* **150**, 73–79.
169. Cunningham, D. S., Liu, Z., Domagalski, N., et al., (2009) Pyruvate Kinase-Deficient *Escherichia coli* exhibits increased plasmid copy number and cyclic AMP levels. *J. Bacteriol.* **191**, 3041–3049.
170. OW, D. S. W., Lee, D. Y., Tung, H. H., Lin-Chao, S. (2009). Plasmids regulation and systems-level effects on *Escherichia coli* metabolism, in *Systems Biology and Biotechnology of Escherichia coli* (Lee, S. Y., Ed), pp. 273–294. Springer Science, New York.
171. Ow, D. S. W., Lee, R. M. Y., Nissom, P. M., et al., (2007) Inactivating FruR global regulator in plasmid-bearing *Escherichia coli* alters metabolic gene expression and improves growth rate. *J. Biotechnol.* **131**, 261–269.
172. Ow, D. S. W., Yap, M. G. S., Oh, S. K. W. (2009) Enhancement of plasmid DNA yields during fed-batch culture of a *fruR*-knockout *Escherichia coli* strain. *Biotechnol. Appl. Biochem.* **52**, 53–59.
173. Cunningham, D. S., Koepsel, R. R., Ataai, M. M., et al., (2009) Factors affecting plasmid production in *Escherichia coli* from a resource allocation standpoint. *Microb. Cell Fact.* **8**, 27.
174. Cooke, G. D., Cranenburgh, R. M., Hanak, J. A. J., et al., (2001) Purification of essentially RNA free plasmid DNA using a modified *Escherichia coli* host strain expressing ribonuclease A. *J. Biotechnol.* **85**, 297–304.
175. Carnes, A. E., Hodgson, C. P., Luke, J. M., et al., (2009). Plasmid DNA production combining antibiotic-free selection, inducible high yield fermentation, and novel autolytic purification. *Biotechnol. Bioeng.* **104**, 505–515.

Part III

Lower Eukaryotes

Chapter 15

Recombinant Protein Production in the Eukaryotic Protozoan Parasite *Leishmania tarentolae*: A Review

Tomoaki Niimi

Abstract

Leishmania tarentolae is a trypanosomatid protozoan parasite of the gecko, and has been established as a new eukaryotic expression system for the production of recombinant proteins. It seems that a protozoan parasite is a curious choice as the expression host; however, Trypanosomatidae are rich in glycoproteins with a pattern of glycosylation closely related to those in mammals and higher vertebrates. Thus, one of the main advantages of a *L. tarentolae* expression system is the mammalian-type posttranslational modification of target proteins. Although there are few examples of recombinant protein expression using this system, it can be an attractive alternative to using mammalian cells. This chapter presents an overview of the newly developed protein expression system based on *L. tarentolae*.

Key words: Trypanosomatidae, Protozoan parasite, Mammalian-type posttranslational modification, Glycosylation

1. Introduction

The protozoan parasites of the genera *Leishmania* and *Trypanosoma* are members of the family Trypanosomatidae (*Eugelenozoa*, *Kinetoplastida*), which comprises unicellular organisms characterized by the presence of a single flagellum and of a DNA-rich, mitochondria-like organelle, the kinetoplast (1–3). They are causative agents of a wide spectrum of tropical diseases responsible for substantial human and livestock morbidity and mortality. Pathogenic species of *Leishmania* cause a diverse group of diseases, collectively called leishmaniases (4). According to the World Health Organization, more than 12 million people in 88 countries are affected by this disease and 350 million people are at risk of infection (http://www.who.int/leishmaniasis/burden/en/). In addition to their importance

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_15, © Springer Science+Business Media, LLC 2012

as agents of diseases, Trypanosomatidae have also been considered as attractive model organisms for studying some intracellular processes because they have unusual mechanisms of gene expression, such as polycistronic transcription and RNA editing. In these organisms, mRNAs are transcribed as polycistronic precursors that are posttranscriptionally processed into individual mRNAs by *trans*-splicing and polyadenylation (5–7). Thus, regulation of protein expression in these organisms occurs predominantly posttranscriptionally through the structure of the intergenic untranslated regions (UTRs). Furthermore, they have been valuable for studying molecular and cellular phenomena, such as glycosylphosphatidylinositol (GPI) anchoring of proteins, antigenic variation, and telomere organization (8–10). Due to their importance for public health, and as model organisms, methods of genetic manipulation for Trypanosomatidae are well-established (6, 11), which has led to several attempts of heterologous gene expression for studying the immunogenicity and biological activity of proteins from parasites (12). It is also known that Trypanosomatidae are rich in glycoproteins, with a pattern of glycosylation closely related to those of mammals and higher vertebrates (13). Possibly due to their parasitic lifestyle, the oligosaccharide structures of their glycoproteins may be similar to those in mammals; therefore, Trypanosomatidae have emerged as potential hosts for heterologous protein production.

Leishmania tarentolae is a nonpathogenic parasite of the white-spotted wall gecko *Tarentola annularis* (14), and has been established as a new eukaryotic expression system for the production of recombinant proteins with a mammalian-type N-glycosylation pattern (15, 16). Using this system, Breitling et al. reported that the yield of green fluorescent protein was up to 30 mg/L of culture. They also succeeded in obtaining biologically active human erythropoietin (EPO); a molecule which requires glycosylation and appropriate modification for activity (17), showing that the *L. tarentolae* expression system can be considered as an alternative to mammalian cells in cultures. This chapter discusses the advantages and disadvantages of this newly developed expression system based on *L. tarentolae*.

2. Genetic Manipulation

2.1. Media and Growth Conditions

L. tarentolae cells can be cultivated in complex media (brain–heart infusion based or yeast–soybean based) or chemically defined media, both supplemented with hemin, which is essential for growth of *Leishmania* (18–20). Hemin is the chlorinated derivative of the heme molecule and traditionally produced from bovine or porcine blood. Therefore, a complete defined medium without animal-derived products is not yet available. The cost of culture medium is not inexpensive.

L. tarentolae cells require aerobic conditions for development. The cells can be maintained as continuous suspension culture in ventilated tissue culture flasks. Conventional static cultures are incubated in the dark at 26°C. Agitated cultures for protein expression are incubated in an orbital shaker at 140 rpm using Erlenmeyer flasks. It is possible to scale up to larger bioreactors with volumes of 1–30 L. *L. tarentolae* cells can be indefinitely grown in vitro with a doubling time of around 5 h and to high cell densities (approximately 5×10^8 cells per mL) in complex media. The cultivation of *L. tarentolae* is much easier than that of mammalian cells.

2.2. Expression Vectors

In eukaryotes, protein-coding genes are transcribed by RNA polymerase II, whereas RNA polymerase I (Pol I) transcribes the genes that encode the three ribosomal RNAs (rRNAs). In Trypanosomatidae, however, translation of Pol I-transcribed protein-coding genes is possible because of *trans*-splicing of polycistronic pre-mRNAs (21). *Trans*-splicing adds a capped 39 nt spliced leader (or mini-exon) to the 5′ end of the mRNA, which is necessary for RNA transport, stability, and translation efficiency. Thus, in these organisms, the higher transcription rate of Pol I is used to ensure high levels of gene expression. In a *L. tarentolae* expression system, integration of the expression cassette into the chromosomal small subunit rRNA locus (ssu) enables the generation of large numbers of transcripts for constitutive expression. The expression cassette of pLEXSY-2 expression vectors for *L. tarentolae* is flanked by two fragments of the small subunit rRNA locus (5′ and 3′ ssu) for homologous recombination.

The minimal vector for gene expression in *Leishmania* consists of a circular plasmid containing a signal for *trans*-splicing of mRNA, followed by a target gene (22). Better expression can often be obtained if the target gene is followed by the 3′-UTR from a highly expressed gene. Since gene regulation in *Leishmania* occurs mainly posttranscriptionally through intergenic UTRs, the choice of suitable UTRs is important for construction of an efficient expression vector. The pLEXSY-2 vectors contain three optimized UTRs: *utr1* derived from the 0.4 kb intergenic region of *L. tarentolae* adenine phosphoribosyl transferase gene; *utr2* from the 1.4 kb intergenic region of *L. tarentolae* calmodulin cluster containing three tandemly arranged calmodulin genes (*camCB*); and *utr3* from the 1.7 kb intergenic region of the *L. major* dihydrofolate reductase-thymidylate synthase gene, flanking the target and marker gene insertion sites, which provide the *trans*-splicing signal for posttranscriptional mRNA processing.

In a *L. tarentolae* expression system, heterologous proteins can either be expressed cytosolically or secreted into the medium. Alternative cloning strategies result in cytosolic or secretory expression of target genes in pLEXSY-2 vectors. For secretion, a signal sequence derived from secreted acid phosphatase of *L. mexicana*

(lmsap) is utilized in pLEXSY-2 vectors (23); however, the native signal sequence has been used successfully for secretion of human EPO and rat proprotein convertase 4 (15, 24). Although the expression efficiency of recombinant EPO with lmsap was significantly higher than that with the native signal sequence, the usage of the native signal sequence enables the production of natively processed proteins at the N-terminus.

In conjunction with the unique organization of the transcription/translation machinery of Trypanosomatidae, the inducible protein expression system based on *L. tarentolae* can also be used (25). In Trypanosomatidae, expression of genes can be driven by not only endogenous RNA polymerase I or II, but also cointegrated promoters for heterologous RNA polymerase (26). Therefore, a combination of bacteriophage T7 RNA polymerase with its promoter controlled by a repressor element responsive to tetracycline (Tet) is used for the inducible expression system. Target genes are inserted into the expression cassette under the control of a T7 promoter with the bacterial Tet responsive element (TRE) and integrated into the chromosomal ornithine decarboxylase (*odc*) locus of the host *L. tarentolae* strain, which is constitutively expressing T7 RNA polymerase and the Tet repressor under the control of endogenous Pol I. When overexpression of heterologous proteins affects the physiology of the host, the inducible expression system may be valuable.

2.3. Gene Transfer and Selection

Following construction in *Escherichia coli*, the expression plasmid is linearized and integrated into the chromosome of *L. tarentolae* by homologous recombination. *Leishmania* cells can be routinely transfected with plasmid DNA by electroporation and the transfected cells can be selected with antibiotics. Currently, four selectable marker genes are available in this system: neomycin phosphotransferase (neo), hygromycin phosphotransferase (hyg), bleomycin resistance protein (ble), and streptothricin acetyltransferase (sat) that confer resistance to G418, hygromycin, bleomycin, and nourseothricin, respectively. Therefore, up to four genes can be expressed simultaneously and it is applicable, for instance, to produce multisubunit proteins (27). Moreover, further increase of expression levels can be obtained with the additional copies of integrated genes after sequential transfection and selection (15, 28).

3. Posttranslational Modifications

The structures of N-linked oligosaccharides from various pathogenic *Leishmania* and *Trypanosoma* species have been well-investigated because they have been implicated in parasite virulence (29).

According to previous studies, the pattern of N-glycosylation is highly variable in different species and at different life cycle stages (30, 31). Most N-linked glycans in these parasites are of the high-mannose type; however, many trypanosomatids retain the enzymatic machinery for making more complex N-glycan structures. As an example, some glycoproteins of *T. brucei* have complex N-linked glycans, primarily biantennary glycans with terminal a-1, 3-linked galactose units or branched poly-*N*-acetyllactosamine sequences similar to those of higher eukaryotes (32, 33). Some proteins of *T. cruzi* are modified with a-linked galactose, fucose, and even sialic acid residues (34, 35). The N-glycosylation profile of *L. tarentolae* was investigated in recombinant human EPO (15). It was exceptionally homogenous, with higher eukaryote-like biantennary N-glycans and the $Man_3GlcNAc_2$ core structure. N-glycans were fully galactosylated and core-α-1, 6-fucosylated, whereas sialylation was missing. Higher branched, tri- and tetra-antennary glycans were not observed, probably due to a lack of N-acetylglucosaminyl transferase IV activity. Overall, the N-glycosylation pathway of *L. tarentolae* is more similar to those in mammals than those in yeast and insect cells (Fig. 1). Genetic engineering of the *L. tarentolae* host may provide a method for producing sialylated or multiantennary heterologous glycoproteins.

L. tarentolae has the potential to perform other posttranslational modifications typically associated with higher eukaryotes. These include processing of signal sequences, protein folding, and disulfide bridge formation. We successfully produced the disulfide-linked heterotrimeric glycoprotein, laminin, which contained complex structures, such as epidermal growth factor (EGF)-like repeat and coiled-coil domains, in a *L. tarentolae* expression system (27).

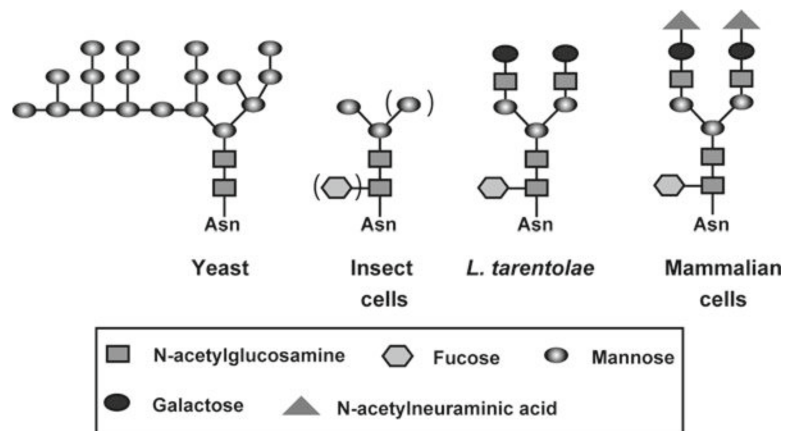


Fig. 1. N-glycosylation patterns of proteins expressed in various expression systems.

Table 1
Comparison of recombinant protein expression systems

System	Growth rate	Productivity	Glycosylation	Protein folding	Availability of genetic systems	Scale-up	Media cost
Prokaryotic							
Bacteria	Very fast	Very high	No	Poor	Very good	Very good	Very low
Eukaryotic							
Yeast	Fast	High	Simple	Fair	Good	Very good	Low
L. tarentolae	Medium	Medium	Mammalian like	Good	Fair	Good	Medium
Insect cells	Slow	High	Non-mammalian like	Very good	Fair	Fair	High
Mammalian cells	Slow	Low	Yes	Very good	Fair	Fair	High

4. Conclusions

Various recombinant protein expression systems have been developed in bacteria, yeast, fungi, insect cells, mammalian cells, transgenic animals, and transgenic plants. Researchers have selected the optimal expression system based on the purpose of the desired protein product. The selection of the expression system depends on numerous factors, such as the expression levels, posttranslational modifications, production cost, availability of genetic systems, and other factors (Table 1) (36, 37). Thus, it is important to understand the advantages and disadvantages of the expression systems. At present, the *L. tarentolae* expression system provides an attractive alternative to mammalian cells for the production of recombinant mammalian proteins. When it is difficult to express a mammalian protein or domain of interest in mammalian cells, it would be advantageous to use the *L. tarentolae* expression system. Furthermore, it can be used for expression of proteins that are difficult to express in bacteria or yeast. Since the *L. tarentolae* expression system has been recently commercialized by Jena Bioscience (http://www.jenabioscience.com), there are still few examples of this recombinant protein expression (15, 24, 27, 28, 38–40). Although further investigation is required to verify a relationship between posttranslational modifications and biological activity of many other proteins, *L. tarentolae* is a promising host for production of proteins with mammalian-type posttranslational modifications. The genomes of several Trypanosomatidae have been sequenced (41–44) and are available from sources, such as TriTrypDB (http://TriTrypDB.org) (45), including the partial sequence of *L. tarentolae*. The availability of the genome sequence should facilitate future improvement of this system for large-scale production of recombinant proteins for industrial and pharmaceutical uses.

Acknowledgments

The preparation of this chapter was supported by a Grant-in-aid (no. 18108003) for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

References

1. Shapiro, T. A., and Englund, P. T. (1995) The structure and replication of kinetoplast DNA. *Annu. Rev. Microbiol.* **49**, 117–143.
2. Morris, J. C., Drew, M. E., Klingbeil, M. M., Motyka, S. A., Saxowsky, T. T., Wang, Z., et al. (2001) Replication of kinetoplast DNA: An update for the new millennium. *Int. J. Parasitol.* **31**, 453–458.
3. Liu, B., Liu, Y., Motyka, S. A., Agbo, E. E., and Englund, P. T. (2005) Fellowship of the rings: the replication of kinetoplast DNA. *Trends Parasitol.* **21**, 363–369.

4. Banuls, A. L., Hide, M., and Prugnolle, F. (2007) *Leishmania* and the leishmaniases: A parasite genetic update and advances in taxonomy, epidemiology and pathogenicity in humans. *Adv. Parasitol.* **64**, 1–109.
5. Simpson, L., and Maslov, D. A. (1999) Evolution of the U-insertion/deletion RNA editing in mitochondria of kinetoplastid protozoa. *Ann. N. Y. Acad. Sci.* **870**, 190–205.
6. Clayton, C. E. (1999) Genetic manipulation of kinetoplastida. *Parasitol. Today* **15**, 372–378.
7. Clayton, C., and Shapira, M. (2007) Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. *Mol. Biochem. Parasitol.* **156**, 93–101.
8. Krakow, J. L., Hereld, D., Bangs, J. D., Hart, G. W., and Englund, P. T. (1986) Identification of a glycolipid precursor of the *Trypanosoma brucei* variant surface glycoprotein. *J. Biol. Chem.* **261**, 12147–12153.
9. Blackburn, E. H. (1991) Structure and function of telomeres. *Nature* **350**, 569–573.
10. Borst, P., and Rudenko, G. (1994) Antigenic variation in African trypanosomes. *Science* **264**, 1872–1873.
11. Beverley, S. M. (2003) Protozomics: Trypanosomatid parasite genetics comes of age. *Nat. Rev. Genet.* **4**, 11–19.
12. Fernandez-Robledo, J. A., and Vasta, G. R. (2010) Production of recombinant proteins from protozoan parasites. *Trends Parasitol.* **26**, 244–254.
13. Ferguson, M. A. (1999) The structure, biosynthesis and functions of glycosylphosphatidylinositol anchors, and the contributions of trypanosome research. *J. Cell Sci.* **112**, 2799–2809.
14. Simpson, L., Frech, G. C., and Maslov, D. A. (1996) RNA editing in trypanosomatid mitochondria. *Methods Enzymol.* **264**, 99–121.
15. Breitling, R., Klingner, S., Callewaert, N., Pietrucha, R., Geyer, A., Ehrlich, G., et al. (2002) Nonpathogenic trypanosomatid protozoa as a platform for protein research and production. *Protein Expr. Purif.* **25**, 209–218.
16. Basile, G., and Peticca, M. (2009) Recombinant protein expression in *Leishmania tarentolae*. *Mol. Biotechnol.* **43**, 273–278.
17. Wasley, L. C., Timony, G., Murtha, P., Stoudemire, J., Dorner, A. J., Caro, J., et al. (1991) The importance of N- and O-linked oligosaccharides for the biosynthesis and in vitro and in vivo biologic activities of erythropoietin. *Blood* 77, 2624–2632.
18. Pal, J. K., and Joshi-Purandare, M. (2001) Dose-dependent differential effect of hemin on protein synthesis and cell proliferation in *Leishmania donovani* promastigotes cultured in vitro. *J. Biosci.* **26**, 225–231.
19. Fritsche, C., Sitz, M., Weiland, N., Breitling, R., and Pohl, H. D. (2007) Characterization of the growth behavior of *Leishmania tarentolae*: A new expression system for recombinant proteins. *J. Basic Microbiol.* **47**, 384–393.
20. Fritsche, C., Sitz, M., Wolf, M., and Pohl, H. D. (2008) Development of a defined medium for heterologous expression in *Leishmania tarentolae*. *J. Basic Microbiol.* **48**, 488–495.
21. Lee, M. G., and Van der Ploeg, L. H. (1997) Transcription of protein-coding genes in trypanosomes by RNA polymerase I. *Annu. Rev. Microbiol.* **51**, 463–489.
22. Curotto de Lafaille, M. A., Laban, A., and Wirth, D. F. (1992) Gene expression in *Leishmania*: Analysis of essential 5′ DNA sequences. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2703–2707.
23. Wiese, M., Ilg, T., Lottspeich, F., and Overath, P. (1995) Ser/Thr-rich repetitive motifs as targets for phosphoglycan modifications in *Leishmania mexicana* secreted acid phosphatase. *EMBO J.* **14**, 1067–1074.
24. Basak, A., Shervani, N. J., Mbikay, M., and Kolajova, M. (2008) Recombinant proprotein convertase 4 (PC4) from *Leishmania tarentolae* expression system: purification, biochemical study and inhibitor design. *Protein Expr. Purif.* **60**, 117–126.
25. Kushnir, S., Gase, K., Breitling, R., and Alexandrov, K. (2005) Development of an inducible protein expression system based on the protozoan host *Leishmania tarentolae*. *Protein Expr. Purif.* **42**, 37–46.
26. Wirtz, E., Hartmann, C., and Clayton, C. (1994) Gene expression mediated by bacteriophage T3 and T7 RNA polymerases in transgenic trypanosomes. *Nucleic Acids Res.* **22**, 3887–3894.
27. Phan, H. P., Sugino, M., and Niimi, T. (2009) The production of recombinant human laminin-332 in a *Leishmania tarentolae* expression system. *Protein Expr. Purif.* **68**, 79–84.
28. Gazdag, E. M., Cirstea, I. C., Breitling, R., Lukes, J., Blankenfeldt, W., and Alexandrov, K. (2010) Purification and crystallization of human Cu/Zn superoxide dismutase recombinantly produced in the protozoan *Leishmania tarentolae*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **66**, 871–877.
29. Descoteaux, A., and Turco, S. J. (1999) Glycoconjugates in *Leishmania* infectivity. *Biochim. Biophys. Acta* **1455**, 341–352.
30. Parodi, A. J. (1993) N-glycosylation in trypanosomatid protozoa. *Glycobiology* **3**, 193–199.

31. McConville, M. J., Mullin, K. A., Ilgoutz, S. C., and Teasdale, R. D. (2002) Secretory pathway of trypanosomatid parasites. *Microbiol. Mol. Biol. Rev.* **66**, 122–154.
32. Zamze, S. (1991) Glycosylation in parasitic protozoa of the trypanosomatidae family. *Glycoconj. J.* **8**, 443–447.
33. Mehlert, A., Zitzmann, N., Richardson, J. M., Treumann, A., and Ferguson, M. A. (1998) The glycosylation of the variant surface glycoproteins and procyclic acidic repetitive proteins of *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **91**, 145–152.
34. Couto, A. S., Goncalves, M. F., Colli, W., and de Lederkremer, R. M. (1990) The N-linked carbohydrate chain of the 85-kilodalton glycoprotein from *Trypanosoma cruzi* trypomastigotes contains sialyl, fucosyl and galactosyl (alpha 1–3)galactose units. *Mol. Biochem. Parasitol.* **39**, 101–107.
35. Parodi, A. J., Labriola, C., and Cazzulo, J. J. (1995) The presence of complex-type oligosaccharides at the C-terminal domain glycosylation site of some molecules of cruzipain. *Mol. Biochem. Parasitol.* **69**, 247–255.
36. Sodoyer, R. (2004) Expression systems for the production of recombinant pharmaceuticals. *BioDrugs* **18**, 51–62.
37. Cha, H. J., Dalal, N. N., and Bentley, W. E. (2005) Secretion of human interleukin-2 fused with green fluorescent protein in recombinant *Pichia pastoris*. *Appl. Biochem. Biotechnol.* **126**, 1–11.
38. Soleimani, M., Mahboudi, F., Davoudi, N., Amanzadeh, A., Azizi, M., Adeli, A., et al. (2007) Expression of human tissue plasminogen activator in the trypanosomatid protozoan *Leishmania tarentolae*. *Biotechnol. Appl. Biochem.* **48**, 55–61.
39. Dortay, H., and Mueller-Roeber, B. (2010) A highly efficient pipeline for protein expression in *Leishmania tarentolae* using infrared fluorescence protein as marker. *Microb. Cell Fact.* **9**, 29.
40. Ben-Abdallah, M., Bondet, V., Fauchereau, F., Beguin, P., Goubran-Botros, H., Pagan, C., et al. (2011) Production of soluble, active acetyl serotonin methyl transferase in *Leishmania tarentolae*. *Protein Expr. Purif.* **75**, 114–118.
41. Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D. C., et al. (2005) The genome of the African trypanosome *Trypanosoma brucei*. *Science* **309**, 416–422.
42. El-Sayed, N. M., Myler, P. J., Bartholomeu, D. C., Nilsson, D., Aggarwal, G., Tran, A. N., et al. (2005) The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science* **309**, 409–415.
43. Ivens, A. C., Peacock, C. S., Worthey, E. A., Murphy, L., Aggarwal, G., Berriman, M., et al. (2005) The genome of the kinetoplastid parasite, *Leishmania major*. *Science* **309**, 436–442.
44. Peacock, C. S., Seeger, K., Harris, D., Murphy, L., Ruiz, J. C., Quail, M. A., et al. (2007) Comparative genomic analysis of three *Leishmania* species that cause diverse human disease. *Nat. Genet.* **39**, 839–847.
45. Aslett, M., Aurrecoechea, C., Berriman, M., Brestelli, J., Brunk, B. P., Carrington, M., et al. (2009) TriTrypDB: A functional genomic resource for the Trypanosomatidae. *Nucleic Acids Res.* **38**, D457-462.

Chapter 16

Expression of Multisubunit Proteins in *Leishmania tarentolae*

Marisa Sugino and Tomoaki Niimi

Abstract

Heterologous gene expression in mammalian cells is the first choice for the production of recombinant proteins when posttranslational modifications affect the biological activity of target proteins. However, the expression efficiency of mammalian cells is relatively low compared to other expression systems, such as *Escherichia coli* or yeast. Recently, a novel protein expression system based on *Leishmania tarentolae*, a protozoan parasite of gecko, was developed. This system allows not only easy handling like *E. coli* and yeast, but also full eukaryotic protein folding and the mammalian-type posttranslational modifications of target proteins. Here, we attempt to produce recombinant human laminin (LM)-332, a large heterotrimeric glycoprotein, in the *L. tarentolae* expression system. A recombinant strain harboring three subunits of LM-332 efficiently formed a heterotrimer and secreted it into the medium. The purified rLM-332 showed similar cell adhesion activity to rLM-332 purified from mammalian cells, indicating its proper folding and assembly. In this chapter, we describe a detailed protocol for multiple gene expression in the *L. tarentolae* expression system.

Key words: Trypanosomatidae, Protozoa, Laminin, Basement membrane

1. Introduction

Over the past few decades, various recombinant protein expression systems have been developed using bacteria, yeast, plant, insect, and mammalian cells. Researchers have chosen systems based on their downstream purpose. When posttranslational modifications affect the biological activity of target proteins, heterologous gene expression in mammalian cells is the first choice. However, the expression efficiency of mammalian cells is relatively low compared to other expression systems, such as *E. coli* or yeast. Recently, a novel protein expression system based on *Leishmania tarentolae*,

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_16, © Springer Science+Business Media, LLC 2012

a unicellular eukaryotic protozoan parasite, has been established (1–3). Compared to mammalian cell cultures, *L. tarentolae* has the advantage of a higher growth rate with doubling time of 4–5 h in agitated cultures. A unique feature of recombinant proteins produced in *L. tarentolae* is the mammalian-type N-glycosylation pattern. Since the *L. tarentolae* expression system has been used for successful expression of various biologically active proteins, it is an alternative to mammalian cells for production of recombinant proteins (1, 4, 5).

Laminins (LMs) are large glycoproteins that are an integral part of the structural architecture of basement membranes. They consist of three subunits, α, β, and γ chains, which bind to each other via disulfide bonds to form a cross-shaped structure (6–8). To date, 5 α, 3 β, and 3 γ chains have been identified that combine into at least 16 heterotrimeric molecules (9). LM-332, which consists of α3, β3, and γ2 chains, is unique among LM family members in that all three subunits have truncated short arms, making LM-332 the smallest molecule. However, it is difficult to express correctly folded heterotrimeric LMs, even LM-332, in *E. coli* and yeast. Thus, mammalian cells have been used to prepare recombinant LMs for use in biochemical studies. Here, we describe the production of recombinant human LM-332 in the *L. tarentolae* expression system (10).

2. Materials

2.1. Construction of Plasmids

1. pLEXSY-2 expression vectors (Jena Bioscience, Fig. 1): pLEXSY-neo2 encoding the aminoglucoside phosphotransferase, pLEXSY-ble2 encoding the bleomycin resistance gene, and pLEXSY-sat2 encoding the streptothricine acetyltransferase (see Note 1).
2. High-fidelity DNA polymerase.
3. Oligonucleotide primers with appropriate restriction sites.
4. Thermal cycler for polymerase chain reaction.
5. Agarose gel equipment.
6. DNA fragment purification kit (Qiagen).
7. Restriction enzymes.
8. *E. coli* strain DH5α.
9. Luria-Bertani (LB) medium: 10 g/L Bacto-tryptone, 5 g/L bacto yeast extract, and 5 g/L NaCl; autoclave at 121°C for 30 min.
10. LB agar plates containing ampicillin at a final concentration of 100 μg/mL.
11. Miniprep and Midiprep plasmid DNA extraction kits (Qiagen).

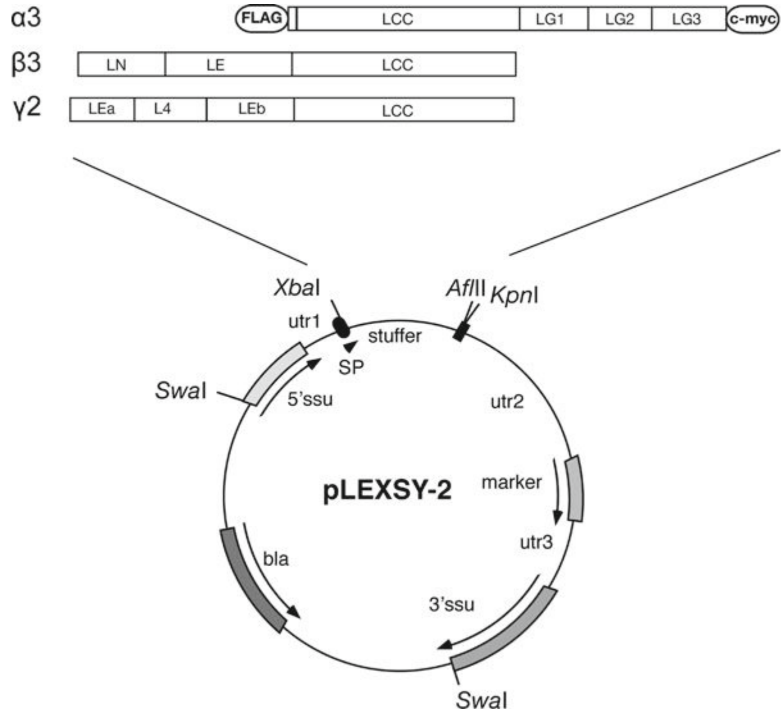


Fig. 1. Construction of expression plasmids for human LM-332 subunits. A map of the pLEXSY-2 vector and the strategy used for plasmid construction are summarized. FLAG and c-*myc* tags were added to the N- and C-terminal ends of the α3 constructs, respectively. The cDNA sequences encoding human LM-332 subunits without signal sequence were inserted in-frame after the signal peptide (SP) of *L. mexicana* secreted acid phosphatase. LM coiled-coil (LCC), LM globular (LG), LM N-terminal (LN), and LM 4 (L4) domains are shown.

2.2. Leishmania Strain, Growth Media, and Additives

1. *L. tarentolae* strain: LEXSY host P10 (Jena Bioscience).
2. LEXSY-BHI liquid medium: Dissolve 37 g LEXSY BHI powder (Jena Bioscience) in 1 L of deionized water and autoclave for 15 min at 121°C. Cool medium to room temperature (RT) before adding of hemin and penicillin/streptomycin.
3. 500× hemin stock solution: 0.25% (v/v) hemin in 50% (v/v) triethanolamine.
4. 200× penicillin/streptomycin stock solution: 10,000 units of penicillin G sodium salt and 10 mg/mL streptomycin sulfate in 0.85% saline.
5. 1,000× G418 stock solution: 50 mg/mL in water.
6. 1,000× bleomycin stock solution: 100 mg/mL in water.
7. 1,000× nourseothricin stock solution: 100 mg/mL in water.
8. LEXSY-BHI agar plates for clonal selection.
   (a) Mix the following components for five plates (50 mL): 35 mL of 2× LEXSY BHI (74 g/L), 10 mL of inactivated

fetal calf serum, 4 mL of 1 M HEPES, pH 7.4, 1 mL of penicillin/streptomycin, 0.2 mL of hemin, and appropriate antibiotics.

(b) Autoclave 50 mL of 2% (w/v) agar and keep at 55°C.

(c) Pour the medium into the warm agar (total 100 mL), mix gently, and distribute 20 mL per 90-mm plastic Petri dish.

(d) Leave the cover off and dry the plate for 10 min after solidifying.

(e) Use the freshly prepared plates immediately.

2.3. Cell Culture, Protein Expression and Detection

1. T25 ventilated tissue culture flask for suspension culture (see Note 2).
2. 125-mL disposable sterile Erlenmeyer flask with vented cap.
3. Incubator at 26°C (see Note 3).
4. Orbital shaker at 26°C.
5. Electroporation device.
6. Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) equipment.
7. Western blotting equipment.
8. Nitrocellulose membrane.
9. Anti-FLAG M2 (Sigma), LM β3 (Santa Cruz Biotechnology), and LM γ2 (Santa Cruz Biotechnology) antibodies.
10. Horseradish peroxidase (HRP)-conjugated secondary antibody (GE Healthcare).
11. Chemiluminescence substrate for peroxidase.
12. Chemiluminescence imaging system.
13. TBS buffer: 50 mM Tris–HCl, pH 7.4, and 150 mM NaCl.

3. Methods

3.1. Preparation of the Expression Plasmids for Transfection

The pLEXSY-2 vectors are used for constitutive expression of target proteins following integration of the expression cassette into the chromosomal 18S rRNA locus of *L. tarentolae* (Fig. 1).

3.1.1. Polymerase Chain Reaction Amplification of cDNAs

The primers are designed to ensure in-frame cloning of the cDNA of interest into the expression cassette. The cDNAs are amplified either by reverse transcription-PCR (RT-PCR) from total RNA of human keratinocytes or by PCR using plasmid DNA as a template. The amplification is performed using high-fidelity DNA polymerase to ensure minimal mutation of the sequence during the PCR.

3.1.2. Cloning of cDNAs into the Expression Vectors

We usually add an epitope tag at both the 5′ and 3′ ends of the cDNA, as one of the tags may not be accessible to an antibody if the expressed protein is processed by proteases in *Leishmania*. We use the pSecTag2A-FLAG vector for adding FLAG- and c-*myc*-tags to the 5′ and 3′ ends of cDNA, respectively (10). Once the amplified cDNAs are inserted into the pSecTag2A-FLAG (laminin α3 subunit) or pcDNA3.1(+) vector (laminin β3 and γ2 subunits), further amplification is performed. The resulting PCR products are then ligated into the *Xba*I-*Afl*II- or *Xba*I-*Kpn*I-digested pLEXSY-2 vectors.

3.2. Transfection

The expression plasmids containing cDNAs encoding human laminin β3, γ2, and α3 chains are sequentially transfected into *L. tarentolae* by electroporation as follows.

1. Linearize approximately 5 μg of the expression plasmid with *Swa*I restriction enzyme. This treatment generates a 2.9 kb fragment from *E. coli* and a larger fragment representing the linearized expression cassette with the target gene.
2. Purify the expression cassette with an agarose gel extraction.
3. One day before transfection, inoculate *L. tarentolae* cells at a 1:10 dilution in 10 mL of LEXSY-BHI medium in a T-flask.
4. On the day of transfection, the cell density should be about 6×10^7 cells/mL ($OD_{600} = 1.4$).
5. Centrifuge the cells at 2,000 × *g* for 5 min at RT and remove 50% of the supernatant volume.
6. Resuspend the pellet in the remaining medium and incubate on ice for 10 min.
7. Chill off linearized plasmid DNA (1–2 μg) in another tube on ice.
8. Add 350 μL volume of prechilled cells to the tube with DNA and transfer a 0.2-cm electroporation cuvette on ice.
9. Apply a pulse of 450 V with 450 microF.
10. Place back the cuvette on ice for 10 min.
11. Transfer the cells to 5 mL of LEXSY-BHI medium.
12. Incubate the electroporated cells at 26°C for 24 h.

3.3. Clonal Selection by Plating on Agar Plates

For establishment of the expression strain, we always use the clonal selection method by plating on agar plates with appropriate antibiotics as follows (see Notes 4 and 5).

1. Withdraw 0.5–1.5 mL from the transfected 5 mL of culture.
2. Pellet the cells for 1 min at 2,000 × *g* at RT.
3. Remove the supernatant and resuspend the cells in 50–100 μL of residual medium.

4. Carefully spread the resuspended cells onto freshly prepared LEXSY-BHI agar plates supplemented with the appropriate selective antibiotics.
5. Seal the plates with parafilm and incubate upside down at 26°C for 5–10 days.
6. Transfer five to ten colonies carefully using a micropipette tip to 96-well plate with 200 μL of selective medium.
7. After a 24–48-h incubation, expand the clones to 1 mL of selective medium in a 24-well plate and can be used for evaluation of protein expression by Western blotting.

3.4. Verification of Correct Transformants by Western Blotting

1. Apply the cell lysates and/or conditioned medium of the selected clones after clonal selection to an SDS-PAGE.
2. Transfer proteins to nitrocellulose membranes.
3. Block the membrane with 5% (w/v) nonfat dry milk in phosphate-buffered saline (PBS) containing 0.1% (v/v) Tween-20 for 1 h at RT.
4. Incubate the membrane with primary antibodies against the FLAG epitope (1:1,000 v/v), LM β3 (1:2,000 v/v), or LM γ2 (1:2,000 v/v) for 1 h at RT.
5. Wash the membrane three times with PBS containing 0.1% (v/v) Tween-20.
6. Incubate with HRP-conjugated secondary antibody (1:2,000 v/v) for 1 h at RT.
7. Wash the membrane three times with PBS containing 0.1% (v/v) Tween-20.
8. Develop the membrane using chemiluminescence reagents and detect on a chemiluminescence imaging system.

3.5. Storage of *L. tarentolae* Host and Recombinants

1. Add 1.2 mL of autoclaved 80% glycerol to a 15-mL tube.
2. Withdraw 3.6 mL of culture grown in LEXSY-BHI medium from the mid-growth phase (see Note 6).
3. Mix with glycerol and dispense 3 × 1.6-mL aliquots into sterile cryovials.
4. Incubate for 10 min at RT.
5. Incubate for 1 h on ice.
6. Incubate for 1 day at –20°C.
7. Transfer to –80°C for long-term storage.

3.6. Protein Expression and Purification

1. Inoculate approximately 3 mL of the static culture from the recombinant strains harboring the three constructs in 30 mL of LEXSY-BHI medium containing half the required concentration of G418, bleomycin, and nourseothricin.

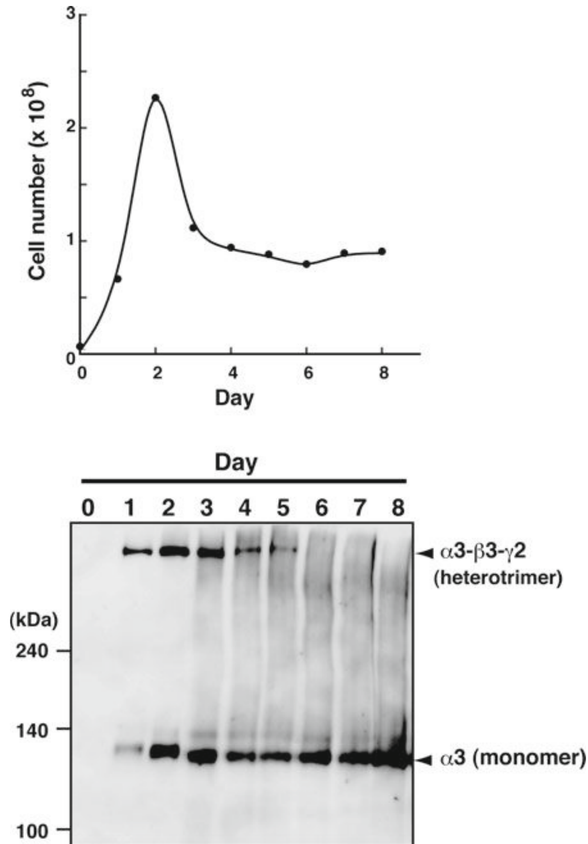


Fig. 2. Production of recombinant human LM-332 in *L. tarentolae* cells. Kinetics of cell growth (*upper panel*) and expression of recombinant LM-332 (*lower panel*) are presented. Plasmids containing cDNA encoding the β3, γ2, and α3 subunits of LM-332 were sequentially transfected into *L. tarentolae* by electroporation. A recombinant strain harboring three constructs was analyzed for the expression of LM-332. Culture medium was analyzed by SDS-PAGE under nonreducing conditions and immunoblotted with anti-FLAG antibody. The bands assumed to be the α3-β3-γ2 heterotrimer and α3 monomer are indicated with arrowheads. Molecular weight markers are indicated to the *left* of the figure.

2. Incubate the cells at 26°C in a 125-mL Erlenmeyer flask with shaking at 140 rpm in an orbital shaker.
3. Check protein expression by Western blotting (Fig. 2) (see Note 7).
4. Harvest the culture medium around 72 h post inoculation by centrifugation at 4,000 × *g* for 15 min.
5. Apply the culture medium to an anti-FLAG M2 affinity column.
6. Wash the column with ten bed volumes of Tris-buffered saline (TBS).
7. Elute the bound proteins with five fractions of FLAG peptide (100 μg/mL) each of half a bed volume.
8. Dialyze the eluate against TBS (see Note 8).

4. Notes

1. A fourth expression vector, pLEXSY-hyg2, encoding hygromycin phosphotransferase is available, which makes it possible to express four genes simultaneously in this system.
2. *L. tarentolae* cells require aerobic conditions. The cells can be maintained continuously in suspension in ventilated tissue culture flasks with regular dilutions at 1:10–1:50.
3. *L. tarentolae* cells must be cultivated in the dark at 26°C because hemin is light sensitive.
4. The frequency of correct integrants is very high in clonal selection by plating on agar plates.
5. The days that colonies begin to appear are different for each clone when using the clonal selection method on agar plates. Some clones can be picked on the fourth day after plating, and some clones need more than 10 days' incubation before they can be picked.
6. The reactivation of *L. tarentolae* strains from glycerol stocks has often failed. Glycerol stock must be prepared from the mid-growth phase culture carefully.
7. It is important to harvest the culture medium in the appropriate growth phase. The LM-332 heterotrimer should be harvested between 2 and 3 days before degradation. If the α3 monomer is required, it may be better to harvest the culture on later days.
8. The final yield of the purified recombinant LM-332 was estimated to be 2.5–5 μg from 30 mL of culture medium. This yield was comparable to that of LM-332-producing mammalian cell lines.

Acknowledgments

This work was supported by a Grant-in-aid (no. 18108003) for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

References

1. Breitling, R., Klingner, S., Callewaert, N., Pietrucha, R., Geyer, A., Ehrlich, G., Hartung, R., Muller, A., Contreras, R., Beverley, S. M., and Alexandrov, K. (2002) Non-pathogenic trypanosomatid protozoa as a platform for protein research and production, *Protein Expr. Purif.* **25**, 209–218.
2. Kushnir, S., Gase, K., Breitling, R., and Alexandrov, K. (2005) Development of an inducible protein expression system based on the protozoan host *Leishmania tarentolae*, *Protein Expr. Purif.* **42**, 37–46.
3. Fritsche, C., Sitz, M., Weiland, N., Breitling, R., and Pohl, H. D. (2007) Characterization of the growth behavior of *Leishmania tarentolae*: A new expression system for recombinant proteins, *J. Basic Microbiol.* **47**, 384–393.

4. Soleimani, M., Mahboudi, F., Davoudi, N., Amanzadeh, A., Azizi, M., Adeli, A., Rastegar, H., Barkhordari, F., and Mohajer-Maghari, B. (2007) Expression of human tissue plasminogen activator in the trypanosomatid protozoan *Leishmania tarentolae*, *Biotechnol. Appl. Biochem.* **48**, 55–61.
5. Basak, A., Shervani, N. J., Mbikay, M., and Kolajova, M. (2008) Recombinant proprotein convertase 4 (PC4) from *Leishmania tarentolae* expression system: purification, biochemical study and inhibitor design, *Protein Expr. Purif.* **60**, 117–126.
6. Tunggal, P., Smyth, N., Paulsson, M., and Ott, M. C. (2000) Laminins: Structure and genetic regulation, *Microsc. Res. Tech.* **51**, 214–227.
7. Li, S., Edgar, D., Fassler, R., Wadsworth, W., and Yurchenco, P. D. (2003) The role of laminin in embryonic cell polarization and tissue organization, *Dev. Cell* **4**, 613–624.
8. Miner, J. H., and Yurchenco, P. D. (2004) Laminin functions in tissue morphogenesis, *Annu. Rev. Cell Dev. Biol.* **20**, 255–284.
9. Aumailley, M., Bruckner-Tuderman, L., Carter, W. G., Deutzmann, R., Edgar, D., Ekblom, P., Engel, J., Engvall, E., Hohenester, E., Jones, J. C., Kleinman, H. K., Marinkovich, M. P., Martin, G. R., Mayer, U., Meneguzzi, G., Miner, J. H., Miyazaki, K., Patarroyo, M., Paulsson, M., Quaranta, V., Sanes, J. R., Sasaki, T., Sekiguchi, K., Sorokin, L. M., Talts, J. F., Tryggvason, K., Uitto, J., Virtanen, I., von der Mark, K., Wewer, U. M., Yamada, Y., and Yurchenco, P. D. (2005) A simplified laminin nomenclature, *Matrix Biol.* **24**, 326–332.
10. Phan, H. P., Sugino, M., and Niimi, T. (2009) The production of recombinant human laminin-332 in a *Leishmania tarentolae* expression system, *Protein Expr. Purif.* **68**, 79–84.

Part IV

Fungi

Chapter 17

Recombinant Protein Production in Yeasts

Diethard Mattanovich, Paola Branduardi, Laura Dato, Brigitte Gasser, Michael Sauer, and Danilo Porro

Abstract

Recombinant protein production is a multibillion-dollar market. The development of a new product begins with the choice of a production host. While one single perfect host for every protein does not exist, several expression systems ranging from bacterial hosts to mammalian cells have been established. Among them, yeast cell factories combine the advantages of being single cells, such as fast growth and easy genetic manipulation, as well as eukaryotic features including a secretory pathway leading to correct protein processing and post-translational modifications. In this respect, especially the engineering of yeast glycosylation to produce glycoproteins of human-like glycan structures is of great interest. Additionally, different attempts of cellular engineering as well as the design of different production processes that are leading to improved productivities are presented. With the advent of cheaper next-generation sequencing techniques, systems biotechnology approaches focusing on genome scale analyses will advance and accelerate yeast cell factories and thus recombinant protein production processes in the near future. In this review we summarize advantages and limitations of the main and most promising yeast hosts, including *Saccharomyces cerevisiae*, *Pichia pastoris*, and *Hansenula polymorpha* as those presently used in large scale production of heterologous proteins.

Key words: Yeast, Heterologous proteins, Expression, Industrial Biotechnology

1. Introduction: the Host Species Commonly Used

High level production of proteins from engineered organisms provides an alternative to protein extraction from natural sources. Natural sources of proteins are often limited, and furthermore the concentration of the desired product is generally low, making its extraction very cost-intensive or even impossible. Besides, extraction might bear the danger of toxic or infectious contamination depending on the natural origin of the protein. With the advent of molecular

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_17, © Springer Science+Business Media, LLC 2012

cloning in the mid-1970s, it became possible to produce foreign proteins in new hosts. Recombinant DNA (rDNA) technologies (genetic, protein, and metabolic engineering) allow the production of a wide range of peptides and proteins from naturally non-producing cells. The first recombinant product on the world market made by means of rDNA was insulin in 1982 (Humulin by Eli Lilly). Since then, more biopharmaceutical proteins such as interferons, erythropoietin, vaccines, and more recently monoclonal antibodies as well as industrial enzymes (for example, used for the treatment of food, feed, detergents, paper-pulp, and health care) followed. Today, recombinant protein production is a multibillion-dollar market. The global sales for biopharmaceutical proteins was 87 billion USD in 2008 and is expected to rise up to 169 billion USD in 2014 (1, 2). The market for industrial enzymes was 5.1 billion USD in 2009 and is predicted to reach 6.5 billion USD in 2013 (3). Indeed, the growing demand in the animal feed and biofuels market, in addition to pharmaceutical enzymes, requires high production of industrial enzymes such as proteases, lipases, and carbohydrases.

In 2009, out of the 151 approved recombinant biopharmaceutical products, 29 monoclonal antibodies contributed to more than 40% of the revenues, followed by vaccines, TNF blockers, hormones such as insulin and erythropoietin. Around 20% of biopharmaceutical proteins are produced in yeast hosts, 30% in *Escherichia coli*, and 50% in higher eukaryotic cells, mainly mammalian cells and hybridomas (4). Currently, nearly all yeast-derived products on the market are produced in *Saccharomyces cerevisiae*; in 2009 the first biopharmaceutical protein produced in a non-*Saccharomyces* yeast was approved by the FDA (Kallikrein inhibitor (Kalbitor by Dynax Inc.) produced in *Pichia pastoris*). Regarding the production of industrial enzymes, the share is almost total of microbial hosts especially *E. coli* (mostly recombinant), *Bacillus subtilis* (natural producer), yeasts (mostly recombinant secretion), and filamentous fungi (often native producers). Especially for the recombinant production of fungal enzymes, secretory expression in yeasts is often the best choice. Despite the fact that a large number of protein drugs are produced in higher eukaryotic cells, there is still a strong interest of biopharmaceutical industry in developing improved microbial production platforms (mainly *E. coli* and yeast systems). This indicates that microbial cells represent convenient and powerful tools for recombinant protein production. In this respect, the dominance of *E. coli* as host for the production of heterologous proteins is clearly a reflection of the quantity and quality of the information available about its genetics, molecular biology, biochemical, physiological, and fermentation technologies. Indeed, recombinant protein fraction can reach up to 50% of total biomass (5) and protocols for high cell density cultivation (HCDC) are established, reaching up to 100 g biomass per liter (6). But, it has to be noted

that bacterial hosts lack the ability of correct post-translational protein processing and, in a lot of cases, heterologous proteins aggregate as inclusion bodies inside the bacterial cells, often causing the proteins to lose their enzymatic activity/3D structure. Due to their incorrect structure any use of such proteins in humans is also excluded, unless they can be correctly refolded in vitro. However, the dogma that *E. coli* cannot secrete or glycosylate recombinant proteins became obsolete due to extensive genetic engineering efforts: both extracellular protein production (reviewed by Nu and Chen (7)) as well as glycosylation (reviewed by Pandhal and Wright (8)) has been reported also for prokaryotic hosts. However, the yields achieved (up to maximum 100 mg/L secreted protein), or the authenticity of the glycans are far from the eukaryotic systems.

Among the microbial eukaryotic hosts systems, yeasts combine the advantages of unicellular organisms (i.e., ease of genetic manipulation and growth) with the capability of a protein processing typical for eukaryotic organisms (i.e., protein folding, assembly, and post-translational modifications), together with the absence of endotoxins as well as oncogenic or viral DNA. Starting from the early 1980s, the majority of recombinant proteins produced in yeasts have been expressed using *S. cerevisiae* (9). As for *E. coli*, this was a reflection of the familiarity of molecular biologists with this yeast combined with the deep knowledge about its genetics, biochemistry, physiology, and fermentation technologies. Furthermore, *S. cerevisiae* is recognized by the American Food and Drug Administration (FDA) as an organism generally regarded as safe (GRAS). However, it shows strong fermentative metabolism and limited recombinant protein productivity. In addition, the proteins produced by *S. cerevisiae* are often hyper-glycosylated and retention of the products within the periplasmic space is frequently observed (10–12), with a consequent partial degradation. These degradation products are generally very difficult to remove from the desired product. Disadvantages such as these have promoted, since the mid 1980s, a search for alternative hosts, trying to exploit the great biodiversity existing among the yeasts, and starting the development of expression systems using the so-called non-conventional yeasts. The most established or prominent examples of non-conventional yeasts are *Hansenula polymorpha* (13–15), *P. pastoris* (reviewed in ref. 16, 17), *Kluyveromyces lactis* (18), *Kluyveromyces marxianus* (19), *Yarrowia lipolytica* (20), *Arxula adeninivorans* (21, 22), *Pichia methanolica* (23), *Schizosaccharomyces pombe* (24), and *Zygosaccharomyces bailii* (25), to mention a few. Some other yeasts, such as the osmotolerant *Zygosaccharomyces rouxii* or xylose-utilizing *Pichia stipitis* have initially been described for protein production (26, 27); however, nowadays their main applications are physiological characterization of salt tolerance or production of organic compounds such as bioethanol, respectively.

The choice of the yeast host is of paramount importance for the success of the whole process. On this topic, different reviews have been published in the last years (15, 28–32). Usually, yeast hosts are divided in two main categories, (a) conventional and non-conventional or (b) Crabtree positive and Crabtree negative. In this respect, *S. cerevisiae* is the only conventional yeast and one of the few Crabtree positive (i.e., producing ethanol under aerobic conditions), together with *Z. rouxii* and *Z. bailii*. With respect to fermentation, *Y. lipolytica* represents an exception as the only non-fermenting yeast among the host species mentioned above. Nevertheless, considering the recent developments of heterologous protein production systems, yeasts should be grouped in two other categories: non-methylotrophic and methylotrophic hosts (Table 1). This review will analyze the advantages and disadvantages of different yeasts according to this grouping.

1.1. Non-Methylotrophic Yeasts

Many of the wild type yeasts listed above, lacking both endotoxins and lytic viruses, are known for their established applications for the production of ethanol, beverages, flavors, enzymes, vitamins, organic and single cell proteins (biomasses). Generally speaking, the main advantage of the non-methylotrophic hosts is related to the familiarity of the molecular biologists, biochemists, microbial physiologists, and fermentation technologists with these yeasts. This is particularly true for *S. cerevisiae* and, to a lesser extent, for *K. lactis* and *Y. lipolytica* for which the entire genome has been fully sequenced. Further, *K. lactis*, like the already mentioned *S. cerevisiae*, is a GRAS organism.

Table 1
Non-methylotrophic and methylotrophic yeast species used for recombinant protein production (for full names and references see text)

Non-methylotrophic	Methylotrophic
S. cerevisiae	*H. polymorpha*
K. lactis	*P. pastoris*
Y. lipolytica	*P. methanolica*
Z. rouxii	*C. boidinii*
Z. bailii	*O. minuta*
S. occidentalis	
P. stipitis	
A. adeninivorans	
K. marxianus	

K. marxianus is a thermotolerant relative of *K. lactis*, which has mainly been utilized to produce endogenous (pectinolytic) or exogenous enzymes (see ref. 33), and exhibits strictly respiratory growth under glucose surplus, thus leading to high biomass concentrations. For many of these yeasts, the tremendous improvements in yeast molecular genetics have yielded a wide range of vectors, selection markers, promoters, terminators, and secretion signals allowing the scientists to understand and solve many problems encountered in the expression of heterologous genes. The success obtained using these yeasts is indicated by the high expression levels that have been obtained in shake-flask cultures. Advanced fermentation techniques (continuous and fed-batch processes) and technologies (computer-controlled fermentations) are also available. An interesting element is also added by the wide range of carbon and energy sources that can be used during the industrial processes by these yeasts (among them: glucose, lactose, maltose, starch, alkanes, and fatty acids). Despite all these considerations, at present, most processes for the production of heterologous proteins by these yeasts are still in the developmental phase. In fact, large-scale processes impose many restrictions and often these expression systems cannot be implemented in large-scale environments.

The amount of genetic, biochemical, physiological information and technologies available for the genus *Zygosaccharomyces* is instead very poor. However, these yeasts seem very promising. Six species, which appear to be evolutionarily quite close to *S. cerevisiae* and not so distant from *K. lactis*, have been classified so far (34). An exceptional resistance to several stresses renders some of the *Zygosaccharomyces* species potentially interesting for industrial purposes. For example *Z. rouxii* is known to be salt tolerant (osmophilic) and *Z. bailii* can tolerate high sugar concentrations (osmotolerant), acidic environments as well as relatively high temperatures of growth (35, 36). Despite being classified as Crabtree positive, ethanol production under aerobic conditions is quite low when compared to that observed from *S. cerevisiae* cells; indeed, these yeasts can grow to high cell densities even without complex fermentation strategies. Developments of these yeasts as hosts for the production of heterologous proteins are quite promising because all of these properties could allow the development of easier and cheaper production strategies.

Y. lipolytica and *A. adeninivorans* represent dimorphic fungi, which are able to change from vegetative growth to formation of (pseudo)hyphae. *A. adeninivorans* is a nitrate-assimilating, thermo- and osmotolerant yeast, which exhibits temperature-dependent polymorphism (change to mycelial structures above 42°C), while *Y. lipolytica* has the capacity to grow on *n*-paraffins and is known as natural secretor of enzymes.

1.2. Methylotrophic Yeasts

Methylotrophic yeast species share a common pathway to metabolize one-carbon compounds, e.g., methanol as carbon and energy source. Initially, this group of yeasts had been employed in the production of feed protein (single cell protein) (37), but soon they attracted the interest as production systems for recombinant proteins (38) because of two major features: they are able to grow to high cell densities even in unsophisticated fermentation processes and, secondly, their high demand for methanol oxidizing enzymes (like alcohol oxidase) endows them with very strong and strictly regulated promoters.

Two species are mainly employed for heterologous protein expression: *P. pastoris* (syn. *Komagataella pastoris*) (reviewed in ref. 16, 17) and *H. polymorpha* (also known as *Pichia angusta*) (reviewed in ref. 15, 22). Additionally, *P. methanolica* (23) and *Candida boidinii* (39, 40) are described in literature as expression systems. More recently another methylotrophic yeast, *Ogataea (Pichia) minuta* was developed as a suitable host for the production of glycoproteins (41). In the following pages, *P. pastoris* and *H. polymorpha* will be discussed in more detail. Although these two yeasts share a common pathway to metabolize methanol, they differ significantly in their specific genetics (42). *P. pastoris* produces two different alcohol oxidases (AOX1 and AOX2), whereas *H. polymorpha* expresses only one methanol oxidase (MOX). The regulation of MOX is somewhat different, as derepression of MOX occurs even in the absence of methanol if there is no glucose present. Furthermore, *H. polymorpha* is able to grow on elevated temperature (42°C) and to assimilate nitrate.

The first generation of expression systems relied on the *AOX*1 promoter (the stronger between the two) or the *MOX* promoter, respectively. With respect to methanol utilization, there are three possible phenotypes for *P. pastoris* compared to two for *H. polymorpha*, namely Mut$^+$ (wild type methanol utilization; both *AOX*1 and *AOX*2 intact), Mut^s (slow methanol utilization; *AOX*1 interrupted, *AOX*2 intact) and Mut$^-$ (deleted methanol utilization, both *AOX*1 and *AOX*2 interrupted). Since methanol is employed as inducer of these expression systems, the methanol utilization phenotype plays an important role with respect to the design of the production process. The technological consequences of these phenotypes will be discussed further ahead.

Eventually, novel promoters were isolated and employed for gene expression, namely the formaldehyde dehydrogenase (*FLD*) promoter (43, 44) and the glyceraldehyde 3-phosphate dehydrogenase (*GAP*) promoter (45) of *P. pastoris* and the formate dehydrogenase (*FMD*) promoter of *H. polymorpha* (46). These promoters are either regulated, as the *FLD* promoter (nitrogen and methanol regulated) or constitutive as the *GAP* promoter. The type of regulation again plays a major role in the design of the fermentation process.

A major drawback of all methylotrophic yeasts as compared to *S. cerevisiae* was for many years the lack of fundamental knowledge on their genetics and molecular biology. Therefore only a limited number of genetic engineering tools were available. However, the genome sequences of *P. pastoris* (47, 48) and *H. polymorpha* (49) recently became public and more will become soon (e.g., *Z. bailii*, PB personal communication), thus accelerating the development of the methylotrophic expression platforms.

2. Transformation and Vector Systems

K. lactis, and *Y. lipolytica* are heterothallic while *P. pastoris*, *P. methanolica*, and *H. polymorpha* are homothallic. Most *Zygosac-charomyces* species appear homothallic, but details still have to be clarified. For the *K. lactis* and *Y. lipolytica* heterothallic strains, the genetic manipulations described for *S. cerevisiae* can be carried on with good efficiency, while for the homothallic strains, complementation and tetrad analyses pose different problems.

The expression of a foreign protein in yeast consists of, firstly, cloning of a foreign protein-coding DNA sequence within an expression cassette containing a yeast promoter and transcriptional termination sequences; secondly, transformation and stable maintenance of this DNA fusion in the host. Cloning of a heterologous gene into the cited yeast hosts can be carried out by three different approaches: spheroplast preparation (50), the lithium acetate method (51) and electroporation (52). The lithium acetate method and electroporation are the methods of choice today. Efficiency of transformation is clearly strain dependent and detailed studies should be carried out when a high efficiency of transformation is required.

Transformants can be selected by (a) complementation of auxotrophic markers (i.e., for example *URA*3, *LEU*2, *TRP*1, *HIS*3, *HIS*4, *ADE*1 being the most commonly used), (b) dominant markers (genes conferring resistance to the presence of antibiotics such as G418, phleomycin, hygromycin, Zeocin, or others in the culture medium), or (c) the so-called autoselection systems. For an industrial process, the expression system should be genetically stable without selection pressure. It is strongly advised against the use of antibiotics during the production process, and auxotrophic marker systems would require chemically defined minimal media (which would not be unusual for a large-scale process), but at high cell densities would also show the tendency of cross feeding between marker-containing and auxotrophic cells that have lost the marker.

For *S. cerevisiae*, the *LEU*2 gene and the G418 resistance gene are the two most popular markers used, whereas for *P. pastoris*, *HIS*4 and Zeocin resistance are mainly employed. Autoselection systems are based on the expression of a vital gene/activity in host

strains lacking such a gene/activity, like *URA*3 in a Δ*fur*1 (uracil phosphoribosyl transferase), *FBA*1 (fructose biphosphate aldolase) in a Δ*fba*1 backgrounds. These systems ensure that plasmid selection is maintained irrespective of culture conditions. While dominant markers can be used with any yeast strain, auxotrophic markers and autoselection systems require the availability of the respective deletion strains. The availability of such deletion strains depends on the abundance of molecular biology research on these species. Table 2 summarizes the most commonly used methods of transformation and selection in the presented yeast expression systems, as well as the progress of genetic engineering and the availability of tools for genetic manipulation.

The heterologous gene can be introduced into the yeast host cells by means of an integrative plasmid as well as by autonomous or episomal circular plasmids. In the first case, the heterologous gene fate is the integration into the chromosomal DNA by means of recombination events. Directed integration requires homology of the DNA introduced with a chromosomal locus. On the other side, heterologous recombination may occur at random positions. In the other cases, the heterologous gene will be replicated due to the replication of a circular plasmid. For the autonomous plasmids, replication is governed by the autonomous replicating sequences (ARS), while the episomal plasmids are based on endogenous circular yeast plasmids like the 2 μm plasmid of the yeast *S. cerevisiae*.

Ideally, an expression vector should be stably maintained in the host cells without the need of any selection pressure and (assuming a direct correlation between gene dosage and expression level) with a high copy number per cell. Cloning of centromeric sequences (CEN) in autonomous plasmids yields stable centromeric plasmids, but those are replicated in one or two copies per cell. ARS-based plasmids without centromeric sequences are generally very unstable and easily lost by growing the host cells without selective pressure. Episomal plasmids are generally maintained with a high copy number per cell (from 1 to 10 up to 100 copies per cell). Unfortunately, endogenous plasmids have been isolated and characterized only for few yeast hosts: *S. cerevisiae, K. lactis, Y. lipolytica, Z. rouxii,* and *Z. bailii* (see Table 2).

Finally, the copy number per cell of the integrative plasmids is related to the cloning strategy. Multiple integration can be obtained by targeting the cloned gene to the ribosomal DNA cluster, constructing concatamers of the expression cassette or simply by chance. Nevertheless, it is very important to underline that a direct correlation between gene dosage and level of expression is not a general feature, but strongly depends on the heterologous protein to be produced. Cloning and transformation of the host cells is only the first step required for the production of a heterologous protein. It has been shown that lower level of expression (transcription and translation) could give higher yield in the production of the heterologous proteins.

Table 2
Characteristics of the most commonly used yeast expression systems: available tools for genetic engineering (including genome sequence), progress in engineering of N-glycosylation and commercial suppliers of the expression platforms

Type	Species	Transformation method	rDNA propagation modus	Ribosomal integration	Dominant selection	Auxotrophic selection	Glycoengineering	Genome sequence	Commercial Supplier
Non-methylotrophic	*S. cerevisiae*	Lithium acetate	Episomal (ARS, CEN, 2 µm plasmids) or genome integration	Yes	G418, hygromycin, nourseothricin	*leu2, ade2, ura3, arg4*, …	Discontinued at *och1 mnn1*→ Man5GlcNAc2	1996 (53)	
	K. lactis	Lithium acetate	Episomal or genome integration	Yes (54)	G418, hygromycin, *amdS*	*ura3, leu2, trp1*	In progress *och1 mnn1*→Man(9–11) GlcNAc2 (55)	Genolevures (56)	New England Biolabs
	P. stipitis	Electroporation	Genome integration	Yes (57)	Hygromycin		n.a.	2007 (58)	
	Y. lipolytica	Lithium acetate or electroporation	Genome integration	Yes (59)	*Phleomycin, hygromycin, SUC2*	*leu2, ura3, lys5, ade1*	In progress *och1*→ Man8-GlcNAc2 (60)	Genolevures (56)	
	Z. rouxii	Lithium acetate plus electroporation	Episomal (centromer or pSR1) or genome integration	n.a.	G418	*ura3, leu2, ade2*	n.a.	Genolevures (56)	
	Z. bailii	Lithium acetate or electroporation	Episomal (ARS, CEN, 2 µm plasmids) or genome integration	Yes (25)	G418, hygromycin, nourseothricin	*leu2*	n.a.	n.a.	
	A. adeninivorans		Genome integration	Yes (61)	G418, hygromycin	*leu2, trp1, lys2, ura3*	n.a.	n.a.	

(continued)

Table 2 (continued)

Type	Species	Transformation method	rDNA propagation modus	Ribosomal integration	Dominant selection	Auxotrophic selection	Glycoengineering	Genome sequence	Commercial Supplier
Methylotrophic	*P. methanolica*	Electroporation	Genome integration, up to 90% non-homologous	n.a.	n.a.	*ade2*	n.a.	n.a.	Invitrogen
	P. pastoris	Electroporation	Genome integration	Yes (62)	Zeocin, G418, FLD	*his4, ade2, ura3, met, arg4*	Complex human N-glycans (63)	2009 (47, 48)	Invitrogen
	H. polymorpha	Electroporation	Genome integration	Yes (64)	G418, hygromycin	*leu2, ura3, arg1, ade1*	In progress *och1, alg3*→Man3-GlcNAc2 (65)	2003 (49)	
	O. minuta	Lithium chloride plus electroporation	Genome integration	n.a.	G418	*ura3, ade1*	In progress *och1*→Man5-GlcNAc2 (41)	n.a.	

The vectors used are often hybrids between yeast-derived and bacterial sequences. The bacterial fraction of these vectors bears an origin of replication for the chosen bacteria (essentially *E. coli*) and selection markers (e.g., antibiotic resistance). Of course, the hybrid fraction is introduced for an easier manipulation of the vector itself.

With regards to the heterologous gene and its codon bias index, it should be noted that codon usage does not appear to be an essential parameter for high production levels even if, in some cases, higher productions have been obtained.

Concluding, no general rules are known to illustrate the transformation efficiency, vector stability, and copy number, although factors such as vector composition, host strain, transformation method, and selective pressure might influence them.

3. Promoters

One of the important key factors for the production of heterologous proteins is the transcription efficiency of the heterologous gene(s) for which a large variety of heterologous and homologous as well as constitutive and inducible yeast promoters are available. Generally, homologous promoters originating from the yeast species used as host are preferred, as heterologous promoters often do not yield good efficiency of expression. This is especially the case for promoters requiring a certain induction strategy.

On one side, inducible promoters can be useful to maintain yeast cultures in an expression off mode during the cell growth phase minimizing selection for non-expressing mutant cells. Such a selection can occur as a result of the added metabolic burden placed on cells expressing high levels of a foreign protein or the toxic effect of a foreign protein on the cells.

On the other side, constitutive promoters are preferred since their use does not imply the development of complex fermentation strategies and/or the use of specific inducers, which could be expensive or could interfere with the isolation of the final product.

In the methylotrophic yeasts, still most processes rely on promoters of the methanol utilization enzymes for recombinant protein production; however, promoters of genes coding for core metabolic enzymes are also used. Table 3 shows some of the most used promoters.

Table 3
Most important promoters used for gene expression in yeasts (31)

Type	Species	Promoters: Constitutive	Galactose induced	Lactose induced	Ethanol induced	Inducible by fatty acids and alkanes	Starch induced	Xylose induced	Methanol induced
Non-methylotrophic	*S. cerevisiae*	GAPDH, PGK, TPI, ENO, α-MP, TEF	GAL1–10, GAL7		ADH2				
	K. lactis	PGK		LAC4	ADH4				
	S. occidentalis	GAM1					AMY1, GAM1		
	Y. lipolytica	TEF, RPS7 XPR2/hp4d				POX2, POT1, ICL1			
	Z. rouxii	GAPDH							
	Z. bailii	TPI							
	P. stipitis							XYL1	
Methylotrophic	*P. pastoris*	GAP, TEF, PGK							AOX1, FLD1
	H. polymorpha								MOX
	C. boidinii								AOD1, DAS
	O. minuta	GAP							AOX1
	P. methanolica								AUG1

Gene nomenclature: α-MP, alpha-mating factor; ADH2, ADH4, alcohol dehydrogenase; AMY1, α-amylase; AOX1, AUG1, AOD1 and MOX, alcohol oxidase in species shown; ENO, enolase; FLD1, formaldehyde dehydrogenase; GAL1-10 and GAL7, Upstream Activating Sequence of the GAL1-10 and GAL7 promoters; GAM1, glucoamylase; GAPDH, GAP, glyceraldehyde-3-phosphate dehydrogenase; LAC4, β-galactosidase; PGK, phosphoglycerate kinase; RPS7, ribosomal protein S7; TEF, translation elongation factor-1a; TPI, triose phospate isomerase; POX2, acyl-CoA oxidase 2; POT1, 3-oxo-acyl-CoA thiolase); ICL1, isocitrate lyase; XPR2, alkaline extracellular protease (inducible by peptones); hp4d, derivative of XPR2

4. Cytoplasmic Versus Secreted Expression

Heterologous proteins are in almost all cases either targeted to the cytoplasm or secreted, preferably into the culture supernatant. While it is not possible to design the optimum expression strategy a priori, some guidelines can be given.

Cytoplasmic expression often leads to very high expression levels, as the potential limitations of the secretion pathways are not involved. Very high expression levels have been reported, e.g., for rubber tree hydroxynitrile lyase (22 g/L; (66)) or tetanus toxin fragment C (12 g/L; (67)). However, several negative aspects are associated with the expression in the cytoplasm. For example, breaking the cells requires additional steps during the downstream processing of the product. The cells must be disrupted usually by high pressure homogenization (and yeast cell walls are known to be very robust), then the lysate needs to be clarified, and the product is only a fraction of the total soluble cellular constituents, so that additional purification steps must be employed. In many cases protein folding and processing – mainly disulfide bond formation – pose a severe limitation for cytoplasmic expression. Some authors have described the formation of recombinant proteins as insoluble aggregates (inclusion bodies) in the cytoplasm of yeasts (68–71). These inclusion bodies may be processed as those obtained in *E. coli* (71).

Alternatively, proteins can be targeted to be secreted essentially by adding an appropriate secretion signal sequence. The current strategies for secretion engineering are mainly related to enhancing protein quantity by (1) engineering the protein folding and quality control system, and (2) engineering the intracellular protein trafficking pathway, as well as enhancing protein quality by (3) minimizing the post-secretory proteolytic degradation, and (4) engineering the post-translational glycosylation pathway (reviewed in ref. 72).

The most commonly used secretion leader sequence targeting newly synthesized polypeptides to the secretory pathway is the *S. cerevisiae* α-mating factor (α-MF) prepro leader (even with non-*Saccharomyces* yeasts). High secretion levels have been described for *S. cerevisiae*, *P. pastoris*, *H. polymorpha*, *K. lactis,* and *Z. bailii* by using the α-MF leader. Several authors observed optimum secretion when the entire processing sequence of prepro α-MF was employed. After signal peptide removal, the pro sequence is cleaved by the Kex2 proteinase in the Golgi, and subsequently two Glu-Ala dipeptides are removed by the Ste13 dipeptidase. However, the dipeptidase cleavage may not be quantitative resulting in a non-homogenous product, which is problematic especially for pharmaceutical applications. Additionally, other yeast signal sequences have been employed (e.g., *S. cerevisiae* invertase, *K. lactis* pre-sequence of the α-subunit

of the K1 killer toxin or *P. pastoris* acid phosphatase – the latter often with lower yields), or the heterologous secretion leader of the gene to express was utilized, which functioned efficiently in several cases even with human genes (e.g., human serum albumin; (73)).

The efficiency of oxidative protein folding within the endoplasmic reticulum (ER) seems to be highly dependent on the product, and is in most cases the rate-limiting step in recombinant protein secretion in yeasts. Several authors describe that significant amounts of product may be retained in the cell even when targeted to secretion (66, 74–76). Unfolded proteins in the ER induce a conserved pathway termed "unfolded protein response" (UPR, recently reviewed by Kohno, (77) and Mori, (78)). Activation of UPR due to protein overproduction has been described for many yeast species including *S. cerevisiae*, *Y. lipolytica*, *P. pastoris*, and *H. polymorpha*. Strategies to overcome secretory bottlenecks including overexpression of ER resident and cytosolic chaperones and folding enzymes have been reviewed recently (79). According to Kauffman et al. (74) in *S. cerevisiae* the UPR is necessary for the cell to decrease the level of intracellular product and to resume normal growth. Disruption of the central regulator for UPR results in a significantly decreased secretory capacity of the cells (80). That suggests that at least some of the components regulated by the UPR are involved in general secretion. In many cases, constitutive up-regulation of the UPR leads to an improved secretion of a heterologous protein (80, 81). However, as discussed here, other secretion limitations may apply as well, so that UPR up-regulation alone is not always sufficient to improve protein secretion.

It was shown for filamentous fungi that genes encoding secreted proteins are specifically down-regulated in response to secretion stress (RESS) (82). While this seems to be a universal concept in filamentous fungi, it seems to be true only in part for yeasts. Generally, yeasts secrete a significantly lower number of endogenous proteins compared to filamentous fungi, which may render RESS unnecessary in yeasts. The low number of secreted proteins makes yeast an attractive host for secretory protein production, as it facilitates subsequent purification (discussed below) and leads to less proteolytic degradation of the product. The secretomes of *K. lactis* and *P. pastoris* have been analyzed experimentally and show only very few or no proteases being secreted in the supernatant, respectively (48, 83). For other yeasts, bioinformatics tools can be used to predict extracellular proteins within the genome sequence, in order to choose the appropriate host.

After being correctly folded and released from the ER and the Golgi, the product may still be retained within the periplasmic space of the host cell. This effect was mainly observed with *S. cerevisiae*, but this may be due to the larger amount of data available for this yeast species. The distribution of a product between the culture supernatant and the periplasm seems to be product dependent.

Modification of the cell wall by deletion of the *GAS*1 gene strongly improved the release of IGF1 into the supernatant in *S. cerevisiae* (84), of *Rhizopus oryzae* lipase in *P. pastoris* (85), and of two heterologous enzymes in *Z. bailii* (86). Additionally, in the case of *R. oryzae* lipase secretion could further be enhanced by combining the *gas1* deletion with constitutive UPR induction (87), indicating that the secretory pathway can be blocked at different sites simultaneously.

One of the putative post-translational modifications that a eukaryotic protein can encounter is represented by N- and/or O-linked glycosylation. This aspect appears to be significant, since approximately 70% of the approved therapeutic proteins are actually glycoproteins, and their pharmacodynamic behavior is influenced by said modifications. In addition, glycosylation results to be critical also for many industrial enzymes, since it can influence the optimization of their enzymatic activity. The glycosylation patterns of yeasts (*S. cerevisiae* as well as other yeasts like *P. pastoris* and *H. polymorpha*) have been reviewed before (88). Briefly, the initial steps of N-glycosylation are very similar if not even identical in yeasts and mammals. A core N-glycan (consisting of 2 *N*-acetylglucosamines and five mannose residues) is transferred to the nascent polypeptide in the ER, using the Asn-X-Ser/Thr consensus sequence. After ER quality control, the nascent proteins are then transferred to the Golgi apparatus, where N-glycan processing differs markedly between yeasts and higher eukaryotes. Generally, the glycan chains of yeasts are less complex compared to those of higher eukaryotes, being limited to the addition of mannose and mannosylphosphate sugars, and tend to be larger in size. *P. pastoris* has been reported to show less hypermannosylation than *S. cerevisiae* and similar observations were made for *H. polymorpha* (for a review, see ref. 89). One reason to this is the different endowment of the yeasts with mannosyltransferases, as has been elucidated from their genome sequences. Attempts to prevent hypermannosylation by inactivation of the responsible mannosyltransferases Och1 and Mnn1 led to severe growth defects in *S. cerevisiae* (90, 91), but was feasible in other yeasts (see refs. in Table 2). Therefore glycoengineering focused mainly on *P. pastoris* (see refs. 92 and 63). Construction of a glycoengineered yeast strain capable of expressing glycoproteins with humanized N-glycosylation structures involved a total of 17 genetic modifications, including deletion of genes of the endogenous pathway, overexpression of suited glycosyltransferases, and finally overexpression of six biosynthetic genes for supply of galactose and sialic acid, which are the terminal sugars in human N-glycans (summarized in ref. 63). These yeasts possess an enormous potential in protein-based therapy as well as in other fields of biotechnological applications (93–95). In combination with extensive folding and secretion engineering, high level secretion of up to 1 g/L full-length antibodies

with correct human N-glycans has been achieved (96). So far, the modification of yeast O-glycosylation (reviewed in ref. 97) has not been successfully performed apart from one exception (98), but will remain a challenge for the next few years.

As a general rule, most proteins that are secreted in their native environment will be more easily produced in an active form, if secreted from the recombinant host. Problems can be encountered with translocation and/or folding. If secretion and/or folding appear to be the bottleneck of production, an increase of transcription by the use of a stronger promoter or the increase of gene dosage will not improve the yield, but may even lead to a reduction of productivity and can pose a severe stress on the host cells. A general technological problem encountered with secretion systems is the harvesting of the supernatants from high cell density fermentations (up to 500 g/L wet biomass). The possible technical solutions for centrifugation of dense cultures at large scale are limited (e.g., decanters) and involve high investment costs. As an alternative, capturing the product out of the total culture broth by STREAMLINE chromatography (99) or crossflow microfiltration has been suggested.

Since the first use of yeasts as hosts for the production of heterologous proteins, several research teams have addressed the problems of endogenous proteases production and developed protease deficient strains. Strains carrying such mutations in the major cellular and/or extracellular proteases are available in different yeast collections and for most platforms on the market. However, also in this case, no general rules can be established in relation to the efficiency of such strains.

From a chronological point of view, increases in the secretion levels of some heterologous proteins from milligrams to grams per liter have been obtained over the first two decades of production. These improvements were obtained mainly thanks to the development of better fermentation processes, which include both an increased concentration of viable cell and/or a reduced protein degradation, and adjustment of the expression vector system. Additionally, secretory pathway engineering based on general knowledge contributed to enhanced production. Looking forward, the main contribution in this respect will come from the yeast systems biotechnology approach, which enables targeting of yet unidentified bottlenecks.

5. Yeast Systems Biotechnology

The unexpectedly fast progress in genome sequencing over the last decade has provided an invaluable source of information on the physiology of microorganisms, including a comprehensive overview on different cellular processes. The modeling of these processes,

together with extensive analyses of "omics data" (i.e., genomics, transcriptomics, proteomics, interactomics, metabolomics, …) to address ideally all metabolic processes of a cell, led to the concept of systems biology (100). These models offer the opportunity to predict cellular processes and are therefore regarded as highly valuable resources for strain optimization (101, 102).

Systems biotechnology is not a purely academic research area as the quantitative description of microbial production cell lines is also already of interest for biotechnological industry. Indeed, by way of example, for the production of heterologous proteins, up to 30-fold increased production in *E. coli* (103), 17-fold more secreted protein in *Bacillus megaterium* (104), 1.4-fold increased secretion in *Aspergillus niger* (105), 1.2-fold increased secretion in *Aspergillus oryzae,* (106) and 2.5-fold higher titer in CHO cells (107) have been described. The importance of systems biotechnology for industrial applications other than protein production has been well documented in the recent years (108–110) too. Applications of genome-wide technologies in yeasts are scarce in the field of recombinant protein production. Approaches, developments, and perspectives have been recently reviewed by Graf et al. (111). Some of the rare examples analyzing cellular responses due to protein over-production are reported for *P. pastoris* (112–115) and *K. lactis* (18). The analysis of the cellular proteome during a fermentation of a chymosin expressing *K. lactis* strain indicated stress during protein production (up-regulation of Hsp26p and Sod2p; (18)), whereas the *P. pastoris* work outlines how environmental factors such as temperature, osmolarity, and pH affect protein expression and secretion on a transcriptomic (112, 113, 115) and proteomic level (114, 115). Alternatively, metabolic flux analyses of protein-producing yeasts were performed. These focused, on the one hand, on the synthesis of high levels of intracellular human superoxide dismutase in *S. cerevisiae* (116), and, on the other, on core metabolic processes of *P. pastoris* during high cell density production processes (117). Both studies revealed an increased demand of energy as metabolic burden upon recombinant protein production, and suggested strain engineering as major objective to overcome this bottleneck. However, so far, for yeast cell factories hardly any reports about actual strain engineering strategies for improved productivity resulted out of all these studies, which may be attributed to the still young age of this platform. One exception is 2.5-fold increase of secretion of a Fab antibody fragment achieved following the overexpression of secretion factors identified by a transcriptomic analysis (118). Based on these results, a similar approach has been applied for a comparative analysis of the effect of the expression of the same antibody fragment in five different hosts of biotechnological importance, namely *S. cerevisiae*, *P. pastoris*, *Trichoderma reesei*, *E. coli*, and *Pseudoalteromonas haloplanktis*, growing in chemostat culture at different temperatures. Data obtained show that in three organisms out of five, a reduction in growth

temperature was reflected on an enhanced recombinant protein production. Although this beneficial effect of temperature has already been demonstrated in the past, the present article provides more reliable data based on transcriptomic data and chemostat cultivation at specific growth rates (119). Summarizing, we should underline that more progress in this field has been made for bacteria and mammalian cells than for yeasts yet. Despite this lag, strain improvement based on systems biotechnology approaches is catching up very fast in the yeast sector.

6. Physiological/ Metabolic Basis for Process Design

Efficient production of a heterologous protein in yeast requires the integrated development of the following items:

1. The genetic determinants ensuring efficient transcription, translation, and the desired localization of the heterologous protein;
2. A host strain optimized for the physiological determinants that maximize the potential of the genetic determinants;
3. A fast, efficient, and reproducible fermentation process;
4. A fast and cheap downstream process.

It is quite obvious that these elements influence each other crosswise and cannot be considered alone. The genetic and physiological determinants are biological parameters, while the fermentation process determines the environment of the producing cells, thus feeding back on the physiology. Furthermore, the fermentation process obviously influences the downstream process by requiring suitable means for biomass separation of high cell density cultures, as well as for purification of rather small amounts of the desired proteins out of a complex matrix.

Generally, heterologous overexpression of proteins causes stress on the cells possibly impeding the production process from the beginning (120). Therefore, one has to set out to reduce negative influences or any detrimental stress reaction of the cells as much as possible, by the best choice of the interacting biological and technical parameters. For example, the knowledge about UPR (see above) is not only being used to engineer cells with improved secretory capacity but also considered during process development. New techniques allow deep understanding of the host cell physiology in order to adapt it to the required process or vice versa.

One approach is to examine expression profiles of stressed cells versus unstressed cells on transcriptomic, proteomic, or metabolomic level, with the aim to better understand the cellular stress reactions. For instance, the expression differences of the methylotrophic yeasts

Table 4
Percentage of the purification costs on the total costs for fermentation bioproducts

Bioproducts	DSP (%)
Single cell protein/biomass	1–5
Organic acids	10–50
Extracellular enzymes	10–30
Antibiotics	20–50
Recombinant proteins	70–95

H. polymorpha and *P. pastoris* upon change of the carbon source from glucose or glycerol (initial batch phase, repression) to methanol (induction, fed-batch phase) as common cultivation conditions were assessed by microarray hybridization (112, 121) and flux analysis (122, 123). Another approach regards the development of quick but sensitive methods to follow the expression profile of specific genes during any desired process (124).

High expression levels have been obtained in shake-flask cultures. However, at present, most large-scale processes for the production of biopharmaceutical proteins in yeasts are still in the development phase. The core of the biotechnological process based on yeast hosts is the stirred tank bioreactor. Any promising result obtained by flask cultures that cannot be reproduced in the bioreactor is – from an industrial point of view – meaningless. The ideal fermentation process should be as short and cheap as possible. These fermentation processes generally require medium to large size reactor vessels (several liters to some 10 m^3), a high amount of aeration and, in turn, a high amount of electricity. Yeast metabolism generates a large amount of heat and thus requires efficient cooling. Yeasts also release high levels of CO_2 into the environment and generally produce low aqueous titers, requiring high cell densities and large volumes, and therefore elaborate recovery procedures. The main costs of the overall process are related to the substrate, the cost of the energy required, ecological costs (e.g., costs of waste disposal) and, above all, costs of product purification. In this respect, Table 4 reports the average influence of the purification costs (DSP: Down Stream Processing) on the total production costs.

Generally speaking, batch or fed-batch are the techniques of choice to obtain high production of recombinant biomass, and hence high production of recombinant proteins, but continuous processes are coming more and more into focus.

7. Process Development for Recombinant Protein Production with Yeasts

As said before, the most common fermentation strategy for recombinant microorganisms is fed-batch, with the aim to obtain high biomass concentrations. The design of an optimized fed-batch protocol depends to a great extent on the physiological parameters of the expressing strain. The major points to consider are: (1) is the host strain Crabtree positive or negative on the desired substrate? And, (2) is the expression of the product constitutive or regulated and, if regulated, by which means?

If the host strain is Crabtree positive, the biomass yield (gram of dry biomass/gram of substrate consumed) in the batch phase will be comparatively low (0.1–0.2 g/g), and the by-products produced (mainly ethanol) should be metabolized to a large extent before starting the fed-batch, which results in an extended time required for the batch phase. The feed rate then needs to be limited so that further aerobic fermentation is prevented. Usually this results in growth rates below 0.1/h. The optimum feed rate must not exceed a critical limit of specific sugar uptake rate, and is best optimized in continuous chemostat cultures (125). The second important concern for optimizing the feed rate is of course product formation, as it is for Crabtree negative yeasts.

The main group of Crabtree negative yeasts employed for recombinant product formation is represented by the methylotrophic yeasts. As described before, the first generation of expression systems developed for methylotrophic yeasts employed alcohol oxidase promoters, which are induced by methanol and repressed by glucose and many other carbon sources. Therefore, methanol is mainly employed as carbon and energy source and as well as inducer. Quite frequently, feedback loop control strategies employing a methanol sensor to maintain a certain methanol concentration (up to 10 g/L) are described. Several groups postulated that high product yields can only be achieved when such methanol concentration is maintained (126, 127). However, this approach will not allow to control the growth rate at a desired low level as in a carbon limited culture. Therefore, methanol limited fed-batch cultures are employed and have been described to yield higher product titers as compared to the above described feedback loop-controlled cultures, as they allow for longer process times with slower biomass growth (16).

Enfors and coworkers have summarized different process strategies for methanol based protein production with *P. pastoris* (128). While standard methanol limited fed-batch (MLFB) leads to significant cell death and lysis, and consequently the release of proteases into the culture broth, resulting in partial degradation of the product. Among the alternative strategies, temperature limited fed-batch (TLFB) turned out to be most productive, avoiding cell death and proteolytic degradation of the product.

Another special case is the use of strains with deleted alcohol oxidases. As described above, most methylotrophic yeasts will either have a Mut^{+} or a Mut^{-} phenotype except for *P. pastoris*, which carries two alcohol oxidases in the wild type: AOX1, the major fraction of alcohol oxidase, and AOX2, which is expressed at much lower levels. Deletion of *AOX*1 leads to the Mut^{s} phenotype, which allows for a slow growth on methanol, so that these strains can be cultivated on a controlled methanol concentration at a low growth rate. There is still controversy in the literature regarding the benefit of Mut^{s} strains for protein production. Kim et al. (129) have described a threefold higher productivity of *Coprinus cinereus* peroxidase in a Mut^{+} strain compared to a Mut^{s} strain. While specific productivity of an antibody ScFv fragment was higher in a Mut^{+} strain, higher final product titers were reached in a Mut^{s} host using a feedback loop-controlled methanol fed-batch strategy (130). Alternatively, Mut^{s} strains and Mut^{-} strains may be cultivated using mixed feed protocols employing a controlled methanol concentration for induction and a limited feed of glycerol as the carbon source (131, 132).

Generally speaking, product formation can be positively growth related (specific product formation rate increases with specific growth rate), growth indifferent or negatively growth related (specific product formation rate decreases with increasing specific growth rate). Positive growth relation has been observed in a number of different process types and can be regarded as a paradigm for heterologous protein production in yeast (75, 133–135). Deviation from this pattern is mainly observed for secreted proteins and may be rather due to limitations in the secretion machinery at high specific growth rates (see ref. 136). Therefore, at first sight, optimization should mean increasing the specific growth rate. However, the most important parameter for judging an industrial process is not the specific product formation rate of the cells but the productivity per fermenter volume and time (defining the costs for using the fermentation plant), and the final product titer as a major parameter for the efforts of the initial purification steps.

The volumetric productivity (or space time yield, STY) can easily be calculated by dividing the total product achieved by the actual volume and the time consumed, and plotted against the fermentation time so that an optimum can be deduced, which is usually reached at an earlier time point than the highest product titer is achieved. Predictive modeling of fed-batch processes has enabled to optimize processes for STY and product titer (135). Essentially the maximization of STY requires an initial growth phase at maximum specific growth phase followed by a phase of continuous decrease of growth rate. Figure 1 displays modeling of different types of fed-batch processes, illustrating the physiological and technical constraints. The major technical limitations are biomass concentration which should not exceed 100 g/L dry

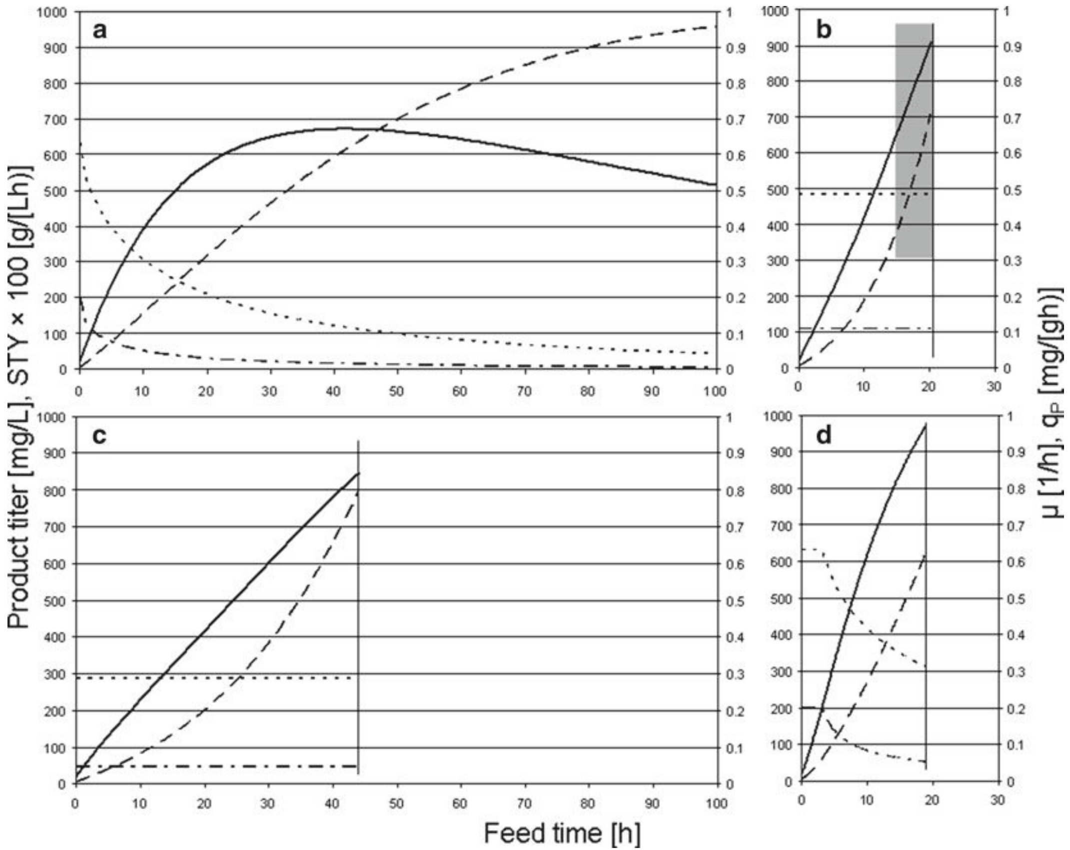


Fig. 1. Model prediction of different fed-batch strategies for production of secreted proteins. The predictions are based on the model described by Maurer et al. (134), assuming a maximum specific productivity $q_p = 1$ mg/(gh). *Curve* STY; *dashed curve* product concentration; *dashed dotted curve* specific growth rate (μm); *dotted curve* specific productivity (q_p). (**a**) Constant slow feed (standard process). (**b**) Fast exponential feed (μ = 0.1/h). *Shaded area*: high oxygen and cooling demand. (**c**) Slow exponential feed (μ = 0.05/h). (**d**) Optimized feed according to Maurer et al. (134). The *vertical lines* mark the accumulation of 100 g/L yeast dry mass in the model processes.

cell weight for production of secreted proteins to allow reasonable cell separation, and heat and oxygen transfer which essentially depend on the substrate feed rate. Therefore the modeled processes are terminated at 100 g/L biomass, and the final stages of exponentially fed cultures are shaded as they are problematic due to limitations of cooling capacity and oxygen transfer. Biological limitations are mainly seen in low specific productivity at low growth, leading to long process times with slow product accumulation in standard processes. The optimal solution begins with an exponential feed phase, followed by linearly increasing feed, so that high oxygen and cooling demand is avoided and rather high specific productivities are maintained while keeping the process time short.

Most protein production processes employing yeasts are based on fed-batch protocols. Continuous processes are not often applied. The reasons for that are divergent: genetic instabilities may lead to

a loss of productivity, the risk of a contamination may be higher, and especially the biopharmaceutical industry fears an unclear regulatory situation concerning batch definition and process lifetime. However, many recombinant expression systems are extremely stable, e.g., when the heterologous gene is integrated into the host genome or when plasmids with stabilizing sequences are employed (see ref. 137). From a theoretical point of view, significantly higher STY can be achieved in continuous culture, therefore this approach should be considered as an attractive alternative to fed-batch. Several authors have achieved promising results by applying continuous cultivation to different yeast species like *S. cerevisiae* (137, 138) or *P. pastoris* (134). However, it should be noted that product instability in the culture conditions might increase in a continuous process, and deleterious effects of the product itself, or its formation, on the host cells may lead to increased problems in continuous culture compared to fed-batch. Therefore it will be necessary to select the optimum process type for every new product.

8. Conclusions and Outlook

The development of a new product begins with the choice of a production host. The main criteria that should be taken into consideration are (1) quality, (2) quantity, (3) yield of the product, and (4) STY of the production process, which are directly connected to the final application of the compound.

In this respect, the key to the success of genetic engineering for enhancing product formation is to identify all the possible limiting steps involved in product biosynthesis. The production of heterologous proteins mainly combines genetic engineering and microbial physiology with the objective of increasing the specific production rate of a desired product. Such an approach is often hampered by the lack of knowledge of the production pathway and its dynamic profile in the producing cells. Therefore, detailed physiological studies are required for a successful production of heterologous proteins, with a special emphasis on studies devoted to the identification of the different physiological determinants that could maximize the potential of the genetic determinants. As recently reviewed by Graf et al. (111), yeast systems biotechnology is covering this gap. In this respect, a major focus of future work should be the quantitative understanding of molecular principles behind protein synthesis, modification, and secretion, derived from basic production strains as well as mutants and rationally engineered strains.

Direct approaches aiming at the extension of substrate range utilization (e.g., starch, lactose, melibiose, xylose, arabinose), improvement of productivity and yield (e.g., avoiding the formation of

by-products), improvement of process performance (e.g., modulating the flocculation process), improvement of cellular properties (e.g., alleviating glucose repression, or the Crabtree effect, modulating the secretion process), or improving product quality (e.g., preventing proteolytic degradation or modulating the glycosylation patterns) are also possible by metabolic engineering applications. Metabolic engineering involves a direct approach to the application of rDNA technology for strain improvement defined as follows: "The directed improvement of product formation or cellular properties through the modification of specific biochemical reaction(s) or the introduction of new one(s) with the use of recombinant DNA technology" (139).

Acknowledgments

This work has been supported by the Austrian BMWFJ, BMVIT, SFG, Standortagentur Tirol and ZIT through the Austrian FFG-COMET- Funding Program and by FAR 2010 to DP and SYSINBIO to PB.

References

1. Goodman, M. (2009) Market watch: Sales of biologics to show robust growth through to 2013. *Nat. Rev. Drug Discov.* **8**, 837.
2. Macdonald, G. (2009) Bio drugs to dominate top ten list by 2014. In-Pharma Technologist. http://www.in-pharmatechnologist.com/Industry-Drivers/Bio-drugs-to-dominate-top-ten-list-by-2014. Accessed 15 October 2010.
3. Freedonia Group (2009) *World Enzymes to 2013*. The Freedonia Group, Inc.
4. Ferrer-Miralles, N., Domingo-Espín, J., Corchero, J., Vázquez, E., and Villaverde, A. (2009) Microbial factories for recombinant pharmaceuticals. *Microb. Cell Fact.* **8**, 17.
5. Panda, A.K. (2003) Bioprocessing of therapeutic proteins from the inclusion bodies of *Escherichia coli*. *Adv. Biochem. Eng. Biotechnol.* **85**, 43–93.
6. Tripathi, N.K., Sathyaseelan, K., Jana, A.M., and Rao, P.V.L. (2009) High yield production of heterologous proteins with *Escherichia coli*. *Defence Science Journal* **59**, 137–146.
7. Ni, Y., and Chen, R. (2009) Extracellular recombinant protein production from *Escherichia coli*. *Biotechnol. Lett.* **31**, 1661–1670.
8. Pandhal, J., and Wright, P.C. (2010) N-Linked glycoengineering for human therapeutic proteins in bacteria. *Biotechnol. Lett.* **32**, 1189–1198.
9. Hitzeman, R.A., Hagie, F.E., Levine, H.L., Goeddel, D.V., Ammerer, G., and Hall, B.D. (1981) Expression of a human gene for interferon in yeast. *Nature* **293**, 717–722.
10. Reiser, J., Glumoff, V., Kalin, M., and Ochsner, U. (1990) Transfer and expression of heterologous genes in yeasts other than *Saccharomyces cerevisiae*. *Adv. Biochem. Eng. Biotechnol.* **43**, 75–102.
11. Romanos, M.A., Scorer, C.A., and Clare, J.J. (1992) Foreign gene expression in yeast: a review. *Yeast* **8**, 423–488.
12. Porro, D., Venturini, M., Brambilla, L., Alberghina, L., and Vanoni, M. (2000) Relating growth dynamics and glucoamylase excretion of individual *Saccharomyces cerevisiae* cells. *J. Microbiol. Methods* **42**, 49–55.
13. Sudbery, P.E., Gleeson, M.A., Veale, R.A., Ledeboer, A.M., and Zoetmulder, M.C. (1988) *Hansenula polymorpha* as a novel yeast system for the expression of heterologous genes. *Biochem. Soc. Trans.* **16**, 1081–1083.
14. van Dijk, R., Faber, K.N., Kiel, J.A., Veenhuis, M., and van der Klei, I. (2000) The methylotrophic yeast *Hansenula polymorpha*: a

versatile cell factory. *Enzyme Microb. Technol.* **26**, 793–800.

15. Gellissen, G., Kunze, G., Gaillardin, C., Cregg, J.M., Berardi, E., Veenhuis, M., and van der Klei, I. (2005) New yeast expression platforms based on methylotrophic *Hansenula polymorpha* and *Pichia pastoris* and on dimorphic *Arxula adeninivorans* and *Yarrowia lipolytica* - a comparison. *FEMS Yeast Res.* **5**, 1079–1096.
16. Cregg, J.M., Cereghino, J.L., Shi, J., and Higgins, D.R. (2000) Recombinant protein expression in *Pichia pastoris. Mol. Biotechnol.* **16**, 23–52.
17. Macauley-Patrick, S., Fazenda, M.L., McNeil, B., and Harvey, L.M. (2005) Heterologous protein production using the *Pichia pastoris* expression system. *Yeast* **22**, 249–270.
18. van Ooyen, A., Dekker, P., Huang, M., Olsthoorn, M., Jacobs, D., Colussi, P., and Taron, C. (2006) Heterologous protein production in the yeast *Kluyveromyces lactis. FEMS Yeast Res.* **6**, 381–392.
19. Fonseca, G.G., Heinzle, E., Wittmann, C., and Gombert, A.K. (2008) The yeast *Kluyveromyces marxianus* and its biotechnological potential. *Appl. Microbiol. Biotechnol.* **79**, 339–354.
20. Madzak, C., Gaillardin, C., and Beckerich, J.M. (2004) Heterologous protein expression and secretion in the non-conventional yeast *Yarrowia lipolytica*: a review. *J. Biotechnol.* **109**, 63–81.
21. Boer, E., Piontek, M., and Kunze, G. (2009) Xplor 2--an optimized transformation/expression system for recombinant protein production in the yeast *Arxula adeninivorans. Appl. Microbiol. Biotechnol.* **84**, 583–594.
22. Stockmann, C., Scheidle, M., Dittrich, B., *et al.* (2009) Process development in *Hansenula polymorpha* and *Arxula adeninivorans*, a reassessment. *Microb. Cell Fact.* **8**: 22.
23. Raymond, C.K., Bukowski, T., Holderman, S.D., Ching, A.F., Vanaja, E., and Stamm, M.R. (1998) Development of the methylotrophic yeast *Pichia methanolica* for the expression of the 65 kilodalton isoform of human glutamate decarboxylase. *Yeast* **14**, 11–23.
24. Takegawa, K., Tohda, H., Sasaki, M., *et al.* (2009) Production of heterologous proteins using the fission-yeast (*Schizosaccharomyces pombe*) expression system. *Biotechnol. Appl. Biochem.* **53**, 227–235.
25. Dato, L., Branduardi, P., Passolunghi, S., Cattaneo, D., Riboldi, L., Frascotti, G., Valli, M., and Porro, D. (2010) Advances in the development of *Zygosaccharomyces bailii* as host for biotechnological productions and construction of the first auxotrophic mutant. *FEMS Yeast Res.* **10**, 894–908.
26. Ogawa, Y., Tatsumi, H., Murakami, S., Ishida, Y., Murakami, K., Masaki, A., Kawabe, H., Arimura, H., Nakano, E., Motai, H., *et al.* 1990. Secretion of *Aspergillus oryzae* alkaline protease in an osmophilic yeast, *Zygosaccharomyces rouxii. Agric. Biol. Chem.* **54**, 2521–2529.
27. Den Haan, R. and Van Zyl, W.H. (2001) Differential expression of the *Trichoderma reesei* beta-xylanase II (xyn2) gene in the xylose-fermenting yeast *Pichia stipitis. Appl. Microbiol. Biotechnol.* **57**, 521–527.
28. Buckholz, R.G., and Gleeson, M.A. (1991) Yeast systems for the commercial production of heterologous proteins. *Biotechnology (N Y)* **9**, 1067–1072.
29. Gellissen, G., and Hollenberg, C.P. (1997) Application of yeasts in gene expression studies: a comparison of *Saccharomyces cerevisiae*, *Hansenula polymorpha* and *Kluyveromyces lactis* – a review. *Gene* **190**, 87–97.
30. Dominguez, A., Ferminan, E., Sanchez, M., Gonzalez, F.J., Perez-Campo, F.M., Garcia, S., Herrero, A.B., San Vicente, A., Cabello, J., Prado, M., Iglesias, F.J., Choupina, A., Burguillo, F.J., Fernandez-Lago, L., and Lopez, M.C. (1998) Non-conventional yeasts as hosts for heterologous protein production. *Int. Microbiol.* **1**, 131–142.
31. Cereghino, G.P., and Cregg, J.M. (1999) Applications of yeast in biotechnology: protein production and genetic analysis. *Curr. Opin. Biotechnol.* **10**, 422–427.
32. Boer, E., Steinborn, G., Kunze, G., and Gellissen, G. (2007) Yeast expression platforms. *Appl. Microbiol. Biotechnol.* **77**, 513–523.
33. Rocha, S.N., Abrahao-Neto, J., Cerdan, M.E., Gonzalez-Siso, M.I. and Gombert, A.K. (2010) Heterologous expression of glucose oxidase in the yeast *Kluyveromyces marxianus. Microb. Cell Fact.* **9**: 4.
34. Kurtzman, C.P. (2003) Phylogenetic circumscription of *Saccharomyces*, *Kluyveromyces* and other members of the *Saccharomycetaceae*, and the proposal of the new genera *Lachancea*, *Nakaseomyces*, *Naumovia*, *Vanderwaltozyma* and *Zygotorulaspora. FEMS Yeast Res.* **4**, 233–245.
35. Makdesi, A.K., and Beuchat, L.R. (1996) Evaluation of media for enumerating heat-stressed, benzoate-resistant *Zygosaccharomyces bailii. Int. J. Food Microbiol.* **33**, 169–181.
36. Sousa, M.J., Miranda, L., Corte-Real, M., and Leao, C. (1996) Transport of acetic acid

in *Zygosaccharomyces bailii*: effects of ethanol and their implications on the resistance of the yeast to acidic environments. *Appl. Environ. Microbiol.* **62**, 3152–3157.

37. Wegener, G.H., and Harder, W. (1987) Methylotrophic yeasts–1986. *Antonie van Leeuwenhoek* **53**, 29–36.
38. Thill, G., Davis, G., Stillman, C., *et al.* (1987) The methylotrophic yeast *Pichia pastoris* as a host for heterologous protein production. In: *Microbial growth on C1 compounds.* Eds: (van Verseveld, H., Duine, W., Nijhoff, J.A., eds.) Dordrecht, pp. 289–296.
39. Yurimoto, H. (2009) Molecular basis of methanol-inducible gene expression and its application in the methylotrophic yeast *Candida boidinii*. *Biosci. Biotechnol. Biochem.* **73**, 793–800.
40. Yurimoto, H., and Sakai, Y. (2009) Methanol-inducible gene expression and heterologous protein production in the methylotrophic yeast *Candida boidinii*. *Biotechnol. Appl. Biochem.* **53**, 85–92.
41. Kuroda, K., Kobayashi, K., Tsumura, H., Komeda, T., Chiba, Y., and Jigami, Y. (2006) Production of Man5GlcNAc2-type sugar chain by the methylotrophic yeast *Ogataea minuta*. *FEMS Yeast Res.* **6**, 1052–1062.
42. Hartner, F.S., Ruth, C., Langenegger, D., *et al.* (2008) Promoter library designed for fine-tuned gene expression in *Pichia pastoris*. *Nucleic Acids Res.* **36**, e76.
43. Shen, S., Sulter, G., Jeffries, T.W., and Cregg, J.M. (1998) A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast *Pichia pastoris*. *Gene* **216**, 93–102.
44. Resina, D., Serrano, A., Valero, F., and Ferrer, P. (2004) Expression of a *Rhizopus oryzae* lipase in *Pichia pastoris* under control of the nitrogen source-regulated formaldehyde dehydrogenase promoter. *J. Biotechnol.* **109**, 103–113.
45. Waterham, H.R., Digan, M.E., Koutz, P.J., Lair, S.V., and Cregg, J.M. (1997) Isolation of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. *Gene* **186**, 37–44.
46. Hollenberg, C.P., and Gellissen, G. (1997) Production of recombinant proteins by methylotrophic yeasts. *Curr. Opin. Biotechnol.* **8**, 554–560.
47. De Schutter, K., Lin, Y.C., Tiels, P., *et al.* (2009) Genome sequence of the recombinant protein production host *Pichia pastoris*. *Nat. Biotechnol.* **27**, 561–566.
48. Mattanovich, D., Graf, A., Stadlmann, J., *et al.* (2009) Genome, secretome and glucose transport highlight unique features of the protein production host *Pichia pastoris*. *Microb. Cell. Fact.* **8**, 29.
49. Ramezani-Rad, M., Hollenberg, C.P., Lauber, J., *et al.* (2003) The *Hansenula polymorpha* (strain CBS4732) genome sequencing and analysis. *FEMS Yeast Res.* **4**, 207–215.
50. Burgers, P.M., and Percival, K.J. (1987) Transformation of yeast spheroplasts without cell fusion. *Anal. Biochem.* **163**, 391–397.
51. Gietz, R.D., and Woods, R.A. (2002) Transformation of Yeast by the LiAc/ss Carrier DNA/PEG. *Method. Meth. in Enzymol.* **350**, 87–96.
52. Sanchez, M., Iglesias, F.J., Santamaria, C., and Dominguez, A. (1993) Transformation of *Kluyveromyces lactis* by electroporation. *Appl. Environm. Microbiol.* **59**, 2087–2092.
53. Goffeau, A., Barrell, B.G., Bussey, H., *et al.* (1996) Life with 6000 genes. *Science* **274**, 563–547.
54. Bergkamp, R.J., Kool, I.M., Geerse, R.H., and Planta, R.J. (1992) Multiple-copy integration of the alpha-galactosidase gene from Cyamopsis tetragonoloba into the ribosomal DNA of *Kluyveromyces lactis*. *Curr. Genet.* **21**, 365–370.
55. Liu, B., Gong, X., Chang, S., *et al.* (2009) Disruption of the *OCH1* and *MNN1* genes decrease N-glycosylation on glycoprotein expressed in *Kluyveromyces lactis*. *J. Biotechnol.* **143**, 95–102.
56. Dujon, B., Sherman, D., Fischer, G., *et al.* (2004) Genome evolution in yeasts. *Nature* **430**, 35–44.
57. Klabunde, J., Kunze, G., Gellissen, G., and Hollenberg, C.P. (2003) Integration of heterologous genes in several yeast species using vectors containing a *Hansenula polymorpha*-derived rDNA-targeting element. *FEMS Yeast Res.* **4**, 185–193.
58. Jeffries, T., Grigoriev, I., Grimwood, J., *et al.* (2007) Genome sequence of the lignocellulose-bioconverting and xylose-fermenting yeast *Pichia stipitis*. *Nat. Biotechnol.* **25**, 319–326.
59. Le Dall, M.T., Nicaud, J.M., and Gaillardin, C. (1994) Multiple-copy integration in the yeast *Yarrowia lipolytica*. *Curr. Genet.* **26**, 38–44.
60. Song, Y., Choi, M.H., Park, J.N., Kim, M.W., Kim, E.J., Kang, H.A. and Kim, J.Y. (2007) Engineering of the yeast *Yarrowia lipolytica* for the production of glycoproteins lacking the outer-chain mannose residues of N-glycans. *Appl. Environ. Microbiol.* **73**, 4446–4454.
61. Wartmann, T., and Kunze, G. (2000) Genetic transformation and biotechnological application of the yeast *Arxula adeninivorans*. *Appl Microbiol Biotechnol* **54**, 619–624.

62. Marx, H., Mecklenbräuker, A., Gasser, B., Sauer, M., and Mattanovich, D. (2009) Directed gene copy number amplification in *Pichia pastoris* by vector integration into the ribosomal DNA locus. *FEMS Yeast Res.* **9**, 1260–1270.
63. Hamilton, S., and Gerngross, T. (2007) Glycosylation engineering in yeast: the advent of fully humanized yeast. *Curr. Opin. Biotechnol.* **18**, 387–392.
64. Cox, H., Mead, D., Sudbery, P., Eland, R.M., Mannazzu, I., and Evans, L. (2000) Constitutive expression of recombinant proteins in the methylotrophic yeast *Hansenula polymorpha* using the PMA1 promoter. *Yeast* **16**, 1191–1203.
65. Oh, D.B., Park, J.S., Kim, M.W., *et al.* (2008) Glycoengineering of the methylotrophic yeast *Hansenula polymorpha* for the production of glycoproteins with trimannosyl core N-glycan by blocking core oligosaccharide assembly. *Biotechnol. J.* **3**, 659–668.
66. Hasslacher, M., Schall, M., Hayn, M., *et al.* (1997) High-level intracellular expression of hydroxynitrile lyase from the tropical rubber tree *Hevea brasiliensis* in microbial hosts. *Protein Expr. Purif.* **11**, 61–71.
67. Clare, J,J,, Rayment, F.B., Ballantine, S.P., Sreekrishna, K., and Romanos, M.A. (1991) High-level expression of tetanus toxin fragment C in *Pichia pastoris* strains containing multiple tandem integrations of the gene. *Bio/Technology* **9**, 455–460.
68. Richardson, P.T., Roberts, L.M., Gould, J.H., and Lord, J.M. (1988) The expression of functional ricin B-chain in *Saccharomyces cerevisiae. Biochim. Biophys. Acta* **950**, 385–394.
69. Binder, M., Schanz, M., and Hartig, A. (1991) Vector-mediated overexpression of catalase A in the yeast *Saccharomyces cerevisiae* induces inclusion body formation. *Eur. J. Cell Biol.* **54**, 305–312.
70. Choi, S.Y., Lee, S.Y., and Bock, R.M. (1993) High level expression in *Saccharomyces cerevisiae* of an artificial gene encoding a repeated tripeptide aspartyl-phenylyalanyl-lysine. *J. Biotechnol.* **30**, 211–223.
71. Weik, R., Francky, A., Striedner, G., Raspor, P., Bayer, K., and Mattanovich, D. (1998) Recombinant expression of alliin lyase from garlic (*Allium sativum*) in bacteria and yeasts. *Planta Med.* **64**, 387–388.
72. Idiris, A., Tohda, H., Kumagai, H., and Takegawa, K. (2010) Engineering of protein secretion in yeast: strategies and impact on protein production. *Appl. Microbiol. Biotechnol.* **86**, 403–417.
73. Barr, K.A., Hopkins, S.A., and Sreekrishna, K. (1992) Protocol for efficient secretion of HSA developed from *Pichia pastoris. Pharm. Eng.* **12**, 48–51.
74. Kauffman, K.J., Pridgen, E.M., Doyle, F.J. 3rd, Dhurjati, P.S., and Robinson, A.S. (2002) Decreased protein expression and intermittent recoveries in BiP levels result from cellular stress during heterologous protein expression in *Saccharomyces cerevisiae. Biotechnol. Prog.* **18**, 942–950.
75. Hohenblum, H., Borth, N., and Mattanovich, D. (2003) Assessing viability and cell-associated product of recombinant protein producing *Pichia pastoris* with flow cytometry. *J. Biotechnol.* **102**, 281–290.
76. Hohenblum, H., Gasser, B., Maurer, M., Borth, N., and Mattanovich, D. (2004) Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant *Pichia pastoris. Biotechnol Bioeng.* **85**, 367–375.
77. Kohno, K. (2010) Stress-sensing mechanisms in the unfolded protein response: similarities and differences between yeast and mammals. *J. Biochem.* **147**, 27–33.
78. Mori, K. (2009) Signalling pathways in the unfolded protein response: development from yeast to mammals. *J. Biochem.* **146**, 743–750.
79. Gasser, B., Saloheimo, M., Rinas, U., *et al.* (2008) Protein folding and conformational stress in microbial cells producing recombinant proteins: a host comparative overview. *Microb. Cell. Fact.* **7**, 11.
80. Valkonen, M., Penttila, M., and Saloheimo, M. (2003) Effects of inactivation and constitutive expression of the unfolded- protein response pathway on protein production in the yeast *Saccharomyces cerevisiae. Appl. Environ. Microbiol.* **69**, 2065–2072.
81. Gasser, B., Maurer, M., Gach, J., Kunert, R., and Mattanovich, D. (2006) Engineering of *Pichia pastoris* for improved production of antibody fragments. *Biotechnol. Bioeng.* **94**, 353–361.
82. Pakula, T.M., Laxell, M., Huuskonen, A., Uusitalo, J., Saloheimo, M., and Penttila, M. (2003) The effects of drugs inhibiting protein secretion in the filamentous fungus *Trichoderma reesei.* Evidence for down-regulation of genes that encode secreted proteins in the stressed cells. *J. Biol. Chem.* **278**, 45011–45020.
83. Swaim, C., Anton, B., Sharma, S., Taron, C., and Benner, J. (2008) Physical and computational analysis of the yeast *Kluyveromyces lactis* secreted proteome. *Proteomics* **8**, 2714–2723.
84. Vai, M., Brambilla, L., Orlandi, I., Rota, N., Ranzi, B.M., Alberghina, L., and Porro, D.

(2000) Improved secretion of native human insulin-like growth factor 1 from *gas1* mutant *Saccharomyces cerevisiae* cells. *Appl. Environ. Microbiol.* **66,** 5477–5479.

85. Marx, H., Sauer, M., Resina, D., *et al.* (2006) Cloning, disruption and protein secretory phenotype of the *GAS1* homologue of *Pichia pastoris. FEMS Microbiol. Lett.* **264**, 40–47.
86. Passolunghi, S., Riboldi, L., Dato, L., Porro, D., and Branduardi, P. (2010) Cloning of the *Zygosaccharomyces bailii GAS1* homologue and effect of cell wall engineering on protein secretory phenotype. *Microb. Cell. Fact.* **9**, 7.
87. Resina, D., Maurer, M., Cos, O., *et al.* (2009) Engineering of bottlenecks in *Rhizopus oryzae* lipase production in *Pichia pastoris* using the nitrogen source-regulated FLD1 promoter. *N. Biotechnol.* **25**, 396–403.
88. Gemmill, T.R., and Trimble, R.B. (1999) Overview of N- and O-linked oligosaccharide structures found in various yeast species. *Biochim. Biophys. Acta.* **1426**, 227–237.
89. Gellissen, G. (2000) Heterologous protein production in methylotrophic yeasts. *Appl. Microbiol. Biotechnol.* **54**, 741–750.
90. Nakayama, K., Nagasu, T., Shimma, Y., Kuromitsu, J., and Jigami, Y. (1992) *OCH1* encodes a novel membrane bound mannosyltransferase: outer chain elongation of asparagine-linked oligosaccharides. *EMBO J.* **11**, 2511–2519.
91. Zhou, J., Zhang, H., Liu, X., Wang, P.G., and Qi, Q. (2007) Influence of N-glycosylation on *Saccharomyces cerevisiae* morphology: a golgi glycosylation mutant shows cell division defects. *Curr. Microbiol.* **55**, 198–204.
92. Jacobs, P.P., Geysens, S., Vervecken, W., Contreras, R., and Callewaert, N. (2009) Engineering complex-type N-glycosylation in *Pichia pastoris* using GlycoSwitch technology. *Nat. Protoc.* **4**, 58–70.
93. Wildt, S., and Gerngross, T.U. (2005) The humanization of N-glycosylation pathways in yeast. *Nat. Rev. Microbiol.* **3**, 119–128.
94. Chiba, Y., and Akeboshi, H. (2009) Glycan engineering and production of 'humanized' glycoprotein in yeast cells. *Biol. Pharm. Bull.* **32**, 786–795.
95. De Pourcq, K., De Schutter, K., and Callewaert, N. (2010) Engineering of glycosylation in yeast and other fungi: current state and perspectives. *Appl. Microbiol. Biotechnol.* **87**, 1617–1631.
96. Potgieter, T.I., Cukan, M., Drummond, J.E., *et al.* (2009) Production of monoclonal antibodies by glycoengineered *Pichia pastoris. J. Biotechnol.* **139**, 318–325.
97. Goto, M. (2007) Protein O-glycosylation in fungi: diverse structures and multiple functions. *Biosci. Biotechnol. Biochem.* **71**, 1415–1427.
98. Chigira, Y., Oka, T., Okajima, T., and Jigami, Y. (2008) Engineering of a mammalian O-glycosylation pathway in the yeast *Saccharomyces cerevisiae*: production of O-fucosylated epidermal growth factor domains. *Glycobiology* **18**, 303–314.
99. Sumi, A., Okuyama, K., Kobayashi, K., Ohtani, W., Ohmura, T., and Yokoyama, K. (1999) Purification of recombinant human serum albumin. Efficient purification using STREAMLINE. *Bioseparation* **8**, 195–200.
100. Westerhoff, H., and Palsson, B. (2004) The evolution of molecular biology into systems biology. *Nat. Biotechnol.* **22**, 1249–1252.
101. Olivares-Hernandez, R., Usaite, R., and Nielsen, J. (2010) Integrative analysis using proteome and transcriptome data from yeast to unravel regulatory patterns at post-transcriptional level. *Biotechnol. Bioeng.* doi:10.1002/bit.22868.
102. Zhang, J., Vemuri, G., and Nielsen, J. (2010) Systems biology of energy homeostasis in yeast. *Curr. Opin. Microbiol.* **13**, 382–388.
103. Han, M., Lee, J., Lee, S., and Yoo, J. (2008) Proteome-level responses of *Escherichia coli* to long-chain fatty acids and use of fatty acid inducible promoter in protein production. *J. Biomed. Biotechnol.* **2008**, 735101.
104. Fürch, T., Wittmann, C., Wang, W., Franco-Lara, E., Jahn, D., and Deckwer, W. (2007) Effect of different carbon sources on central metabolic fluxes and the recombinant production of a hydrolase from *Thermobifida fusca* in *Bacillus megaterium. J. Biotechnol.* **132**, 385–394.
105. Wang, Y., Xue, W., Sims, A.H., *et al.* (2008) Isolation of four pepsin-like protease genes from *Aspergillus niger* and analysis of the effect of disruptions on heterologous laccase expression. *Fungal. Genet. Biol.* **45**, 17–27.
106. Kimura, S., Maruyama, J., Takeuchi, M., and Kitamoto, K. (2008) Monitoring global gene expression of proteases and improvement of human lysozyme production in the nptB gene disruptant of Aspergillus oryzae. *Biosci Biotechnol Biochem* **72**, 499–505.
107. Wong, D., Wong, K., Nissom, P., Heng, C., and Yap, M. (2006) Targeting early apoptotic genes in batch and fed-batch CHO cell cultures. *Biotechnol. Bioeng.* **95**, 350–361.
108. Pizarro, F., Vargas, F., and Agosin, E. (2007) A systems biology perspective of wine fermentations. *Yeast* **24**, 977–991.

109. Takors, R., Bathe, B., Rieping, M., Hans, S., Kelle, R., and Huthmacher, K. (2007) Systems biology for industrial strains and fermentation processes--example: amino acids. *J. Biotechnol.* **129**, 181–190.
110. Mukhopadhyay, A., Redding, A., Rutherford, B., and Keasling, J. (2008) Importance of systems biology in engineering microbes for biofuel production. *Curr. Opin. Biotechnol.* **19**, 228–234.
111. Graf, A., Dragosits, M., Gasser, B., and Mattanovich, D. (2009) Yeast systems biotechnology for the production of heterologous proteins. *FEMS Yeast Res.* **9**, 335–348.
112. Sauer, M., Branduardi, P., Gasser, B., Valli, M., Maurer, M., Porro, D., and Mattanovich, D. (2004) Differential gene expression in recombinant *Pichia pastoris* analysed by heterologous DNA microarray hybridisation. *Microb. Cell. Fact.* **3**, 17.
113. Gasser, B., Maurer, M., Rautio, J., *et al.* (2007) Monitoring of transcriptional regulation in *Pichia pastoris* under protein production conditions. *BMC Genomics* **8**, 179.
114. Dragosits, M., Stadlmann, J., Albiol, J., *et al.* (2009) The effect of temperature on the proteome of recombinant *Pichia pastoris. J. Proteome Res.* **8**, 1380–1392.
115. Dragosits, M., Stadlmann, J., Graf, A., *et al.* (2010) The response to unfolded protein is involved in osmotolerance of *Pichia pastoris. BMC Genomics* **11**, 207.
116. Gonzalez, R., Andrews, B., Molitor, J., and Asenjo, J. (2003) Metabolic analysis of the synthesis of high levels of intracellular human SOD in *Saccharomyces cerevisiae* rhSOD 2060 411 SGA122. *Biotechnol. Bioeng.* **82**, 152–169.
117. Heyland, J., Fu, J., Blank, L.M., and Schmid, A. (2010) Quantitative physiology of *Pichia pastoris* during glucose-limited high-cell density fed-batch cultivation for recombinant protein production. *Biotechnol. Bioeng.* **107**, 357–368.
118. Gasser, B., Sauer, M., Maurer, M., Stadlmayr, G. and Mattanovich, D. (2007) Transcriptomics-based identification of novel factors enhancing heterologous protein secretion in yeasts. *Appl. Environ. Microbiol.* **73**: 6499–6507.
119. Dragosits, M., Frascotti, G., Bernard-Granger, L., *et al* (2010) Influence of growth temperature on the production of antibody Fab fragments in different microbes: a host comparative analysis, *Biotechnology Progress*, accepted for publication.
120. Mattanovich, D., Gasser, B., Hohenblum, H., and Sauer, M. (2004) Stress in recombinant protein producing yeasts. *J. Biotechnol.* **113**, 121–135.
121. van Zutphen, T., Baerends, R.J., Susanna, K.A., de Jong, A., Kuipers, O.P., Veenhuis, M., and van der Klei, I.J. (2010) Adaptation of *Hansenula polymorpha* to methanol: a transcriptome analysis. *BMC Genomics* **11**, 1.
122. Solà, A., Maaheimo, H., Ylönen, K., Ferrer, P. and Szyperski, T. (2004) Amino acid biosynthesis and metabolic flux profiling of *Pichia pastoris. Eur. J. Biochem.* **271**, 2462–2470.
123. Solà, A., Jouhten, P., Maaheimo, H., Sánchez-Ferrando, F., Szyperski, T. and Ferrer, P. (2007) Metabolic flux profiling of *Pichia pastoris* grown on glycerol/methanol mixtures in chemostat cultures at low and high dilution rates. *Microbiology* **153**, 281–290.
124. Rautio, J.J., Kataja, K., Satokari, R., Penttila, M., Soderlund, H., and Saloheimo, M. (2006) Rapid and multiplexed transcript analysis of microbial cultures using capillary electophoresis-detectable oligonucleotide probe pools. *J. Microbiol. Methods.* **65**, 404–416.
125. Weik, R., Striedner, G., Francky, A., Raspor, P., Bayer, K., and Mattanovich, D. (1999) Induction of oxidofermentative ethanol formation in recombinant cells of *Saccharomyces cerevisiae* yeasts. *Food Technol. Biotechnol.* **37**, 191–194.
126. Hong, F., Meinander, N.Q. and Jonsson, L.J. (2002) Fermentation strategies for improved heterologous expression of laccase in *Pichia pastoris. Biotechnol. Bioeng.* **79**, 438–449.
127. Zhang, W., Smith, L.A., Plantz, B.A., Schlegel, V.L., and Meagher, M.M. (2002) Design of methanol feed control in *Pichia pastoris* fermentations based upon a growth model. *Biotechnol Prog.* **18**, 1392–1399.
128. Jahic, M., Veide, A., Charoenrat,T., Teeri, T., and Enfors, S.O. (2006) Process technology for production and recovery of heterologous proteins with *Pichia pastoris. Biotechnol. Prog.* **22**, 1465–1473.
129. Kim, S.J., Lee, J.A., Kim, Y.H., and Song, B.K. (2009) Optimization of the functional expression of *Coprinus cinereus* peroxidase in *Pichia pastoris* by varying the host and promoter. *J. Microbiol. Biotechnol.* **19**, 966–971.
130. Pla, I.A., Damasceno, L.M., Vannelli, T., Ritter, G., Batt, C.A., and Shuler, M.L. (2006) Evaluation of Mut+ and MutS *Pichia pastoris* phenotypes for high level extracellular scFv expression under feedback control of the methanol concentration. *Biotechnol. Prog.* **22**, 881–888.
131. d'Anjou, M.C., and Daugulis, A.J. (2001) A rational approach to improving productivity in recombinant *Pichia pastoris* fermentation. *Biotechnol. Bioeng.* **72**, 1–11.

132. Jungo, C., Marison, I., and von Stockar, U. (2007) Mixed feeds of glycerol and methanol can improve the performance of *Pichia pastoris* cultures: A quantitative study based on concentration gradients in transient continuous cultures. *J. Biotechnol.* **128**, 824–837.
133. Kobayashi, K., Kuwae, S., Ohya, T., Ohda, T., Ohyama, M., and Tomomitsu, K. (2000) High level secretion of recombinant human serum albumin by fed-batch fermentation of the methylotrophic yeast, *Pichia pastoris*, based on optimal methanol feeding strategy. *J. Biosci. Bioeng.* **90**, 280–288.
134. Curvers, S., Brixius, P., Klauser, T., Thommes, J., Weuster-Botz, D., Takors, R., and Wandrey, C. (2001) Human chymotrypsinogen B production with *Pichia pastoris* by integrated development of fermentation and downstream processing. Part 1. Fermentation. *Biotechnol Prog* **17**, 495–502.
135. Maurer, M., Kuhleitner, M., Gasser, B., and Mattanovich, D. (2006) Versatile modeling and optimization of fed batch processes for the production of secreted heterologous proteins with *Pichia pastoris. Microb. Cell Fact.* **5**, 37.
136. Potgieter, T.I., Kersey, S.D., Mallem, M.R., Nylen, A.C., and d'Anjou, M. (2010) Antibody expression kinetics in glycoengineered *Pichia pastoris. Biotechnol. Bioeng.* **106**, 918–927.
137. Goodey, A.R. (1993) The production of heterologous plasma proteins. *Trends Biotechnol.* **11**, 430–433.
138. Porro, D., Martegani, E., Ranzi, B.M., and Alberghina, L. (1991) Heterologous gene expression in continuous cultures of budding yeast. *Appl. Microbiol. Biotechnol.* **34**, 632–636.
139. Stephanopoulos, G., Aristodou, A., and Nielsen, J. (1998) *Metabolic engineering.* Academic Press, Inc., San Diego, Calif.

Chapter 18

Yeasts as a Tool for Heterologous Gene Expression

Raja Mokdad-Gargouri, Salma Abdelmoula-Soussi, Nadia Hadiji-Abbès, Ines Yacoubi-Hadj Amor, Istabrak Borchani-Chabchoub, and Ali Gargouri

Abstract

The yeasts *Saccharomyces cerevisiae* and *Pichia pastoris* are attractive hosts for production of human proteins. The main advantages offered by these systems are the well-developed and easily accessible genetic tools, rapid growth, the simple and inexpensive culture media, and many of the cellular and metabolic processes found in higher eukaryotes are conserved in both yeast species. In this chapter, we describe the production of two proteins of therapeutic interest: the human P53 tumor suppressor and the viral HBsAg in *P. pastoris* and *S. cerevisiae* using the strong and inducible promoters AOX1 and Gal10/Cyc1, respectively. Besides the production as a goal of both expressions, we also report on an unexpected result that has occurred in *S. cerevisiae*: The overexpression of human p53 induces yeast cell death with characteristic markers of apoptosis, such as the externalization of phosphatidylserines and DNA strand cleavage.

Key words: Heterologous expression, *Saccharomyces cerevisiae*, *Pichia pastoris*, Tumor suppressor gene P53, HBsAg

1. Introduction

The yeast *Saccharomyces cerevisiae* has been extensively used for the production of foreign proteins (1–3). Several characteristics of this system account for its popularity: safety as an organism, the well-developed and easily accessible genetic tools, rapid growth and simple and inexpensive culturing conditions, easiness of DNA manipulation for in vivo genetic modification, eukaryotic post-translational modification and secretion pathways, etc.

A wide variety of proteins from fungal and mammalian species have been successfully produced in *S. cerevisiae*, even though it is not the most efficient yeast for secretion of a high protein level in the

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_18, © Springer Science+Business Media, LLC 2012

extracellular medium, but hopefully, the secretion system of yeast has been improved, and the secretion titer of some heterologous proteins has been increased from milligrams to grams per liter (4–7).

Later on, several other gene expression systems were developed in other yeasts such as *Pichia pastoris*, *Hansenula polymorpha*, *Klyveromyces lactis*, each one presenting different advantages and drawbacks (8–11). Nevertheless, *Pichia pastoris* offered much more useful properties as a host for foreign gene expression: (1) a strong and methanol-inducible promoter (AOX1) to drive gene expression, (2) growth at high cell density, and (3) the efficient secretion of the recombinant protein into the growth medium. *P. pastoris* secretes low levels of endogenous proteins into the media, and consequently, the heterologous expressed protein would be highly represented in the total secreted proteins, which would facilitate the subsequent purification steps (9–12). Although very similar in several points, *Pichia* or *Saccharomyces* differ greatly in terms of transforming vectors. Until now, only integrating vectors are available in *Pichia* system, while various types of vectors exist in *S. cerevisiae*: integrating, replicating ARS, replicating episomal (based on 2 μ system), centromeric, and YAC (yeast artificial chromosome).

In this paper, we describe the production in *S. cerevisiae* and *P. pastoris* of two therapeutic proteins: the human tumor suppressor P53 and the antigen S of the hepatitis B virus (HBsAg). The tumor suppressor P53 protects human cells against anarchic proliferation by inducing cell cycle arrest and apoptosis in response to genotoxic stress (13, 14). The P53 is inactivated by missense point mutations in about 50% of human cancers and most of these genetic variations are located in the central DNA-binding domain (15). P53 has been expressed in various eukaryotic systems and the resulting recombinant protein was used as antigen to detect auto antibodies in the sera of cancers' patients (16, 17). More interestingly, its successful expression in yeast allowed Flaman et al. to elaborate a useful tool, the FASAY (Functional Analysis of Separated Alleles on Yeast) to assess the functionality of mutant forms of P53, which are isolated from human cancer cells (18). This assay is based on the action of p53 as a transcription factor on a target reporter gene which is placed under the control of the p53 response element. Even though the expression of p53 has been carried out since 1982 (19), we demonstrated that the overexpression of the wild type p53 induced death of *S. cerevisiae* cells on minimal medium (20).

Hepatitis B virus (HBV) is one of the main etiological agents of acute and chronic liver disease. It still constitutes a major public health problem since more than two billion people worldwide have a history with this infection and 350 millions being chronic carriers (21). As there are no effective treatments for hepatitis B infection, preventive vaccines are the only way for its control (22). The surface antigen of hepatitis B virus (HBsAg) represents the basal component of vaccines against hepatitis B virus infection and was successfully expressed in yeast (23–26).

2. Materials

2.1. Strains and Vectors

1. *Saccharomyces cerevisiae* strain W303-1B [*Mat*α, *leu2*-3112 *trp1-1* can1-100 *ura3-1 ade2-1 his3*-11,15].
2. *Pichia pastoris* strains X33 (Mut+, His+) and KM71 (Muts, His+, Arg+) purchased from Invitrogen (San Diego, CA, USA).
3. The YepDP1-8 (provided by D. Pompon) contains the Gal10/Cyc1 promoter, inducible by the galactose, the 2 μ replicative origin, and the URA3 marker.
4. The pPICZαC vector (Invitrogen) contains the AOX1 promoter, inducible by the methanol, and the Zeocin gene as marker.

2.2. Media for S. cerevisiae and P. pastoris Strains

1. Minimal Medium: 0.67% yeast nitrogen base (Difco), 2% glucose. Tryptophan, histidine, leucine, and adenine were added to a concentration of 40, 20, 120, and 40 mg/L, respectively.
2. Rich Medium: 1% Bacto peptone, 1% yeast extract, 2% glucose.
3. BMGH: 100 mM $KHPO_4$ pH 6, 1.34% Yeast nitrogen base, 0.000041% biotin, 0.0041% histidine, 1% glycerol.
4. MDH: 1.34% Yeast nitrogen base, 0.000041% biotin, 2% glucose, 0.0041% histidine.
5. MMH: 1.34% Yeast nitrogen base, 0.000041% biotin, 0.5% methanol, 0.0041% histidine.
6. YPD: 1% Yeast extract (Difco), 2% Bacto peptone (Difco), 2% glucose.
7. YPDZ: YPD + 100 μg/mL Zeocin.

2.3. Buffers

1. TE: 10 mM Tris–HCl, 1 mM EDTA pH 8.
2. LiCl/TE: 100 mM LiCl in TE buffer.
3. Lysis buffer: 1 M sorbitol, 50 mM citric acid, 150 mM K_2HPO_4, 10 mM EDTA, 0.1% β-mercaptoethanol, 0.3 mg/mL zymolyase.
4. Permeabilization solution: 0.1% Triton X-100, 0.1% sodium citrate.
5. SCK buffer: 1 M sorbitol, 50 mM citric acid, 150 mM K_2HPO_4.
6. Buffer A: 1 M sorbitol, 50 mM citric acid, 150 mM K_2HPO_4, 10 mM EDTA, pH 8, 0.1% β mercaptoethanol, 0.3 mg/mL zymolyase pH 7.
7. Buffer B: 150 mM NaCl, 10 mM Tris–HCl pH 7.5, 5 mM EDTA, pH 8, 1% Sarkosyl.

8. Hybridization solution: 50% formamide, 1% Denhardt, 1% SDS, 6× SSC (150 mM NaCl, 15 mM sodium citrate).
9. Breaking buffer for *P. pastoris*: 50 mM Tris–HCl pH 8, 5 mM EDTA pH 8, 150 mM NaCl, 0.1% Nonidet P-40, 1 mM dithiothreitol, 1 mM PMSF, 2 mg/mL leupeptine, 1 mg/mL pepstatin. Store at 4°C.
10. Breaking buffer for *S. cerevisiae*: 50 mM $NaHPO_4$, pH 7.4, 1 mM EDTA pH 8, 5% glycerol.

2.4. SDS-Polyacrylamide Gel Electrophoresis

1. Separating buffer: 3 M Tris–HCl, pH 8.8.
2. Stacking buffer: 1 M Tris–HCl, pH 6.8.
3. Acrylamide/Bis solution (29:1) and *N,N,N,N′*-Tetramethyl-ethylenediamine (TEMED).
4. Ammonium persulfate: prepare 10% solution in water and store at 4°C for 1 week.
5. Running buffer (5×): 25 mM Tris-base, 250 mM glycine, pH 8.3, 0.1% SDS.
6. Prestained molecular weight markers, Broad range (Fermentas).

2.5. Western Blotting

1. Transfer buffer: 39 mM glycine, 48 mM Tris (do not adjust pH), 0.037% SDS, 20% methanol.
2. Hybond P membrane (Amersham-Biosciences) and 3MM chromatography paper.
3. PBS (Phosphate-Buffered Saline): 100 mM NaCl, 80 mM Na_2HPO_4, 20 mM $NaH_2PO_4{\cdot}2\ H_2O$, pH 7.4.
4. PBS-T: PBS supplemented with 0.1% Tween 20.
5. Blocking buffer: 5% (w/v) non fat dry milk in PBS-T.
6. Primary antibodies dilution buffer: PBS-T supplemented with 1% (w/v) fraction V bovine serum albumin (BSA).
7. Secondary antibody: Anti-mouse IgG conjugated to horseradish peroxidase (Bio-Rad).
8. Enhanced chemiluminescence Plus (ECL Plus kit) reagents (Amersham – Biosciences).

3. Methods

3.1. Transformation of S. cerevisiae Strain W303 1-B by Lithium Chloride

1. Preculture the yeast strain in 5 mL of liquid medium YPD and incubate overnight on a rotary shaker at 200 rpm and 30°C. Dilute aliquot of the culture in 10 mL of sterile H_2O and pipette a small amount into the haemocytometer and count. Inoculate 100 mL of YPD with the preculture to give a starting

cell density of 5×10^6 cells/mL. Incubate at 30°C and 200 rpm until the OD 600 nm reaches 4 U/mL. Harvest the cells by centrifugation at $3,000 \times g$ for 5 min.

2. Wash the cells in 25 mL of sterile water and resuspend in 1/200 (v/v) of 0.1 M LiCl, 10 mM Tris–HCl pH 7.5, 1 mM EDTA and incubate for 1 h at 30°C under agitation.
3. Mix in a 1.5-mL sterile tube 100 μL of the cells suspension, 40 μg carrier DNA, 1–5 μg DNA and incubate for 30 min at 30°C under agitation.
4. Add 700 μL of 40% PEG 4000, 100 mM LiCl/TE and incubate for 1 h at 30°C and then 5 min at 42°C (optional).
5. Centrifuge for 2 s and wash twice in TE without mixing.
6. Resuspend in TE, spread on selective MM, and incubate at 30°C for 3 days (see Note 1).

3.2. RNA Extraction and Northern Blot Analysis

Total RNA can be isolated from yeast cells harvested from 100 mL culture of recombinant strains as follows:

1. Break the cells with 1 g alumina in 10 mL TE buffer. After two phenol–chloroform extractions, precipitate the nucleic acids overnight at −20°C with 2 volumes of ethanol.
2. Resuspend the pellet in 10 mL TE buffer and precipitate the RNA by an equal volume of 8 M LiCl at 4°C overnight.
3. Fractionate 20 μg of total RNA by electrophoresis on a 1.5% formaldehyde agarose gel and transfer to a nylon membrane (Amersham-Biosciences). Hybridize the filters to a ^{32}P-labeled gene probe, using a Rediprime kit (Amersham-Biosciences).
4. Hybridize at 42°C overnight followed by two 10 min washes at room temperature with 2× SSC, 0.1% SDS and two 15 min washes at 42°C in 2× SSC, 0.1% SDS. Figure 1 shows a representative example of a Northern blot to visualize the mRNA in a recombinant yeast strain transformed with the HBV S gene (see Note 2).

3.3. Assessment of the Recombinant Protein Expression by Western Blot

1. Grow recombinant yeast strains on MM containing 2% glucose or raffinose until the OD reaches 4 U/mL and then add 2% galactose as an inducer of the Gal10/Cyc1 promoter to drive p53 expression.
2. Harvest an appropriate volume (1 mL) of the culture after 3, 5, 7, 9, and 24 h of methanol induction. Incubate pellets on ice and suspend in 100 μL of lysis buffer.
3. Add acid-washed glass beads (size 0.5 mm) v/v and vortex for eight cycles of 30 s on ice.
4. Separate the protein lysates by 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and transfer onto a polyvinylidene difluoride (PVDF) membrane.

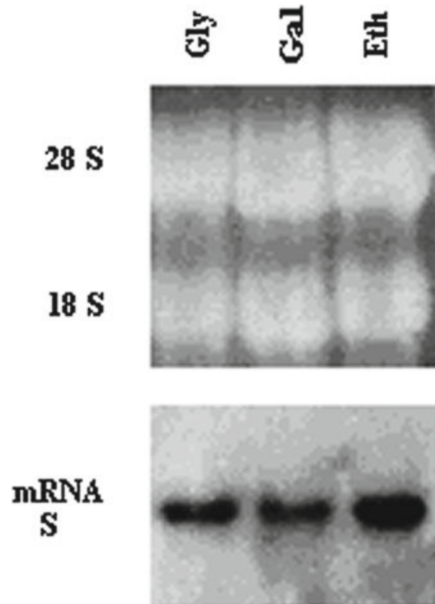


Fig. 1. Northern blot analysis of recombinant yeast strain (W303/HBVS) cultivated on different carbon sources 2% glycerol (Gly), 2% galactose (Gal), and 2% ethanol (Eth). The *upper part* represents the 28S and 18S rRNA on ethidium bromide stained gel, while the lower part shows the autoradiograph of a Northern blot using the HBV S gene as a probe.

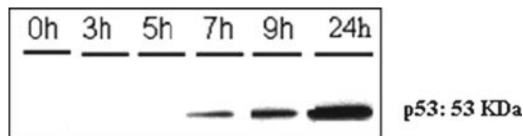


Fig. 2. Western blot analysis of yeast cells expressing human p53. The W303 yeast strain, transformed with wtp53 (W303/wtp53), was grown on MM containing 2% raffinose as a carbon source. Once the OD 600 nm reached 4 U/mL, 2% galactose was added to induce p53 expression driven by the yeast Gal10/Cyc1 inducible promoter. Aliquots were taken at indicated times for protein extraction and Western blot analysis using monoclonal p53-antibody (DO-1 Santa Cruz).

5. Carry out the immunodetection with the monoclonal antibody DO-1 (Santa Cruz Biotechnology) diluted at 1:1,000 (v/v), followed by an anti-mouse immunoglobulin antiserum conjugated to horseradish peroxidase (Bio-Rad).
6. Detect by chemiluminescence using the ECL Plus kit (Amersham-Biosciences). Figure 2 illustrates a representative example of expression of human p53 in *S. cerevisiae* driven by the Gal10/Cyc1 promoter after galactose induction (see Note 3).

3.4. Effect of the Recombinant Protein Expression of Yeast Cell Growth

To study the effect of human p53 expression (wt form and mutant forms) you can carry out a growth kinetic and measure cell viability as follows:

1. Grow yeast precultures to the stationary phase on glucose-containing MM and then dilute to 5×10^6 cells/mL.
2. Harvest cells by centrifugation, wash with distilled sterile water, and use them to inoculate the appropriate cultures.
3. Maintain cultures under continuous agitation (150 rpm) at 30°C and measure the optical density at 600 nm (OD600) every 2 h.
4. For kinetic studies on solid medium, serial dilutions of different recombinant yeast clones can be prepared (10^{-1} to 10^{-6}) from the same precultures, and then 5 μL from each dilution can be spotted onto solid RM and MM containing either glucose or galactose and incubated at 30°C. The number of viable cells can be determined by counting the colony-forming units (CFU). The percentage of recovered viable cells after *p53* gene induction by galactose can be calculated as follows: number of viable cells in the galactose-containing solid MM over the number of viable cells in the glucose-containing solid MM X100.

3.5. Annexin V and Propidium Iodide Staining

Exposed phosphatidylserine and internal propidium iodide (PI) staining can be detected using the Annexin-V-FLUOS Staining Kit (Roche).

1. Yeast cell walls can be digested by zymolyase as described below for the TUNEL assay.
2. Observations are carried out using an LSMS 5 PASCAL confocal microscope (Carl Zeiss).

3.6. Terminal Deoxynucleotidyl Transferase Mediated dUTP Nick End Labeling Assay

DNA strand breaks are detected by terminal deoxynucleotidyl transferase mediated dUTP nick end labeling (TUNEL) with the In Situ Cell Death Detection kit-Fluorescein (Roche).

1. Yeast cells are fixed with 3.7% (v/v) formaldehyde for 30 min at room temperature. Wash the cells three times with phosphate-buffered saline (PBS) and incubate with zymolyase in lysis buffer.
2. Apply 10 μL of the cell suspension to a polylysine-coated slide and allow to dry for 30 min at 37°C. Rinse the slides with PBS, incubate in permeabilization solution for 2 min on ice, and rinse twice with PBS. For a positive control of the TUNEL assay, 30 U of DNAse I can be applied to the same permeabilized cells.
3. Place the microscope slides in a humidified box for 1 h at 37°C and wash twice for 15 min in PBS. Subsequently incubate the

slides with 10 µL of TUNEL reaction mixture for 60 min and finally rinse three times with PBS. Observations can be carried out using an LSMS 5 PASCAL confocal microscope (Carl Zeiss, Germany).

4. Quantify of the apoptotic cells by TUNEL staining by fluorescence-activated cell sorting (FACS) analysis on an EPICS*s* XL flow cytometer (Beckman Coulter). Count at least 15,000 cells per run.

3.7. Gene Expression in the Methylotrophic Yeast Pichia pastoris

1. Culture the host strain (X33 or KM71) in 5 mL of YPD at 30°C overnight.
2. Inoculate 100 mL of YPD with 0.2 mL of preculture and incubate in the same condition until the OD 600 nm reaches 1.3–1.5.
3. Centrifuge 5 min at 3,000 × *g*, 4°C and wash the cells pellet in ice-cold 50 mL of sterile H_2O.
4. Centrifuge 5 min at 1,500 × *g* at 4°C and dissolve the pellet in 50 mL of sterile cold ddH_2O.
5. Centrifuge and dissolve the pellet in 4 mL of ice-cold solution of 1 M sorbitol and then in 0.2 mL of the same solution. The competent cells are kept on ice until use.
6. Mix 80 µL of competent cells with linearized vector DNA (1–2 µg) and incubate on ice for 5 min, and then transfer to an ice-cold 0.2 cm electroporation cuvette. Be sure to tap the cells down to the bottom of the cuvette.
7. Incubate the cuvette with the cells and DNA on ice for 5 min.
8. Pulse the cells according to the manufacturer's instructions for yeast (i.e., *Pichia pastoris* 25 µF, 200 Ω, 1.5 kV).
9. Immediately add 1 mL of ice-cold 1 M sorbitol to the cuvette. Transfer the cuvette contents to a sterile 15-mL tube, and incubate at 30°C without shaking for 1–2 h.
10. Spread 10, 25, 50, 100, and 200 µL on separate labeled YPDS plates containing 100 µg/mL Zeocin and incubate plates at 30°C until colonies form (3–5 days).

3.8. Genomic DNA Extraction

Genomic DNA was extracted from 1 mL of stationary culture of *Pichia* clones.

1. Culture was centrifuged for 5 min at 5,000 × *g*, then washed in SCK buffer and incubated in buffer A for 2 h at 37°C to generate spheroplasts.
2. Centrifuge 15 min at 12,000 × *g* and cell pellet was lysed in buffer B by pipetting.

3. Extract DNA twice by phenol–chloroform (v/v) followed by phenol–chloroform extraction.
4. Genomic DNA was recovered by 2 volumes of ethanol precipitation at −20°C overnight (see Note 4).

3.9. Quantitative PCR for Determining Gene Copy Number

The ICyler (Bio-Rad) was used for real time PCR (qPCR) amplification and detection.

1. Prepare reactions in triplicates, in 25 μL mixture in MicroAmp optical 96-well reaction plates (Bio-Rad). Each reaction contains 5 μL of template DNA (100 ng), 12.5 μL of SYBR Green PCR Master Mix (Bio-Rad), and 12.5 pmol of forward and reverse p53 primers. Measure concentrations of plasmid and genomic DNA by absorbance at 260 and 280 nm with a spectrophotometer. Use nucleic acid with a ratio of absorbance at 260/280 nm ranging between 1.8 and 2.0 indicating minimum protein contamination.
2. Make serial tenfold dilutions of plasmid DNA (pPICZaC/p53) in triplicates to establish the standard curve. Include a negative control in all experimental runs by substituting the template with ddH_2O (see Note 5).
3. Carry out real-time PCR amplification with the following cycling conditions: 95°C for 3 min, followed by 40 cycles of 95°C for 30 s, 55°C for 30 s, and 72°C for 45 s.

3.10. Assessment by Western Blot of Human P53 in Recombinant *P. pastoris* Strains

Recombinant *Pichia* strains harboring one or more copies of the human p53 cDNA in the genome were constructed as described in ref. 27. The transgene was introduced in two genetic backgrounds of *Pichia*, Mut+ and Mut^s (respectively wild type and slow methanol utilizing strains).

1. Cultivate the selected strains on BMGY until OD600 reaches a value of 6 to generate biomass (see Note 5).
2. Harvest cells by centrifugation at 3,000 × *g* for 5 min at room temperature and resuspend in a volume of BMMY until OD600 reaches 1, and then further incubate to induce the expression of the recombinant protein.
3. Add methanol to a final concentration of 0.5% every 24 h to maintain promoter induction.
4. Take 1 mL aliquots periodically at 24, 48, and 72 h; centrifuge at 3,000 × *g* for 5 min. The supernatant containing the extracellular proteins is stored for analysis and the pellet is resuspended in 100 μL of breaking buffer.
5. The cells are lysed by adding a volume of glass beads (0.5 mm diameter, Sigma); vortex the mixture eight to ten times for 30 s each time.

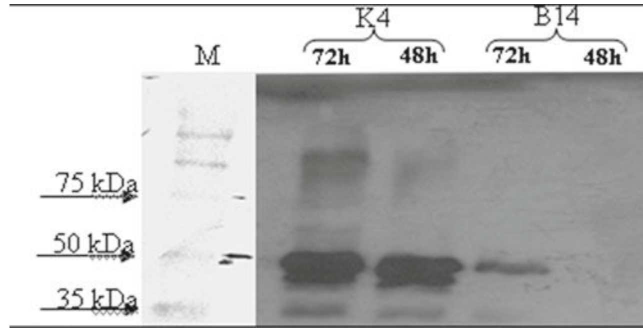


Fig. 3. Comparative expression of human p53 in the two genetic contexts of *P. pastoris* (Mut^s and Mut^+). Total proteins were extracted from recombinant strains K4 (Mut^s) and B14 (Mut^+) after 48 h and 72 h of methanol induction, fractioned on SDS-PAGE 10%, transferred onto a PVDF membrane and analyzed for p53 expression by Western blot with a p53-specific monoclonal antibody.

6. Centrifuge the mixture at 13,000 × *g* for 15 min at 4°C.
7. Remove the supernatant and determine protein concentration using the Bradford assay. We usually obtain 1–5 μg/μL, and load a SDS-PAGE for Western blot using 20 μg of total protein. We have compared the level of P53 in both genetic background of *Pichia* (Muts and Mut+) and concluded that the production of p53 is significantly higher in the Muts than Mut+ context (Fig. 3).

4. Notes

1. It is important to wash the cells pellet twice with TE to get rid of the PEG, which is toxic for yeast cells.
2. RNA could be extracted by other methods, as the method we describe below for the extraction of *Pichia*'s DNA, which is based on the protoplasts formation by zymolyase. Protoplasts are then isolated and lysed in TE and sarkosyl. This step replaces the cell grinding in alumina. Except that, all the following steps are the same.
3. The solution containing the first antibody diluted in PBS-T can be conserved at −20°C and used up to three times.
4. Even if it is not specified, all techniques presented works well with either *Pichia* or *Saccharomyces* cells: Proteins, DNA, and RNA extraction and their analysis by Western, Southern/PCR and Northern/RT-PCR, qPCR analysis of copy number, growth kinetics on solid or liquid media, etc.

5. For *Pichia* strains, it is important to have adequate aeration for growth; in general use 1:5 ratio of media to flask volume.
6. We have found that the serial dilutions of DNA should be left at 4°C overnight, before use in the qPCR assay, to ensure complete dissolution of DNA.

Acknowledgments

The authors would like to thank Lamia Jemal and Mosbah Dardouri for technical assistance. This work was supported by grant of the Tunisian Ministry of Higher Education and Scientific Research.

References

1. Romanos M.A., Scorer C.A., Clare J.J. (1992) Foreign gene expression in yeast. *Yeast*, **8**, 423–488.
2. Buckholz R.G., Gleeson M.A. (1991) Yeast systems for the commercial production of heterologous proteins. *Biotechnology (N Y)*. **9,** 1067–1072.
3. Sudbery P. E. (1996) The expression of recombinant proteins in yeasts. *Cur Op Biotech* 517–524.
4. Robinson A., Hines V., Wittrup K. (1994) Protein disulfide isomerase overexpression increases secretion of foreign proteins in *Saccharomyces cerevisiae*. *Biotechnology (NY)* **12**, 381–384.
5. Wentz A.E., Shusta E.V. (2007) A novel high-throughput screen reveals yeast genes that increase secretion of heterologous proteins. *Appl Environ Microbiol* 7**3**, 1189–1198.
6. Wentz A.E., Shusta E.V. (2008) Enhanced secretion of heterologous proteins from yeast by overexpression of ribosomal subunit RPP0. *Biotechnol Prog* **24**, 748–756.
7. Idiris A., Tohda H., Kumagai H., Takegawa K. (2010) Engineering of protein secretion in yeast: strategies and impact on protein production. *Appl Microbiol Biotechnol* **86**, 403–417.
8. Cereghino J. L., Cregg J.M. (2000) Heterologous protein expression in the methylotrophic yeast Pichia pastoris, *FEMS Microbiol. Rev*. **24**, 45–66.
9. Macauley-Patrick S., Fazenda M. L., McNeil B., Harvey L. M. (2005) Heterologous protein production using the Pichia pastoris expression system. *Yeast* **22**, 249–270.
10. Cregg J.M., Cereghino J. L., Shi J., Higgins D.R. (2000) Recombinant protein expression in *Pichia pastoris*. *Mol. Biotechnol.* **16**, 23–52.
11. Gellissen G., Kunze G., Gaillardin C., Cregg J.M., Berardi E., Veenhuis M., van der Klei I. (2005) New yeast expression platforms based on methylotrophic *Hansenula polymorpha* and *Pichia pastoris* and on dimorphic *Arxula adeninivorans* and *Yarrowia lipolytica*: a comparison. *FEMS Yeast Res* **5**, 1079–1096.
12. Hollenberg C.P., Gellissent G. (2009) Production of recombinant proteins by methylotrophic yeasts *Biotechnology Advances* **27**, 297–306.
13. Levine A.J. (1997) P53 the cellular gatekeeper for the growth and division. *Cell* **88**, 323–331.
14. Lane D.P. (1992) P53, guardian of the genome *Nature* **358**, 15–16.
15. Guimaraes D.P., Hainaut P. (2002) TP53: a key gene in human cancer *Biochimie* **84** 83–93.
16. Hammel P., Boissier B., Chaumette M.T., Piedbois P., Rotman N., Kouyoumdjian J.C., et al. (2000) Detection and monitoring of serum p53 antibodies in patients with colorectal cancer. Gut **40,** 356–361.
17. Kuli A, Sirotkovi -Skerlev M, Jelisavac-Cosi S, Herceg D, Kovac Z, Vrbanec D. (2010) Anti-p53 antibodies in serum: relationship to tumor biology and prognosis of breast cancer patients. Med Oncol. **27**, 887–893.
18. Flaman J. M, Frebourg T., Moreau V., Charbonnier F., Martin C., Chappuis P. et al. (1995) A simple p53 functional assay for screening cell lines, blood, and tumors. *Proc Natl Acad Sci USA* **92**, 3963–3967.
19. Nigro J., Sikorski R., Reed S., Vogelstein B. (1992) Human p53 and CDC2Hs genes combine to inhibit the proliferation of *Saccharomyces cerevisiae*. *Mol Cell Biol* **12**, 1357–1365.

20. Yacoubi-Hadj Amor I., Smaoui K., Chaabene I., Mabrouk I., Djemal L., Elleuch H. et al. (2008). Human p53 induces cell death and downregulates thioredoxin expression in *Saccharomyces cerevisiae*. *FEMS Yeast Res* **8**, 1254–1262.
21. State of the World's Vaccines and Immunization, WHO/UNICEF Bulletin, Geneva, 1996.
22. Rogers S.A., Dienstag J.L., Liang T.J. (1997) Hepatitis B virus-clinical disease, prevention and therapy, in: R.A. Willson (Ed.), Viral Hepatitis, Dekker, New York, 119–146.
23. Hitzeman R.A., Chen C.Y., Hagie F.E., Patzer E.J., Liu C.C., Estell D.A. et al. (1983) Expression of hepatitis B surface antigen gene in yeast. *Nucleic Acids Res.* **11**, 2745–2763.
24. Imamura T., Araki M., Miyanohara A., Nakao J., Yonemura H., Ohtomo N., Matsubara K. (1987) Expression of hepatitis B virus middle and large surface antigen genes in *Saccharomyces cerevisiae*. *J. Virol.* **61**, 3543–3549.
25. Vassileva A., Chugh D.A, Swaminathan S., Khanna N. (2001) Expression of hepatitis B surface antigen in the methylotrophic yeast *Pichia pastoris* using the GAP promoter, *J. Biotechnol.* **88**, 21–35.
26. Hadiji-Abbes N., Borchani-Chabchoub I., Triki H., Ellouz R., Gargouri A., Mokdad-Gargouri R. (2009). Expression of HBsAg and preS2-S protein in different yeast based system: A comparative analysis. *Prot Exp & Purif.* **66**, 131–137.
27. Abdelmoula-Souissi S., Rekik L., Gargouri A., Mokdad-Gargouri R. (2007) High-level expression of human tumour suppressor P53 in the methylotrophic yeast: Pichia pastoris. *Prot. Exp Purif.* **54**, 283–288.

Chapter 19

The Cre/Lox System: A Practical Tool to Efficiently Eliminate Selectable Markers in Fungal Endophytes

Simona Florea, Caroline Machado, Kalina Andreeva, and Christopher L. Schardl

Abstract

Interspecific hybridization is a common evolutionary process for the many epichloid fungi that consequently possess multiple gene copies accumulated from their parental strains. Serial gene manipulations in such strains are impeded by the limited availability of selectable resistance marker genes. Therefore, we developed a method for marker elimination suitable for a range of filamentous fungi that allows the reuse of the same marker for successive manipulations, and can also generate gene knockout mutants free of any foreign genes. For epichloae, the complete elimination of the marker gene from the genome would mitigate public concerns and regulatory hurdles to the use of such fungal strains in field experiments.

Key words: Cre/lox, Transient expression, Transient transfection, Epichloë, Endophyte, Marker excision

1. Introduction

The type I topoisomerase enzyme, Cre recombinase, catalyzes the reciprocal recombination of a pair of directly repeated loxP sites, and has been extensively used to excise the selectable marker flanked by these two sites (1, 2). The excisions are precise, leaving only the small (34 bp) *loxP* sequence in the genome. Numerous studies have reported the use of Cre/lox technology to eliminate the marker genes used for transformation in a number of different organisms (3). We have developed an efficient method to eliminate the marker gene from fungal genomes based on transient expression of Cre recombinase following introduction of the *cre* gene into protoplasts by standard transfection procedures. This method

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_19, © Springer Science+Business Media, LLC 2012

offers benefits over previous methods in that it requires no additional selection or counterselection systems, is rapid, and can be used to generate genetically modified fungal strains that are devoid of foreign genes. Expression of Cre recombinase during transient transfection with a plasmid harboring *cre* under the control of a constitutive fungal promoter is sufficient for marker gene removal with acceptable frequency for PCR screening. The procedure was applied to several species of fungal endophytes. Following transient transfection and recombination, the efficiency of marker excision varied from 0.5 to 2.1%, demonstrating that this is a facile method for eliminating selectable markers from genetically modified fungal strains (see Note 1).

2. Materials

2.1. Equipment and Consumables

1. Centrifuge Allegra25R (Beckman Coulter).
2. TS 5.1-500 swinging-bucket rotor (Beckman Coulter).
3. Biological Safety Cabinet (NuAire).
4. Microscope BH2 (Olympus).
5. Refrigerated orbital shaking incubator M1281-0010 (New Brunswick Scientific).
6. Tissue pulverizer Geno/Grinder2000 (SpexCertiPrep).
7. Borosilicate glass tubes size 16×150 mm (Fisher Scientific).
8. Glass centrifuge tubes 50 mL, Corning, sterile (E&K Scientific).
9. Erlenmeyer flasks 300 mL, Kimax29966, sterile (Kimble).
10. Falcon tubes, No. 2059, sterile (Becton Dickenson).
11. Glass centrifuge tubes, 30 mL, Corex, sterile (Krackeler Scientific).
12. Hemocytometer (Bright-Line improved Neubauer) (American Optical Corporation).
13. Hemocytometer cover glass CG-200 V (Propper).
14. Miracloth, sterile (Calbiochem).
15. Parafilm (UltiDent Scientific).
16. Steriflip filter tube, 50 mL, sterile (Millipore).
17. Oak Ridge tube, 50 mL, sterile (PGC Scientific).

2.2. Transformation Solutions and Media

1. Complete regeneration medium (CRM): 1 g NH_4NO_3, 1 g KH_2PO_4, 1 g NaCl, 0.46 g $MgSO_4 \cdot 7H_2O$, 0.13 g $CaCl_2 \cdot 2H_2O$, 1 g yeast extract, 12 g dehydrated potato dextrose broth, 1 g peptone, 1 g casein hydrolysate (acid), and 304 g sucrose. Mix all ingredients in 600 mL deionized water. After all of the components are dissolved, bring to pH 6.0 with NaOH, and then

add deionized water to bring the volume to 1 L. Add 7 g/L agarose to the solution, and then autoclave.

2. Osmotic solution: 1.2 M $MgSO_4 \cdot 7H_2O$ and 10 mM $NaHPO_4$, pH 6.0. Filter sterilize and store at room temperature (RT).
3. ST solution: 0.6 M sorbitol and 0.1 M Tris–HCl, pH 7.4. Autoclave and store at RT.
4. STC solution: 1 M sorbitol, 50 mM Tris–HCl, pH 7.4, and 50 mM $CaCl_2$. Autoclave and store at RT. Cool on ice before use.
5. Polyethylene glycol (PEG) solution: 60% w/v PEG 3350. Autoclave and store at RT.
6. 3× PEG amendments solution: 1.8 M KCl, 150 mM $CaCl_2$, and 150 mM Tris–HCl, pH 7.4. Autoclave and store at RT.
7. PEG mix: On the day of transfection, mix two parts of PEG solution with one part of 3× PEG amendments solution.
8. Potato dextrose agar (PDA): Prepared as indicated by the supplier.
9. Potato dextrose broth (PDB): Prepared as indicated by the supplier.
10. Cell wall digestion mixture: 5 mg/mL lysing enzymes, 5 mg/mL driselase, and 3 mg/mL bovine serum albumin (BSA) in osmotic solution (10 mL).

3. Methods

The *Neotyphodium coenophialum* transformed strain e7133 (4), harboring a *hygromycin B phosphotransferase* (antibiotic resistance) marker gene flanked by loxP sites (loxP::*hph*::loxP), was subjected to transient transfection with a plasmid (pKAES175) carrying the *cre* gene under the control of a constitutive fungal promoter (P_{tubB}-*cre*) (see Note 1). A simple PCR screen of the DNA extracted from a half colony of each transiently transfected isolate allowed identification of those that underwent marker elimination. For an easy screen, the PCR primers were designed to give different fragment sizes when the flanked marker gene was present or absent (eliminated).

The spores of *Neotyphodium* species used in this study are uninucleate (5), whereas their protoplasts might be multinucleate. Therefore, purification of the nuclear genotypes through single-spore isolation is required for the identification of pure isolates (e7202 and e7204) that had lost the marker gene due to Cre activity. As a method of confirmation that the *hph* marker had been lost, PDA with hygromycin B (Calbiochem) at a concentration previously

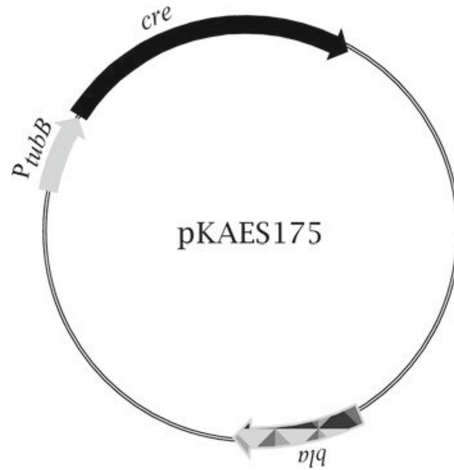


Fig. 1. Plasmid used for transient transfection resulting in marker excision. The plasmid carries a *cre* gene under the control of *PtubB*, the promoter of a constitutively expressed β-tubulin gene of the fungus, *Epichloë typhina*. The *bla* (β-lactamase) gene confers ampicillin resistance to transformed *Escherichia coli*. The plasmid lacks any selectable marker for fungi.

determined to stop growth of the wild-type fungus was inoculated with the fungal isolates identified by PCR as having the marker gene eliminated.

3.1. Plasmid DNA Preparation

1. Culture *Escherichia coli* with pKAES175, (Fig. 1; or any plasmid containing a *cre* gene modified for expression in the fungus).
2. Extract plasmid DNA, using Qiaprep Spin Miniprep Kit (Qiagen) or other bacterial plasmid DNA extraction method. If necessary, concentrate by a standard protocol (e.g., NaOAc/ethanol precipitation) to at least 0.5 μg/μL in H_2O.

3.2. Protoplast Isolation

1. Grow fungal mycelium in PDB (in a 300-mL Erlenmeyer flask) in an orbital shaking incubator (200 rpm) at 20°C for 6–10 days.
2. Prepare the cell wall digestion mixture (Subheading 2.2, step 10) on the day of transformation. Allow the enzyme mixture to dissolve in the osmotic solution, and then filter with Steriflip filter tube.
3. For protoplast isolation (see Notes 2–4), harvest mycelium by centrifuging the culture (from step 1) in 50-mL Corning centrifuge tubes at 4°C for 20 min at 3,400 × *g*.
4. After centrifugation, discard the supernatant and pour the filtered cell wall digestion mixture (from step 2) over the mycelium. Close the tube and shake it to make sure that the mycelium is evenly dispersed in the digestion mixture.

5. Place the tube in horizontal position on a shaker and incubate at 30°C while shaking for 2–3 h (depending on the fungal strain).
6. Separate the protoplasts from undigested fungal mass by filtering through miracloth in a sterile (autoclaved) 30 mL glass centrifuge tube (Corex).
7. Overlay the protoplast suspension (from step 6) with 10 mL of ST separation buffer added drop by drop down the side of the tube.
8. Centrifuge at 3,400 × *g* for 20 min at 4°C in a swinging bucket rotor.
9. Remove the protoplast layer (white cloud) formed at the interface and transfer it to a sterile Oak Ridge tube containing 5 mL of cold STC solution. From this point until transient transfection, keep protoplast cool (on ice and in refrigerated centrifuge).
10. Centrifuge the protoplast-STC suspension at 3,400 × *g* for 10 min at 4°C. Carefully remove the supernatant with a wide-bore pipette, and discard the supernatant.
11. Resuspend the protoplast pellet in 5 mL STC solution. Do not vortex because the protoplasts are delicate; slowly suspend the pellet using a sterile 1-mL pipette.
12. Dilute 50 μL of protoplast suspension (from step 11) into 5 mL of STC and then add 10 μL of the diluted solution at the edges of the coverslip on each side of a hemocytometer (the suspension is drawn onto the grid of the hemocytometer by capillary action). Count the spherically shaped protoplasts on each grid and average the numbers, and then calculate protoplast yield according to the formula for the hemocytometer. For example, in a Neubauer hemocytometer, a complete grid (1 × 1 mm, 0.1-mm depth) contains 100-nL volume. If c protoplasts are counted over an entire grid, then the undiluted protoplast suspension contains $10^6 c$/mL ($c \times 10^4 \times 10^2$ dilution factor/mL).
13. Repeat step 10.
14. Resuspend the pellet in a volume of STC solution (based on the calculation from step 11) to obtain circa 5×10^6 protoplasts in 100 μL STC.
15. Keep the tube of protoplasts on ice until used for transformation/transient transfection.

3.3. Protoplast Transient Transfection

1. In a sterile borosilicate glass tube, add 100 μL protoplast suspension (containing circa 5×10^6 protoplasts), mix slowly with 25 μL PEG mix, and then add 5–7 μg of the plasmid DNA.
2. Incubate the protoplast–PEG–DNA mixture on ice for 30 min.

3. After 30 min, add 1 mL of PEG mix to the tube and mix gently by pipetting. Do NOT vortex (see Note 5), and then incubate for 20 min at RT.
4. Prepare 15–20 Petri plates (Fisher Scientific) by adding 20 mL of CRM medium (without antibiotic) per plate, and let the medium solidify.
5. Prepare the same number (15–20) of Falcon tubes containing 7 mL CRM. Close the tubes with the cap and keep them at 45–50°C to prevent the medium from solidifying.
6. Add 1–2 μL of the PEG mix containing 900–4,000 protoplasts (from step 3) to each Falcon tube, mix gently by inverting several times, and then quickly pour over the plates containing CRM. Spread it over the plate evenly.
7. Allow the medium to solidify completely, then wrap the plates with parafilm (UltiDent Scientific), and incubate them upside down at 22°C.

3.4. Selection of Transiently Transfected Protoplasts (Fig. 2)

1. Allow the protoplasts to recover until colonies are barely visible by eye (incubation time depending on the fungal strains used).
2. Randomly pick as many colonies as possible (e.g., 500 colonies) and transfer them to PDA plates without antibiotic.
3. Allow the colonies to grow to circa 0.5 cm in diameter.
4. Extract DNA from half of each colony using DNeasy 96 Plant Kit (Qiagen) or a DNA extraction technique of choice.
5. Screen the DNA by PCR for marker gene excision.
6. Conduct two to three sequential single-spore isolations from the colonies indicated by PCR as having marker gene eliminated (this procedure is required for purification of the nuclear genotypes).
7. Screen the single-spore isolates by PCR to identify the purified nuclear genotypes that had undergone marker elimination (Fig. 3).

3.5. Sensitivity Test

The isolates identified as having the marker eliminated from fungal genome are grown on medium containing or lacking antibiotic. The growth of the isolates with marker eliminated is arrested on medium containing the selective concentration of antibiotic (i.e., hygromycin B), although they grow normally on medium where the antibiotic has been omitted (Fig. 4).

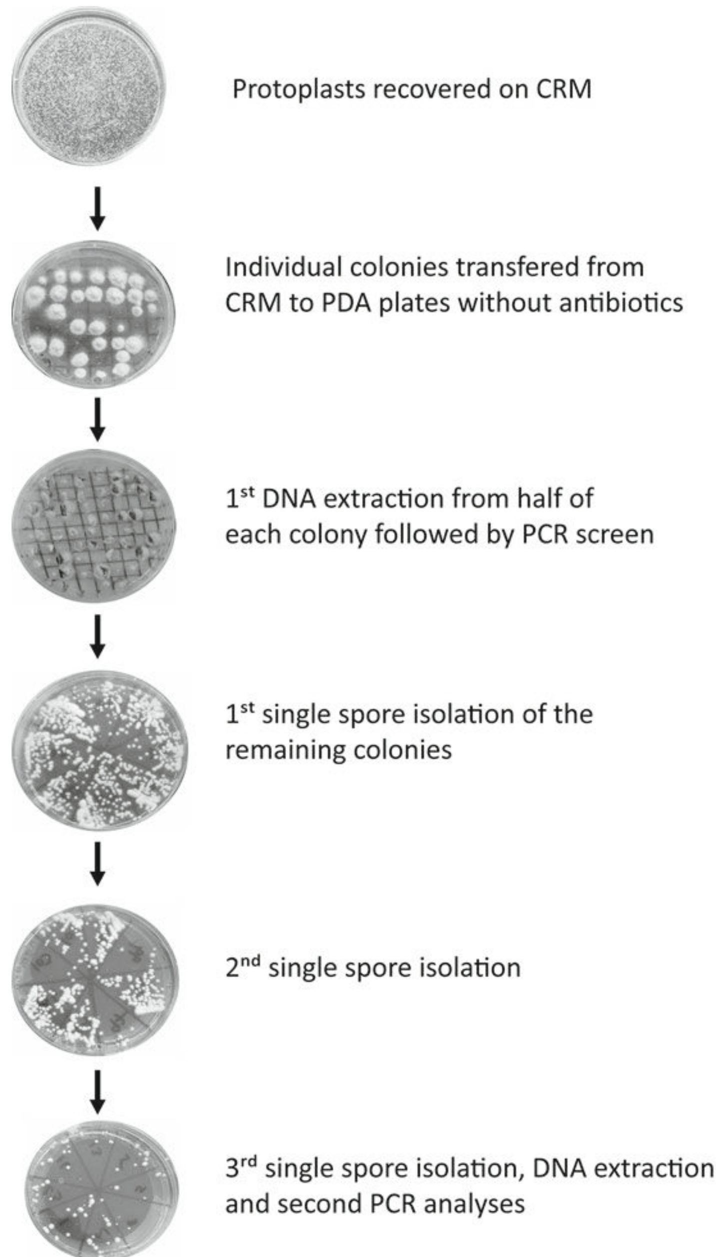


Fig. 2. Steps in identification of fungal isolates having undergone marker excision after transient transfection with plasmid pKAES175 (see Fig. 1).

4. Notes

1. Procedures outlined here are specific for *Epichloë* and *Neotyphodium* species. For other fungi, various parameters differ, such as growth conditions, growth medium, incubation time, osmotic solution, complete regeneration medium, and

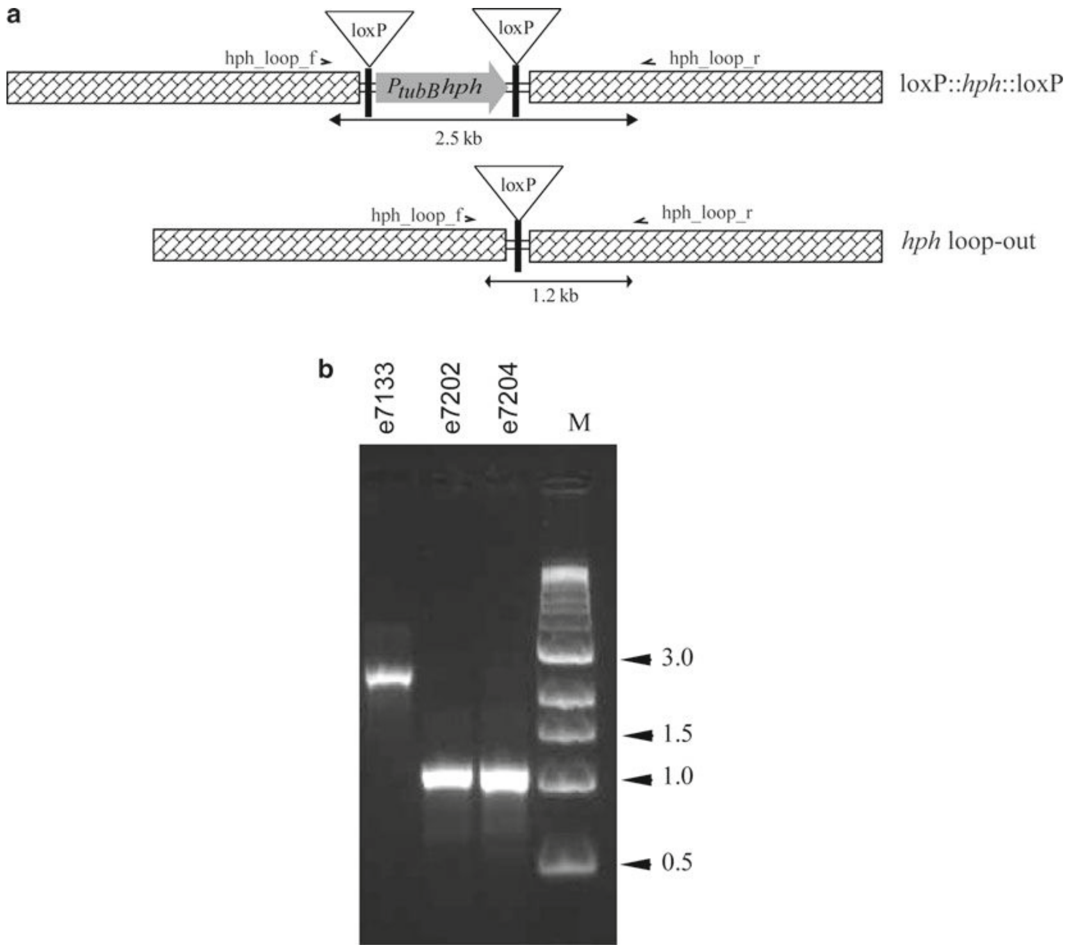


Fig. 3. (**a**) Maps of the targeting vector and locus after marker exchange mutagenesis with loxP::*hph*::loxP, and the locus after excision of *hph* by the activity of Cre recombinase (*hph* loop-out). PCR with the hph_loop_f and hph_loop_r primers flanking the loxP::*hph*::loxP insertion site amplifies differently sized fragments when the marker gene is present or absent (eliminated). (**b**) PCR screen for excision of *hph* after transient transfection of *Neotyphodium coenophialum* e7133, a transformant with the loxP::*hph*::loxP marker. Lanes contain size markers indicated in kb (lane M), PCR products with DNA templates from e7133, and hygromycin B-sensitive isolates e7202 and e7204 derived from transient transfection with pKAES175. The Cre-mediated excision of *hph* is indicated by the 1.2-kb PCR product.

components of cell wall digestion mixture. To apply this approach to another fungus, consult published methods for protoplast isolation and transformation or transfection of that species.

2. Cover the mouths of borosilicate glass tubes with aluminum foil, wrap the glass and plastic tubes, flasks and miracloth in aluminum foil, and sterilize them by autoclaving in a dry cycle.
3. The protoplast isolation and transformation procedures should be performed in a biological safety cabinet.
4. Keep the protoplasts on ice at all times.
5. Do not vortex the mixture containing the protoplasts at any time.

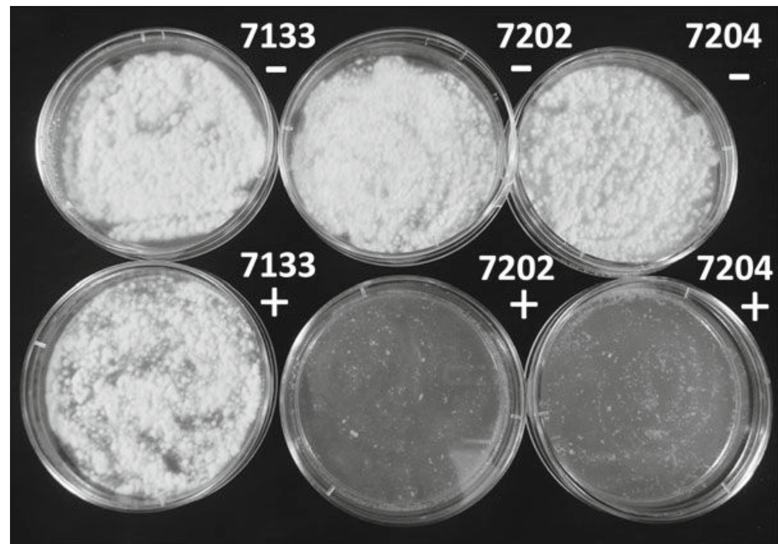


Fig. 4. Response to hygromycin B of isolates derived from *N. coenophialum* e7133 (loxP::*hph*::loxP) after transient transfection with pKAES175. Isolate e7133 grows on PDA containing (+) or lacking (–) hygromycin B. The e7202 and e7204 isolates fail to grow on hygromycin B-containing medium.

Acknowledgments

We thank Walter Hollin for technical support. Funding was provided by the US Department of Agriculture grants 2005-35319-16141, 2007-10021743, and 2008-35318-04549.

References

1. Sauer, B. (1987) Functional expression of the Cre-lox site-specific recombination system in the yeast *Saccharomyces cerevisiae*. *Molec. Cell. Biol.* **7**, 2087–2096.
2. Sternberg, N. and Hamilton, D. (1981) Bacteriophage P1 site-specific recombination. *J. Molec. Biol.* **150**, 467–468.
3. Schwenk, F., Sauer, B., Kukoc, N., Hoess, R., Muller, W., Kocks, C., Kuhn, R. and Rajewsky, K. (1997) Generation of Cre recombinase-specific monoclonal antibodies, able to characterize the pattern of Cre expression in cre-transgenic mouse strains. *J. Immunol. Methods* **207**, 203–212.
4. Florea, S., Andreeva, K., Machado, C., Mirabito, P. M. and Schardl, C. L. (2009) Elimination of marker genes from transformed filamentous fungi by unselected transient transfection with a Cre-expressing plasmid. *Fungal Genet. Biol.* **46**, 721–730.
5. Kuldau, G. A., Tsai, H. F. and Schardl, C. L. (1999) Genome sizes of *Epichloë* species and anamorphic hybrids. *Mycologia* **91**, 776–782.

Chapter 20

Aptamer-Regulated Expression of Essential Genes in Yeast

Beatrix Suess, Karl-Dieter Entian, Peter Kötter, and Julia E. Weigand

Abstract

Conditional gene expression systems are important tools for the functional analysis of essential genes. Tetracycline (tc)-binding aptamers can be exploited as artificial riboswitches for the efficient control of gene expression by inserting them into the 5′ untranslated region of an mRNA. The ligand-bound form of those mRNAs inhibits gene expression by interfering with translation initiation. In contrast to previous tc-dependent regulatory systems, where tc inhibits or activates transcription upon binding to the repressor protein TetR, the tc-binding aptamer system inhibits translation of the respective mRNA. We describe here a simple and powerful PCR-based strategy which allows easy tagging of any target gene in yeast using a tc aptamer-containing insertion cassette. The expression window can be adjusted with different promoters and protein synthesis is rapidly switched off.

Key words: Aptamer, Conditional gene expression, Essential genes, Yeast, Riboswitch

1. Introduction

Conditional gene expression systems are important tools to analyze the function of essential genes. For this purpose, inducible promoters have been employed using either endogenous or heterologous sequences. In yeast, promoters responding to galactose or methionine have been used, which necessitate a change in growth conditions, leading to side effects potentially obscuring functional analysis (1, 2). A widely used engineered expression system relies on the bacterial tetracycline (tc) repressor TetR (3–5). Here, gene expression is dependent on the, for yeast harmless, antibiotic tc. Alternative to the use of inducible promoters, proteins can be temperature sensitized either by mutation or fusion to a heat-inducible

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_20, © Springer Science+Business Media, LLC 2012

degron cassette (6, 7). A temperature increase, however, also provides severe second side effects due to heat-shock reactions. All of these systems need a certain strain background which can hinder their customized application.

Riboswitches are RNA control elements, which regulate gene expression in response to small molecule concentrations (8). They sense their effector by direct RNA–ligand interaction, rendering auxiliary protein factors unnecessary. This direct response makes riboswitches ideal models for the engineering of synthetic RNA-based gene regulation systems. In contrast to natural riboswitches, which sense cellular metabolites, aptamers responding to nonmetabolizable ligands have been employed to build up engineered riboswitches in different organisms acting on various levels of gene expression (9, 10).

We successfully applied a tc-binding aptamer for the control of translation initiation and pre-mRNA splicing in yeast (11–13). Translational control is achieved by insertion of one to three aptamer copies into the 5′ untranslated region (UTR) of a yeast gene (Fig. 1). In the absence of tc, downstream genes are expressed, but upon tc binding the aptamer–ligand complex is able to interfere either with cap binding or scanning of the 5′ UTR. Using tc-binding aptamer regulation, factors up to 40-fold have been achieved. In addition, gene repression is tight enough to allow the conditional expression of essential genes (14).

In this chapter, we describe an easy method for the conditional expression of yeast genes. Genes are tagged by chromosomal integration of a universal insertion cassette. Repression of gene expression can be followed either by Western blot analysis or, for essential genes, by serial dilution growth assays. Due to the use of direct RNA–small molecule interactions, the method is independent on the genetic background of the yeast strain.

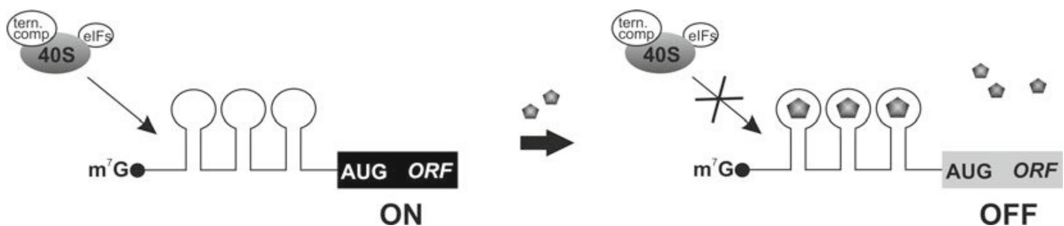


Fig. 1. Regulation of translation initiation by the tc aptamer. Three copies of the tc aptamer are inserted into the 5′ UTR of an mRNA. *Left*: Downstream genes are expressed only in the absence of the ligand. *Right*: Upon tc binding, the aptamer–ligand complex interferes with translation initiation; the expression of downstream genes is repressed.

2. Materials

2.1. Expression Vectors Containing the Universal Insertion Cassettes

The vectors pADH1-tc3-3xHA and pTDH3-tc3-3xHA contain a universal insertion cassette for the genomic tagging of yeast genes (Fig. 2). Three copies of the tc aptamer are followed by three repeats of the HA tag (see Note 1), both under the control of constitutive promoters with different expression strengths (either from alcohol dehydrogenase 1, *ADH1*, or triosephosphate dehydrogenase 3, *TDH3*, see Note 2). A kanamycin resistance cassette flanked by *loxP* sites is included as selection marker for chromosomal integration. The *loxP* sites can be used to excise the kanamycin resistance allowing the reusage of the insertion cassette in an already tagged strain (15). In addition, both vectors carry the gene for the β-lactamase, conferring resistance to the antibiotic ampicillin and an origin of replication for the amplification in *Escherichia coli*. The vectors (accession numbers P30598 and P30600) are deposited at the EUROSCARF yeast strain collection (http://web.unifrankfurt.de/fb15/%20mikro/euroscarf/index.html), which is available from Scientific Research and Development (Oberursel).

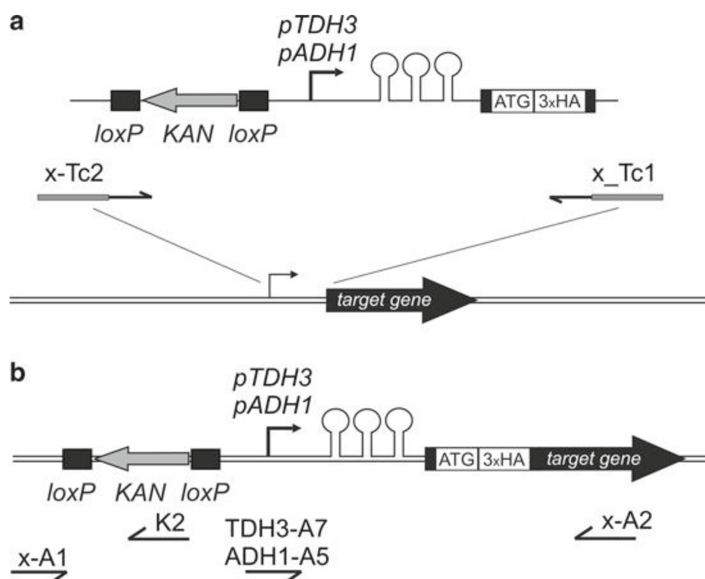


Fig. 2. Universal insertion cassettes for tc aptamer-mediated regulation. (**a**) Schematic representation of the insertion cassette with a kanamycin-resistance gene (*KAN*) flanked by *loxP* sites. Three copies of the tc aptamer followed by an HA tag (3xHA) for protein quantification are expressed by a constitutive promoter. All plasmid-encoded insertion cassettes can be PCR amplified with a single pair of target gene-specific primers (x-Tc2 and x-Tc1). x-Tc2 shares 40 bases homology with the promoter and x-Tc1 40 bases homology with the open reading frame of the target gene (*grey boxes*). The double lane represents the target chromosomal locus. (**b**) Chromosomal situation after integration: The gene of interest is expressed by a heterologous promoter (either *pADH1* or *pTDH3*) and controlled by the tc aptamer in its 5′ UTR.

2.2. Yeast Strain, Medium

1. *Saccharomyces cerevisiae* strain: CEN.PK122: *MATa/MATα MAL2-8c/MAL2-8c SUC2/SUC2* ((16), see Note 3).
2. YPD medium: 1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) glucose. Autoclave for 20 min and store at room temperature (RT) until use. For plates, add 1.5% (w/v) agar prior autoclaving. Let the medium cool to around 60°C before adding antibiotics.
3. Tc stock solution: 100 mM tc solution in water. Make fresh solution every time. Protect tc solution and tc-containing medium from light. Add 200 mg/mL geneticin (G418, Roth) directly as powder to the medium.
4. K-Acetate (KAc) plates: Dissolve 1% KAc and 1.5% agar in water. Autoclave for 20 min and store at RT until use.
5. 0.91% NaCl (normal saline): Prepare in H_2O, autoclave for 20 min and store at RT until use.

2.3. Instrument and Reagents for Tetrad Dissection

1. A SINGER micromanipulator is recommended for tetrad dissection (http://www.singerinst.co.uk).
2. 1.5 mg/mL zymolyase: 1.5 mg zymolyase 20 T (20,000 units/g); dissolve in 1 mL 0.5 M sorbitol. Store at –20°C until use.

2.4. Reagents for Molecular Biology

1. PCR reagents: 10× PCR buffer (100 mM Tris–HCl, pH 9.0, 500 mM KCl, 1% Triton X-100, 15 mM $MgCl_2$); 10× dNTPs (0.2 mM each); polymerase mix (3 U *Taq* polymerase with 0.5 U *Pfu* polymerase or other polymerases with proofreading activity).
2. Ethanol precipitation: 3 M NaAc, 98% (v/v) ethanol; 70% (v/v) ethanol.
3. TE buffer: 10 mM Tris–HCl and 1 mM EDTA, pH 8.0.
4. Spin columns for PCR product purification (e.g., Promega, Wizard SV Gel and PCR Clean-up System).
5. 1 M LiAc: 10.2 g of LiAc dihydrate in 100 mL water; sterilize by autoclaving.
6. 50% (w/v) PEG3350: Prepare in H_2O, autoclave for 20 min and store at RT until use.
7. 5 mg/mL carrier DNA: Salmon sperm DNA; denature the carrier DNA for 5 min at 98°C and cool down immediately on ice/water before use.

2.5. Reagents for Western Blotting

1. Polyvinylidenedifluoride (PVDF) membrane (Millipore).
2. 12CA5 anti-HA monoclonal antibody (Roche).
3. Anti-mouse IgG-conjugated horseradish peroxidase (BioRad).

4. ECL Plus Western Blotting Detection System (GE Healthcare).
5. TBST: 10 mM Tris–HCl, pH 8.0, 150 mM NaCl, and 0.05% Tween 20.
6. 5% nonfat dry milk in TBST.

3. Methods

3.1. Amplification of the Insertion Cassettes for Homologous Recombination

The insertion cassettes can be amplified by PCR using the same set of primer pair. x-Tc2 anneals upstream of the resistance cassette and attaches sequences homologous to the upstream region of the target open reading frame (ORF) to the insertion cassette (see Note 4 and Fig. 2). x-Tc1 anneals downstream of the HA tag (see Note 5) and attaches sequences homologous to the ORF to the insertion cassette (Fig. 2). For a successful chromosomal integration, at least 40 nucleotides (nt) of homologous sequences are needed [x-Tc2: 5′-(40 nt homologous sequence)-GCATAGGC-CACTAGTGGATCTG-3′; x-Tc1: 5′-(40 nt homologous sequence)-AAGCTTCGTACGAGCGTAATC-3′]. Note that an alternative start codon is introduced by the insertion cassette. Therefore, the homologous sequence of x-Tc1 has to start with the second codon of the ORF.

After PCR amplification, analyze the reactions by agarose gel electrophoresis. Concentrate the PCR products by ethanol precipitation. Add 1/10 vol. 3 M sodium acetate and 2.5 vol. 98% ethanol to the PCR reaction, vortex and centrifuge at 16,000 × *g* for 15 min at RT. Wash pellet with 70% ethanol and resuspend the air-dried pellet in 10–15 μL of TE. Use 1–5 μg of PCR product to transform *S. cerevisiae*.

3.2. Genomic Integration of the Insertion Cassette

Prepare competent yeast cells as follows:

1. Grow yeast overnight to an OD_{600} of 0.8–2.0 in 100 mL YPD medium.
2. Harvest cells by centrifugation at 3,000 × *g* for 5 min at RT.
3. Wash cells in 40 mL sterile H_2O and collect by centrifugation at 3,000 × *g* for 5 min at RT.
4. Wash cells in 20 mL 0.1 M LiAc and collect by centrifugation at 3,000 × *g* for 5 min at RT.
5. Remove the supernatant and resuspend cells in 1 mL 0.1 M LiAc.
6. Transfer the sample to a microfuge tube and incubate for 10 min at RT.
7. Centrifuge cells at 3,000 × *g* for 1 min.

8. Remove the supernatant and resuspend the cells in 400 μL 0.1 M LiAc.

Transformation:

1. Mix 1–5 μL of DNA (1–5 μg of the amplified insertion cassette) with 50 μL competent yeast cells.
2. Add 10 μL denatured carrier DNA and 300 μL PEG3350. Mix carefully by pipetting and incubate for 30 min at 30°C.
3. Heat shock for 15 min at 42°C.
4. Add 800 μL sterile H_2O, and mix carefully by pipetting.
5. Collect cells by centrifugation at 3,000 × *g* for 1 min.
6. Resuspend the cells in 1 mL YPD medium and incubate for 2–4 h at 30°C.
7. Harvest cells by centrifugation at 3,000 × *g* for 1 min.
8. Resuspend the pellet in 100 μL YPD medium.
9. Plate on YPD plus 200 mg/L G418 plates. Incubate plates for 1–2 days at 30°C (see Note 6).

PCR analysis of G418-resistant transformants/segregants:
Correct integration of the tc aptamer cassette in front of the target gene can be verified by colony PCR of G418 resistant transformants or segregants. The method works very well for the amplification of PCR products smaller than 1 kb.

1. Transfer a small amount of cell material from a colony with a sterile tooth pick in 50 μL 1× PCR buffer and heat for 10–15 min at 96°C.
2. Quickly cool the cell suspension on ice/water.
3. Start a separate 50 μL PCR reaction for each integration site. For the upstream integration site, use the primer combination x-A1/K2, and for the downstream integration site use TDH3-A7/x-A2 or ADH1-A5/x-A2 depending on the promoter (see Note 7).
4. Analyze 10 μL of the PCR reaction by agarose gel electrophoresis. Strains showing the predicted PCR fragments should be further analyzed by sequencing to verify the correct sequence of the tc aptamers and the HA tags.
5. Depending on the promoter used, purify the TDH3-A7/ORFx-A2 or ADH1-A5/ORFx-A2 PCR product, respectively, using spin columns. Sequence purified PCR products either with primer TDH3-A7 or ADH1-A5.

3.3. Tetrad Analysis

1. Grow heterozygous diploid cells carrying a wild-type and a tc aptamer-modified allele of the target gene overnight at 30°C in 5 mL YPD medium.

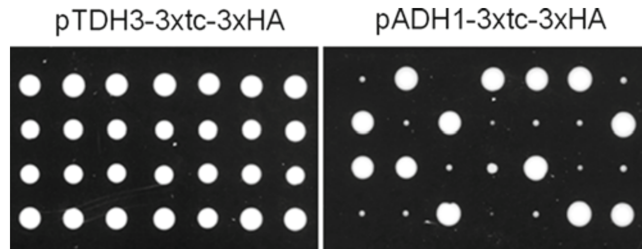


Fig. 3. Segregation pattern of tc aptamer-regulated *ESF2* gene. *E*ighteen *S* rRNA *F*actor 2 (*ESF2*) encodes for an essential nucleolar protein involved in pre-18S rRNA processing. The diploid strain CEN.PK122 was transformed with the tc3-3xHA aptamer insertion cassettes. After meiosis, the viability of the tc-regulated *ESF2* gene was followed by tetrad dissection in the absence of tc. For the *TDH3* promoter (pTDH3)-driven construct, four viable segregants with the same growth behavior were observed. In contrast, the *ADH1* promoter (pADH1)-driven construct results in four viable segregants, but the pADH1-tc3-3xHA segregants showed reduced growth when compared to wild-type segregants (2:2 segregation). It shows that the pADH1-driven construct cannot fully complement the *ESF2* endogenous promoter.

2. Harvest 1 mL culture by centrifugation at 3,000 × *g* for 5 min.
3. After removing the supernatant, resuspend the yeast cells in 100 μL sterile H_2O and spot the suspension on KAc plates.
4. Incubate at 30°C for 2 days and track the appearance of tetrads by microscopy (see Note 8).
5. At this stage, spores can be dissected by a SINGER micromanipulator after removing the ascus wall enzymatically. Resuspend a loopful of sporulated cells in 100 μL zymolyase solution and incubate at RT for 5 min.
6. Add carefully 800 μL sterile H_2O to stop the reaction.
7. Spread a loopful of cells across the full width of a YPD agar plate and dissect tetrads using a micromanipulator.
8. For spore germination and colony growth, incubate for 2–4 days at 30°C (Fig. 3).
9. Subject spores to phenotypic analysis by replica plating on YPD + G418 plates to identify the wild-type and tc aptamer-controlled segregants (see Note 9).

3.4. Serial Dilution Growth Assay

1. Grow cells overnight in YPD medium to an OD_{600} of 1–2.
2. Dilute cells in normal saline to an OD_{600} of 1 followed by tenfold serial dilutions.
3. Spot 5 μL from the diluted cultures on YPD plates in the absence (control) and presence of different tc concentrations. Recommended tc concentrations are 50, 100, 250, and 500 μM.
4. Verify growth differences after incubation of the plates for 2–4 days at 30°C in the dark (Fig. 4). In certain cases, lower tc

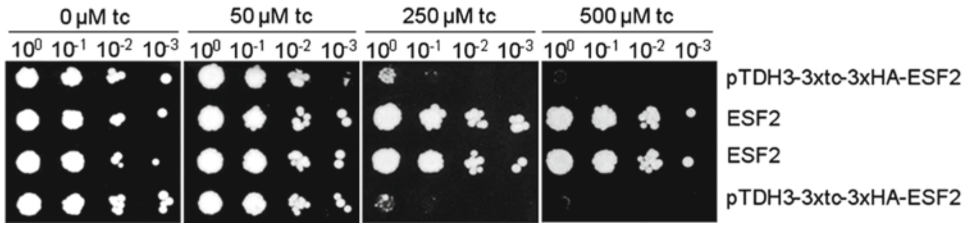


Fig. 4. Serial dilution growth assays performed from wild-type and pTDH3-tc3-3xHA insertion mutants. Dilutions were spotted on YPD agar plates containing increasing concentrations of tc. Colonies were grown for 3 days at 30°C.

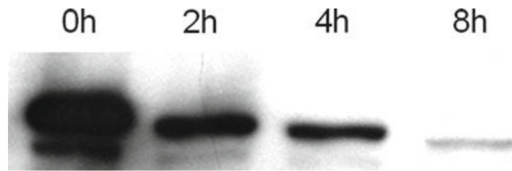


Fig. 5. Repression kinetic of tc aptamer-controlled Esf1 expression. *E*ighteen *S* rRNA *F*actor 1 (*ESF1*) encodes for an essential nucleolar protein involved in pre-18S rRNA processing. 150 μg protein extracts from pTDH3-tc3-3xHA-ESF1 cells (time points after treatment with 250 μM tc are indicated) were loaded on a 10% SDS polyacrylamide gel. After protein separation and transfer on PVDF membrane, 3xHA-Esf1 levels were detected by Western blot analysis.

concentrations are necessary to observe growth differences between wild-type and mutant proteins both under the translational control of the tc aptamer (17).

3.5. Protein Detection by Western Blot

1. Grow HA-tagged strains in 50 mL YPD medium overnight at 30°C to an OD_{600} of 1–2.
2. Next day, inoculate 100 mL fresh YPD medium with the overnight culture to an OD_{600} of 0.1 and incubate for 2.5 h at 30°C.
3. After splitting the culture into two halves, treat one aliquot with 250 μM tc and incubate for further 8–12 h at 30°C.
4. Harvest the cells (e.g., 5 mL every 2 h) and prepare protein extracts using glass beads.
5. Determine total protein concentration with the micro-biuret method (18) or comparable methods.
6. Separate equal amounts of protein (50–150 μg) on a SDS polyacrylamide gel and transfer proteins onto PVDF membrane by semidry blotting.
7. After blocking the membrane with 5% nonfat dry milk for 30 min, add 1:5,000 diluted 12CA5 anti-HA monoclonal antibody for at least 2 h at RT or overnight at 4°C.
8. Remove the first antibody and incubate the membrane for additional 30 min in 5% nonfat dry milk before adding the anti-mouse IgG horseradish peroxidase antibody (1:30,000). Incubate for 30 min at RT. Detect the HA-tagged proteins (Fig. 5) using the ECL chemoluminescence protocol (GE Healthcare).

4. Notes

1. Three repeats of the HA tag are sufficient for the detection of recombinant expressed proteins. The vectors are also available with six repeats of the HA tag which allows detection of smaller amounts of protein.

2. The tc aptamer leads to decreased translation efficiency even in the absence of its ligand. Therefore, it is necessary to drive transcription from strong promoters to compensate for the aptamer-mediated decrease. The transcription strength of the *TDH3* promoter is two- to five-fold higher compared to the *ADH1* promoter (19, 20). It is recommended to use the pTDH3-tc3-3xHA cassette first. If tetrad dissection shows four uniformly growing colonies, the pADH1 cassette should be used additionally to adjust the translation rate more to the wild-type situation. Figure 3 shows an example in which the *ADH1* promoter is not sufficient to complement the substitution of the endogenous promoter resulting in slower growing colonies compared to the wild type.

3. The approach is strain independent; thus, every *S. cerevisiae* strain can be used. Information about further useful *S. cerevisiae* strains can be found on the EUROSCARF Web page under http://web.uni-frankfurt.de/fb15/mikro/euroscarf/data/ref_str.html.

4. x-Tc2 should be designed in a way to delete at least part of the endogenous promoter sequence. If no information about the endogenous promoter is available, it is advisable to delete 200 bp as most yeast 5′ UTRs are below 100 nt (21).

5. Some proteins are negatively influenced by an N-terminal tag. In this case, the HA tag should not be attached. For amplification of the insertion cassette without the HA tag, the primer dHA should be used: 5′-(40 nt homologous sequence)-CATATGTTCTCGAGGCCTAGG-3′. The homologous sequence should start with the second codon of the ORF as described above.

6. A typical transformation plate shows colonies heterogeneous in size 2 days after incubation at 30°C. A few large colonies and much more small colonies are observable. Generally, transformants that form large colonies have integrated the tc aptamer cassette into their genome, whereas the small colonies are the result of transient transformation without integration of the tc aptamer cassette into the genome. High background in the form of small colonies often leads to depletion of nutrients from the medium and prevents growth of the real transformants. To prevent this, it is recommended to replica the transformation plate after 1–2 days on a fresh YPD+G418 plate and incubate further for 1 day at 30°C.

7. Sequences of the common primers used for verification of the correct integration site:

   K2: 5′-GTTTCATTTGATGCTCGATGAG-3′

   THD3-A7: 5′-CTTCTGCTCTCTCTGATTTGG-3′

   ADH1-A5: 5′-CTACTCTCTAATGAGCAACGG-3′

8. This method works very well with the CEN.PK strain background. An adaptation of the methods is necessary when other yeast strain backgrounds, like BY or FY, are used (22).

9. If no micromanipulation system is available, tetrad analysis is also offered commercially (http://www.srd-biotec.de).

Acknowledgments

This work was supported by the Aventis Foundation and the Deutsche Forschungsgemeinschaft (SU402/1-2, SFB579, and the Excellence Cluster: Macromolecular Complexes).

References

1. Johnston, M. and Davis, R.W. (1984) Sequences that regulate the divergent *GAL1-GAL10* promoter in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **4**, 1440–1448
2. Mountain, H.A., Bystrom, A.S., Larsen, J.T. et al., (1991) Four major transcriptional responses in the methionine/threonine biosynthetic pathway of *Saccharomyces cerevisiae*. *Yeast* **7**, 781–803
3. Belli, G., Gari, E., Aldea, M. et al., (1998) Functional analysis of yeast essential genes using a promoter-substitution cassette and the tetracycline-regulatable dual expression system. *Yeast* **14**, 1127–1138
4. Belli, G., Gari, E., Piedrafita, L., Aldea, M. et al., (1998) An activator/repressor dual system allows tight tetracycline-regulated gene expression in budding yeast. *Nucleic Acids Res.* **26**, 942–947
5. Gari, E., Piedrafita, L., Aldea, M.et al., (1997) A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in *Saccharomyces cerevisiae*. *Yeast* **13**, 837–848
6. Dohmen, R.J., Wu, P. Varshavsky, A. (1994) Heat-inducible degron: a method for constructing temperature-sensitive mutants. *Science* **263**, 1273–1276
7. Sanchez-Diaz, A., Kanemaki, M., Marchesi, V. et al., (2004) Rapid depletion of budding yeast proteins by fusion to a heat-inducible degron. *Sci. STKE* **2004**, PL8
8. Roth, A. and Breaker, R.R. (2009) The structural and functional diversity of metabolite-binding riboswitches. *Annu. Rev. Biochem.* **78**, 305–334
9. Suess, B. and Weigand, J.E. (2008) Engineered riboswitches: overview, problems and trends. *RNA Biol.* **5**, 24–29
10. Weigand, J.E. and Suess, B. (2009) Aptamers and riboswitches: perspectives in biotechnology. *Appl. Microbiol. Biotechnol.* **85**, 229–236
11. Hanson, S., Berthelot, K., Fink, B., et al., (2003) Tetracycline-aptamer-mediated translational regulation in yeast. *Mol. Microbiol.* **49**, 1627–1637
12. Suess, B., Hanson, S., Berens, C., et al., (2003) Conditional gene expression by controlling translation with tetracycline-binding aptamers. *Nucleic Acids Res.* **31**, 1853–1858
13. Weigand, J.E. and Suess, B. (2007) Tetracycline aptamer-controlled regulation of pre-mRNA splicing in yeast. *Nucleic Acids Res.* **35**, 4179–4185
14. Kötter, P., Weigand, J.E., Meyer, B., et al., (2009) A fast and efficient translational control system for conditional expression of yeast genes. *Nucleic Acids Res.* **37**, e120
15. Guldener, U., Heck, S., Fielder, et al., (1996) A new efficient gene disruption cassette for

repeated use in budding yeast. *Nucleic Acids Res.* **24**, 2519–2524

16. Entian, K.D. and Kötter, P. (2007) Yeast Genetic Strain and Plasmid Collections. In: Stansfield, I. and Stark, M.R.J. (eds) Methods in Microbiology Vol. 36, Yeast Gene Analysis. Academic Press Ltd., San Diego.
17. Meyer, B., Wurm, J.P., Kötter, P., et al., (2010) The Bowen-Conradi syndrome protein Nep1 (Emg1) has a dual role in eukaryotic ribosome biogenesis, as an essential assembly factor and in the methylation of Ψ 1191 in yeast 18S rRNA. *Nucleic Acids Res.* doi: 10.1093/nar/gkq931
18. Itzhaki, R.F. and Gill, D.M. (1964) A Micro-Biuret Method for Estimating Proteins. *Anal. Biochem.* **9**, 401–410
19. Fraenkel, D.G. (2003) The top genes: on the distance from transcript to function in yeast glycolysis. *Curr. Opin. Microbiol.* **6**, 198–201
20. Partow, S., Siewers, V., Bjorn, S., et al., (2010) Characterization of different promoters for designing a new expression vector in *Saccharomyces cerevisiae. Yeast* **27**, 955–964
21. David, L., Huber, W., Granovskaia, M., et al., (2006) A high-resolution map of transcription in the yeast genome. *Proc. Natl. Acad. Sci. USA* **103**, 5320–5325
22. Stansfield, I. and Stark, M.R.J. (2007) Yeast Genetics and Strain Construction. In: Stansfield, I. and Stark, M.R.J. (eds) Methods in Microbiology Vol. 36, Yeast Gene Analysis. Academic Press Ltd., San Diego.

Chapter 21

Cloning and Expression of Hemicellulases from *Aspergillus nidulans* in *Pichia pastoris*

Prasanna Vasu, Stefan Bauer, and Brett J. Savary

Abstract

The methylotrophic yeast *Pichia pastoris* is increasingly used for heterologous expression of high quality proteins in laboratory-scale (milligram) quantities. Commercially available polysaccharide-active enzyme preparations have limited applications in plant cell wall research due to their heterogeneous mix of hydrolytic activities. *P. pastoris* provides an ideal in vitro expression system for producing monocomponent enzymes, since it lacks endogenous plant cell wall-active enzymes and can perform eukaryotic post-translational modifications (i.e., glycosylation). We have routinely prepared cDNA constructs from *Aspergillus nidulans* encoding a broad array of hydrolases active on various linkages contained in plant cell wall polysaccharides. The cDNAs were inserted into the pPICZα C shuttle vector (Invitrogen) in-frame with the *Saccharomyces cerevisiae* α-secretion factor and expressed under the transcriptional control of the highly inducible alcohol oxidase 1 (AOX1) promoter. The enzyme products were efficiently secreted into buffered complex methanol medium (BMMY) as C-terminal his-tagged proteins for simple one-step affinity purification. The insertion of the c-*Myc* epitope enabled easy immunodetection. Here we present the detailed protocols for primer design, cloning, expression, and activity assays for a representative set of xylan-acting hemicellulases produced in *P. pastoris*.

Key words: *Pichia pastoris*, *Aspergillus nidulans*, Protein expression, Plant cell wall hemicellulase, Enzyme assay, Xylanase, Xylosidase

1. Introduction

Pichia pastoris has become a well-established system for heterologous expression of quality proteins in high levels. *P. pastoris* is a single-celled eukaryotic methylotrophic yeast, capable of metabolizing methanol as the carbon source. Transformation in *P. pastoris* uses a similar set of well-defined molecular and genetic manipulation techniques established for *Saccharomyces cerevisiae*. However, unlike *S. cerevisiae, P. pastoris* can grow to very high cell densities

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_21, © Springer Science+Business Media, LLC 2012

in minimal medium owing to its respiratory growth rather than the fermentative growth (1). Additionally, *Pichia* can perform suitable post-translational modifications (i.e., glycosylation) and is thus generally regarded as simple, fast, and readily manipulated eukaryotic protein expression system (1, 2). The expression of protein is tightly regulated by the methanol-inducible alcohol oxidase 1 (AOX1) promoter, which improves protein yields. Furthermore, this expression system has high genetic stability (i.e., decreased gene deletion) and is capable of efficiently secreting correctly folded heterologous proteins to the culture medium (3). Extracellular secretion of heterologous protein is the preferred option for *Pichia* expression, rather than intracellular cytoplasmic accumulation (4), as it simplifies the downstream processing for protein recovery. The availability of a commercial kit (Invitrogen) also promotes the versatility of *Pichia* as an in vitro expression system.

Commercially available enzyme preparations can have limited applications in plant cell wall polysaccharide research due to their heterogeneous mix of hydrolytic activities (5). Since *P. pastoris* lacks endogenous plant cell wall active enzymes and is able to secrete expressed protein in high levels with significantly low level of endogenous proteins, they are suitable for recombinant expression of cell wall active enzymes (4, 5). Being a saprophytic filamentous ascomycete fungus, *Aspergillus nidulans* secretes a variety of cell wall-active enzymes to survive on dead and decaying plant materials. Thus *A. nidulans* is expected to contain genes encoding diverse hydrolases active on various plant cell wall polymers. The complete genome sequence of *A. nidulans* is now available (6, 7), and a comprehensive set of *A. nidulans* gene sequences (ORFs) were identified based on homology to known cell wall active enzymes (8). We successfully cloned respective complementary DNAs (cDNAs) from *A. nidulans* into *P. pastoris* and expressed them in frame with the *S. cerevisiae* α-secretion factor (mating factor) under the transcriptional control of highly inducible AOX1 promoter of the pPICZα C shuttle vector (8). The enzymes were secreted directly into buffered complex methanol medium and then purified efficiently by use of a C-terminal polyhistidine-tag with metal-chelate affinity chromatography media. The recombinant enzymes also contained a c-*Myc* epitope for easy immunodetection by dot-blot. The one-step purification using metal-chelation chromatography provided milligram quantities of high purity cell wall active enzymes (9). These enzymes were devoid of any contaminating side activities on the respective cell wall polymers, and are thus "monocomponent" (8–10). We present in this chapter the detailed protocols for genomic DNA and mRNA isolation from *A. nidulans*, primer design, cloning and expression in *P. pastoris*, and subsequent purification and activity assays for a representative set of xylan-acting enzymes.

2. Materials

1. *A. nidulans* FGSC A4 (Glasgow WT, from the Fungal Genomic Stock Center, University of Missouri, Kansas City, MO).
2. YPD: 1% Yeast extract, 2% peptone, and 2% glucose (dextrose).
3. Sorbitol buffer: 1 M sorbitol and 0.1 M Na_2 EDTA pH 8.0.
4. Zymolyase 100T (MP Biomedicals).
5. Yeast resuspension buffer: 50 mM Tris–HCl pH 7.4, 20 mM Na_2 EDTA pH 8.0.
6. 10% SDS: dissolve 100 g of electrophoresis-grade sodium dodecyl sulfate in 900 mL water by heating to 68°C, adjust pH to 7.2 with HCl and make up volume to 1 L.
7. 5 M potassium acetate: Mix 60 mL of 5 M potassium acetate (49.07 g potassium acetate in a total volume of 100 mL water), 11.5 mL of glacial acetic acid and 28.5 mL of water.
8. 1 M Tris–Cl: Dissolve 12.11 g Tris base [Tris(hydroxymethyl) aminomethane] in 80 mL of water, adjust pH to 7.4 or pH 8.0 with HCl (6 M) and make the volume to 100 mL.
9. 0.5 M Na_2 EDTA (pH 8.0): Add 18.61 g of Na_2 EDTA·2 H_2O in 80 mL of water, stir vigorously, and adjust the pH to 8.0 with NaOH pellets. Adjust volume to 100 mL.
10. TE buffer (pH 8.0 and pH 7.4): 10 m M Tris–HCl, pH 8.0 or pH 7.4, 1 mM Na_2 EDTA.
11. RNase (e.g., Invitrogen).
12. 3 M sodium acetate: Dissolve 408.3 g sodium acetate trihydrate in 800 mL of water. Adjust pH to 7.0 with dilute acetic acid and adjust the volume to 1 L. Sterilize by autoclaving.
13. Minimal medium: 1 L containing 6.0 g $NaNO_3$, 1.52 g KH_2PO_4, 0.52 g KCl, and 0.52 g $MgSO_4$.
14. Hutner's trace elements: 100 mL contained 5.0 g Na_2 EDTA·2 H_2O, 2.2 g $ZnSO_4 \cdot 7H_2O$, 1.14 g H_3BO_3, 0.5 g $MnCl_2 \cdot 4H_2O$, 0.5 g $FeSO_4 \cdot 7H_2O$, 0.16 g $CoCl_2 \cdot 6H_2O$, 0.16 g $CuSO_4 \cdot 5H_2O$, and 0.11 g $(NH_4)_6MO_7O_{24} \cdot 4H_2O$. Adjust pH to 6.5–6.8 with KOH.
15. Complete medium: Contained minimal medium supplemented with 0.5% yeast extract, 1% peptone, 2% glucose, and Hutner's trace elements (1 mL/L), pH 4.5.
16. Conical and baffled flasks (e.g., Pyrex).
17. Rotary shaking incubator (e.g., Innova 4000, New Brunswick Scientific).
18. Miracloth (Calbiochem).

19. Citrus pectin, larch wood xylan, gum arabic (e.g., Sigma-Aldrich).
20. RNase-free centrifugal tubes.
21. Centrifuges (e.g., Sorvall high speed, and Eppendorf microcentrifuge).
22. Liquid nitrogen.
23. Mortar and pestle.
24. TRIzol® reagent (e.g., Invitrogen).
25. Ethanol: 100% and 75% prepared with DEPC (diethylpyrocarbonate)-treated water.
26. RNAase-free water.
27. Temperature-controllable water bath.
28. Turbo DNA-free™ Kit containing Turbo DNase, 10× Turbo DNase buffer and DNase Inactivation Reagent (Applied Biosystems).
29. UV-Vis Spectrophotometer.
30. SuperScript® First-Strand Synthesis Kit for RT-PCR containing SuperScript™ II reverse transcriptase, 5× First strand buffer, DTT (dithiothreitol), RNaseOUT™, Oligonucleotide (dT) 12–18 primers, dNTP mix, 25 mM $MgCl_2$, DEPC-treated water, RNase H (Invitrogen).
31. PCR Purification Kit (Qiagen).
32. Primer design software "Oligo6" (current version is "Oligo7").
33. Taq polymerase (e.g., Invitrogen).
34. 10× PCR buffer: 750 mM Tris–HCl pH 8.5, 200 mM $(NH_4)_2SO_4$, and 0.1% Tween-20.
35. PCR ready-to-use mixture: 100 μL 10× PCR buffer, 80 μL 25 mM $MgCl_2$, 20 μL dNTP mix (10 mM each), and 800 μL of water.
36. Thermocycler.
37. DNA molecular mass ladder of known concentration (e.g., MassRuler™ DNA Ladder Mix, ready-to-use, 80–10,000 bp, Fermentas).
38. Electrophoresis assembly for DNA electrophoresis. (e.g., BioRad SUB® Cell GT and PowerPac™ Basic).
39. 50× TAE running buffer: 24.2% Tris–HCl pH 7.5, 5.71% (w/v) acetic acid and 3.72% EDTA·2 H_2O, adjust the pH to 8.0. Dilute 50 times (e.g., Research Organics).
40. Expression vector pPICZα C (Invitrogen).

41. 10× Buffer Y+(Fermentas): 330 mM Tris–acetate pH 7.9, 660 mM potassium acetate, 100 mM magnesium acetate, and 1 mg/mL bovine serum albumin (BSA).
42. Restriction enzymes *(Xba*I, *Pml*I, and *Pme*I).
43. Alkaline phosphatase (Shrimp AP, Fermentas).
44. Ligation Kit (Fermentas): 10× ligation buffer, 50% PEG 4000, and T4 ligase.
45. Chemically competent TOP10 *Escherichia coli* strain (Invitrogen).
46. SOC medium (Invitrogen): 2% Tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MnCl_2$, 10 mM $MgSO_4$, 20 mM glucose.
47. Low salt LB plates: 1% Tryptone, 0.5% yeast extract, 0.5% NaCl, and 1.5% agar.
48. Zeocin™ (Invitrogen).
49. TE buffer: 10 mM Tris–HCl pH 8.0, 1 mM EDTA.
50. Centrifugal vacuum concentrator (e.g., Speed-Vac from Savant).
51. Big Dye® Terminator Kit (Applied Biosystems).
52. Alkaline lysis buffer (0.2 N NaOH, 1% SDS).
53. Phenol:chloroform:isoamyl alcohol (25:24:1, as a ready-to-use mixture saturated with 100 mM Tris–HCl pH 8.0; contains ~0.1% 8-hydroxyquinoline, e.g.,, from Sigma-Aldrich).
54. Centri-Sep™ spin columns (Princeton Separations).
55. 10× Buffer B+(Fermentas): 100 mM Tris–HCl pH 7.5, 100 mM $MgCl_2$ and 1 mg/mL BSA.
56. *P. pastoris* X-33 (Invitrogen).
57. YPDS plates: 1% Yeast extract, 2% peptone, 2% dextrose, 1 M sorbitol, 2% agar and 100 μg/mL of Zeocin.
58. Electroporation cuvettes: Gene Pulser/Micro Pulser, 0.2 cm gap (BioRad).
59. Electroporation device *E. coli* Pulser™ Transformation Apparatus (BioRad).
60. Falcon conical tubes with screw-cap lid, 50 mL, sterile.
61. Nitrocellulose membrane (e.g., Schleicher & Schuell).
62. Instant non-fat dry milk powder (e.g., Carnation from Nestlé).
63. TBS buffer: 1.21 g Tris–HCl (10 mM) pH 7.4, 8.77 g NaCl (150 mM) in 1 L.

64. Anti-c-*Myc* antibody produced in rabbit (Sigma-Aldrich), ImmuniPure® Antibody, host: Goat Anti-, Antigen: Rabbit lgG (H+L), label: horseradish peroxidase (Thermo Scientific).
65. SuperSignal West Pico Chemiluminescence Kit (Thermo Scientific).
66. Imaging system (e.g., LAS-4000, Fujifilm).
67. Yeast nitrogen base with ammonium sulfate (YNB, Difco).
68. BMGY: 10 g Yeast extract, 20 g peptone, 13.4 g yeast nitrogen base, 10 mL of glycerol and 100 mM phosphate buffer pH 6.0 in 1 L.
69. BMMY: BMGY where glycerol was replaced with 5 mL of 100% methanol.
70. MMY: Is BMMY, where phosphate buffer was omitted.
71. BMM: Is BMMY without yeast extract and peptone.
72. MM: Is BMM where phosphate buffer was omitted.
73. Centrifuge bottles, 250 mL.
74. Stirred cell ultrafiltration unit with 10 kDa membrane (e.g., Amicon).
75. Nickel-chelation affinity chromatography media (e.g., His Bind Resin, Novagen).
76. Centricon centrifugal ultrafiltration devices, 10 kDa NMWCO, 2 mL (Millipore).
77. NuPAGE electrophoresis system consisting of Bis-Tris gels, LDS sample loading buffer, reducing agent, antioxidants, MOPS buffer (Invitrogen).
78. SimplyBlue™ Safestain (Invitrogen).
79. Lowry reagent A: 2.0 g NaOH and 10.0 g Na_2CO_3 in 500 mL of water.
80. Lowry reagent B: 1.0 g $CuSO_4{\cdot}5H_2O$ and 2.0 g sodium citrate in 100 mL of water.
81. F-C reagent (Folin Ciocalteu phenol reagent, Sigma-Aldrich).
82. BSA (bovine serum albumin): Pre-diluted protein assay standards (Pierce).
83. Various *p*NP-α-glycosides or *p*NP-β-glycosides substrates (Sigma-Aldrich).
84. Bicinchonic acid (BCA) Reagent A: 10.856 g Na_2CO_3, 4.84 g $NaHCO_3$ and 0.3884 g BCA sodium salt in 200 mL of water.
85. BCA Reagent B: 0.2496 g $CuSO_4{\cdot}5H_2O$ and 0.2524 g L-serine in 200 mL of water.

3. Methods

The standard molecular cloning methods (11) are used as required throughout this chapter in combination with the *Pichia* protocols (12) and protocols in the Invitrogen manuals (13, 14).

3.1. Preparation of Genomic and cDNA from A. nidulans

3.1.1. Preparation of Genomic DNA from A. nidulans

1. Grow *A. nidulans* in 10 mL YPD medium at 37°C with shaking at 150 rpm overnight.
2. Centrifuge 5 mL of the culture media at 2,000 × *g* for 5 min and remove supernatant.
3. Resuspend cells in a microfuge tube with 0.5 mL sorbitol buffer and add 20 μL of Zymolyase 100T (2.5 mg/mL in sorbitol buffer). Incubate for 1 h at 37°C.
4. Centrifuge at 16,800 × *g* for 1 min and remove supernatant.
5. Resuspend cells in 0.5 mL yeast resuspension buffer, and add 50 μL of 10% SDS.
6. Invert the closed tube several times rapidly and incubate at 65°C for 30 min.
7. Add 0.2 mL of 5 M potassium acetate and store the tube on ice for 1 h.
8. Centrifuge at 16,800 × *g* for 5 min at 4°C to pellet the cell debris and transfer supernatant to a fresh microfuge tube using a wide-bore pipet (e.g., cut the tip of a plastic pipet tip).
9. Add an equal volume of isopropanol, mix well, and keep at room temperature (RT) for 5 min to precipitate DNA. Do not allow the reaction to go more than 5 min!
10. Centrifuge at 16,800 × *g* for 10 s to recover the precipitated DNA. Remove the supernatant by aspiration and air-dry the pellet for 10 min.
11. Dissolve pellet in 0.3 mL TE buffer pH 8.0 containing 20 μg/mL RNase, and incubate at 37°C for 30 min.
12. Add 30 μL of 3 M sodium acetate pH 7.0, mix the solution, and add 0.2 mL isopropanol. Mix again and centrifuge at 16,800 × *g* for 20 s.
13. Remove supernatant by aspiration and air-dry pellet for 10 min. Dissolve in 150 μL TE buffer pH 7.4.

3.1.2. Preparation of cDNA from A. nidulans mRNA

1. Inoculate *A. nidulans* mycelia from a culture plate into 1 L of complete medium pH 4.5 in 2.8 L Fernbach flasks.
2. Incubate at 37°C for 3 days with shaking at 200 rpm.
3. Recover mycelium by filtering with sterile Miracloth.
4. Transfer mycelium to minimal medium pH 4.5 containing Hutner's trace elements (1 mL/L) and 0.5% (w/v) pectin,

0.5% (w/v) larch wood xylan or 2% (w/v) gum arabic as the carbon source (see Note 1).

5. Incubate at 37°C for 24 h with shaking at 200 rpm.
6. Pellet cells in aliquots by centrifugation in RNase-free centrifuge tubes at 12,000 × *g* for 5 min at 4°C (see Note 2). Do not wash the pellet.
7. Freeze the pellet in centrifuge tube using liquid nitrogen and grind pellet with mortar and pestle, which was cooled in liquid nitrogen before use.
8. Add TRIzol® reagent (ten times the volume of the tissue mass) and incubate for 5 min at RT.
9. Add 0.2 mL of chloroform for each 1 mL of TRIzol reagent used, mix capped tubes vigorously by hand for 15 s and incubate for 3 min at RT.
10. Centrifuge at 12,000 × *g* for 10 min at 4°C and transfer aqueous (upper) phase to a fresh tube.
11. Add 0.5 mL of isopropanol for each 1 mL of TRIzol reagent used, mix, and incubate for 10 min at RT followed by centrifugation at 12,000 × *g* for 10 min at 4°C.
12. Remove supernatant and wash RNA pellet with 1 mL of 75% ethanol for each 1 mL TRIzol reagent used, centrifuge at 7,000 × *g* for 10 min at 4°C.
13. Discard supernatant, briefly dry pellet for 10 min and dissolve in RNase-free water.
14. Remove DNA using the Turbo DNA-free™ Kit by mixing 87 μL of RNA solution, 10 μL of 10× Turbo DNase buffer, 3 μL of Turbo DNase, and incubate at 37°C for 60 min.
15. Add 20 μL of DNase Inactivation Reagent and incubate at RT for 5 min, with intermittent mixing (two to three times) during this period.
16. Centrifuge at 10,000 × *g* for 1.5 min and transfer supernatant containing RNA into new tube, then immediately proceed to the next step (RT-PCR) or store at −80°C.
17. Determine RNA concentration by measuring A260/280 ratio of the solution using a spectrophotometer. The ratio should usually be around 2.0.
18. To perform first-strand cDNA synthesis, set up first-strand cDNA synthesis reaction mixture in a 0.2 mL nuclease-free microfuge tube. Mix 1–5 μg of RNA with 1 μL of Oligo(dT) 12–18 primer (500 μg/mL) and 1 μL of dNTP mix (10 mM each of dATP, dGTP, dCTP, and dTTP), and add DEPC-treated water to 12 μL. Mix well by shaking. For convenience, all the following heating steps (steps 19–23) can be performed in the thermocycler.

19. Incubate at 65°C for 5 min and place immediately on ice.
20. Add the following to each tube on ice: 4 μL of 5× First-Strand Buffer, 2 μL of 0.1 M DTT, 1 μL of RNaseOUT™ (40 U/μL). Mix the contents of the tube gently by shaking and incubate at 42°C for 2 min.
21. Add 1 μL (200 U) of SuperScript™ II Reverse Transcriptase, mix gently by pipetting up and down, and incubate at 42°C for 50 min.
22. Inactivate the reaction by heating at 70°C for 15 min.
23. To remove RNA complementary to the cDNA, add 1 μL (2 U) of *E. coli* RNase H and incubate at 37°C for 20 min.
24. Use PCR purification Kit (Qiagen) according to the manufacturer's instruction to purify the genomic DNA or cDNA products.

3.2. Primer Design

Based on the sequence homology (available from the NCBI taxonomy browser) to known fungal cell wall hydrolase genes (7), we identified a large set of sequences from the *A. nidulans* genome database (8). Those which are significantly best hits to known cell wall hydrolases were selected and used to design primers for cloning these enzymes into *P. pastoris* (5, 8). We chose pPICZα C (see Note 3) (13) to clone all the cell wall hydrolases. You can similarly identify your gene homolog of interest from *A. nidulans.*

For primer design, we use the primer design software Oligo6, which allows the analysis of existing restriction sites within the target sequence as well as possible intra- or intermolecular primer–primer interactions and other criteria (see Note 4). Since most genomic DNA (from Subheading 3.1.1) and cDNA (from Subheading 3.1.2) sequences lacked *Pml*I and *Xba*I sites, these restriction enzyme sites can be used to clone most genes of interest into the multiple cloning site (MCS) of pPICZα C.

The complete coding sequence of each target gene is cloned, including any signal sequence. For this the 5′ (forward or upper) primer sequence is annealed at the ATG start codon. Additional nucleotides matching the template sequence of the genes are appended at the 3′ end of the primer to reach the desired target melting temperature range of 56–60°C. In order to be in line with the reading frame, the restriction site for *Pml*I (CACGTG) requires two nucleotides be added before the ATG codon. The nucleotides A and T are used since their contribution to the melting temperature was lowest. For enzyme recognition, five additional nucleotides (GAAAG) adjacent to the restriction site are attached. As a result, the common design for a primer using *Pml*I restriction site is GAAAG**CACGTG**AT**ATG**X_n, where X_n is 12–21 nucleotides from the target template's 5′ end.

A similar strategy is followed for the 3′ (reverse or lower) primer containing the *Xba*I site (TCTAGA). The stop codon in the target sequence is deleted, including one additional nucleotide from the preceding codon triplet, in order to be in line with the reading frame and to be contiguous with the c-*Myc* and 6× histidine tag coding sequence. The primer sequence therefore ends 1 nucleotide before the stop codon of the target sequence. Again, the primer sequence matching the target sequence consists of 15–24 nucleotides so that the resulting melting temperature will be in the range of 56–60°C, matching the upper primer. The *Xba*I primer has the general structure TACAG**TCTAGA**Y_n, where Y_n is 15–24 nucleotides from the target template's 3′ end, excluding the four last nucleotides.

3.3. Polymerase Chain Reaction

Genomic DNA or cDNAs are amplified by polymerase chain reaction (PCR) using gene specific primers generating *Pml*I and *Xba*I restriction sites at 5′ and 3′ ends, respectively. PCR is performed using the designed primers. Store all PCR solutions at −20°C and keep on ice when thawed. The PCR reaction mixture contains 18.2 μL Ready-to-use mixture, 0.8 μL forward primer (10 μM), 0.8 μL reverse primer (10 μM), 0.1 μL genomic DNA or cDNA, and 0.12 μL Taq polymerase. To obtain maximum efficiency, it is important to work with an optimal PCR program. A typical Touchdown PCR program used is detailed below (see Note 5):

1. 96°C for 2 min.
2. 94°C for 20 s.
3. 65°C for 15 s (−1°C/cycle).
4. 72°C for 1 min (extend 1 min for each additional kb product).
5. 9× to step 2 (i.e., repeat step 2–5, nine times).
6. 94°C for 20 s.
7. 55°C for 20 s.
8. 72°C for 1 min (extend 1 min for each additional kb product).
9. 30× to step 6 (i.e., repeat step 6–9, 30 times).
10. 72°C for 30 min.
11. 4°C ∞.

Check and confirm the size of the PCR product by running a 1% agarose gel electrophoresis (see Note 6). Once the presence of a selected sequence has been confirmed in a genomic or cDNA preparation, set up 12 additional PCR reactions targeting this sequence, then combine after completion of the cycles.

3.4. Cloning into E. coli

1. Use PCR Purification Kit to purify the PCR products. PCR results can be checked on a 1% agarose gel.

2. Set up restriction digest reaction mixture (20 μL reaction) by mixing 10–16 μL of purified PCR product or pPICZα C vector with 2 μL of buffer Y + and 1 μL each of *Xba*I and *Pml*I. Add water to make up to 20 μL.
3. Mix well by pipetting gently up and down several times (important!); briefly centrifuge to collect the liquid at the bottom and incubate overnight at 37°C.
4. For dephosphorylation of pPICZα C plasmid, add 2.5 μL of 10× reaction buffer and 2 μL shrimp alkaline phosphatase (2 U) only to the digest containing the pPICZα C vector (20 μL), incubate at 37°C for 1 h and heat inactivate at 65°C for 20 min. This helps in reducing self-ligation due to incomplete restriction reaction.
5. Use PCR Purification Kit to purify all reactions, check aliquots on 1% agarose gel and estimate the concentration by comparing to the DNA molecular weight ladder.
6. To set up ligation reaction, mix 2 μL of 10× ligation buffer, 2 μL of 50% PEG 4000, 3 μL of T4 DNA ligase with 13 μL of water containing purified PCR product (from step 3 of this section) and pPICZα C vector (from step 4 of this section). The molar ratio of PCR product (50–400 ng) to vector should be around 1 : 3. Incubate at RT for 1 h and heat inactivate at 65°C for 10 min.
7. For transformation, pipet 5 μL of the ligation mixture into a vial containing 100 μL of TOP10 chemically competent *E. coli* cells, and place on ice for 30 min. Incubate at 42°C (heat shock) for exactly 30 s and immediately place the cells back on ice. Add 200 μL of SOC medium and incubate the vial at 37°C for 3 h with moderate shaking.
8. Spread 10, 50 and 200 μL of this media on to different low salt LB plates (pre-heated to 37°C) containing Zeocin (25 μg/mL) and incubate at 37°C overnight or until colonies are visible.
9. For each transformation, transfer 24–48 colonies to a new LB plate (daughter plate) containing Zeocin and perform colony PCR (see Note 7) using 5′AOX1 and lower PCR primer, respectively.
10. Estimate the size of the PCR product on a 1% agarose gel (Fig. 1a).

3.5. Plasmid Extraction and Sequencing

1. Pick candidates (inoculate aseptically using a sterile toothpick or 10 μL micropipet tip) showing expected size of the insert (and thus having the correct orientation) from the daughter plate and grow in 5 mL of low salt LB media containing Zeocin (25 μg/mL) at 37°C overnight with vigorous shaking.
2. Pellet the cells by centrifugation (16,800 × *g* for 10 min) and extract plasmid DNA by alkaline lysis. Resuspend pellet in

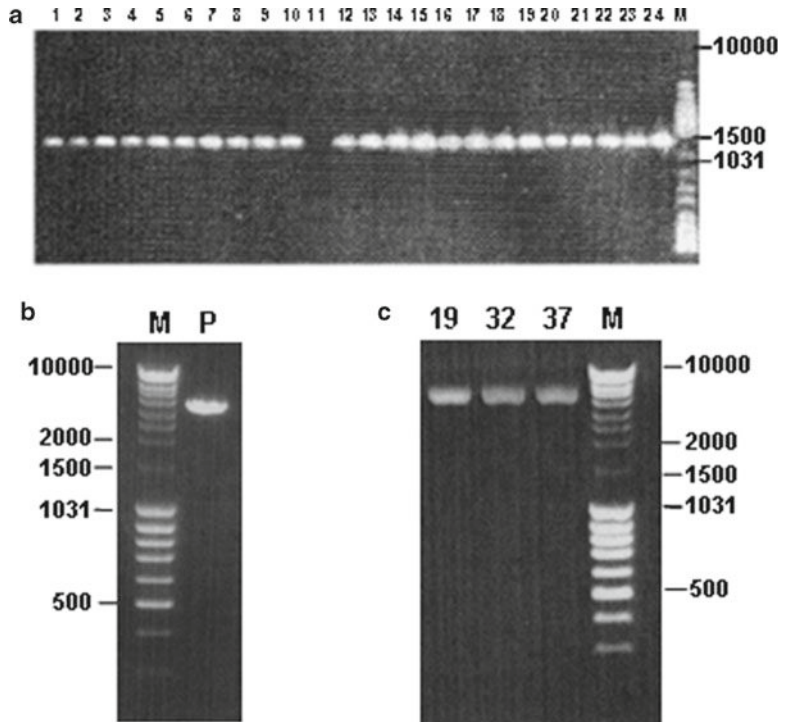


Fig. 1. Example of a colony PCR and confirmation for a typical construct. (**a**) Verification of correct transformants by colony PCR. One out of twenty-four colonies is negative. (**b**) Plasmid pPICZα C (P) digested with the restriction enzymes; *Pml*I and *Xba*I. (**c**) Based on the screening of 24–48 colonies for each construct, the best three colonies selected were grown overnight in low salt LB media containing Zeocin (25 μg/mL). Mini-prep DNA extractions (alkaline extraction/phenol/chloroform) were made and pellets dissolved in 25 μL of RNase water (20 μg/mL). DNA extracted (1 μL) from three different colonies (lanes 19, 32, 37) was loaded. Molecular weight ladders (M) are used to determine the correct size of the recombinant pPICZα C plasmid containing inserts. The increase in the size of the construct is observed.

100 μL 25 mM Tris–HCl pH 8.0 containing 10 mM EDTA and 150 mM glucose. Add 200 μL of freshly prepared lysis buffer (0.2 N NaOH, 1% SDS) and gently mix by inverting the tube several times (do not vortex!) and place on ice for 5 min.

3. Add 150 μL of ice-cold potassium acetate (pH 4.8), gently mix by inverting the tube several times, and let the tube stand on ice for 5 min. Centrifuge at 16,800 × *g* for 10 min and collect the supernatant.
4. Add an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) and mix by vortexing; then centrifuge (16,800 × *g* for 10 min) to achieve phase separation. Collect aqueous upper phase and add an equal volume of chloroform. Vortex and centrifuge (16,800 × *g* for 10 min) and collect the upper phase.
5. Add 0.1 volumes of 3 M sodium acetate pH 5.2 and 2 volumes of 100% ethanol and incubate at –20°C for 30 min. Centrifuge

at 16,800 × *g* for 10 min, discard supernatant, and wash the pellet with 200 μL of 70% ethanol by inverting the tube several times.

6. Pellet the cells by centrifugation (16,800 × *g* for 10 min) and air-dry the pellet for 10 min at RT and resuspend pellet in 50 μL of TE buffer containing 20 μg/mL RNAse. Check and verify the size and concentration of DNA on a 1% agarose gel using 2 μL solution (Fig. 1b–c). Store the DNA solution at −20°C if not used immediately.
7. Set up 20 μL of sequencing reaction in PCR tube using the Big Dye® Terminator Kit. Mix 8 μL Terminator Ready Reaction Mix, 2 μL primer (1.6 μM), 200–500 ng plasmid and water up to 20 μL.
8. Perform the PCR reaction. A typical PCR reaction temperature profile is 25 cycles of 96°C for 30 s, 50°C for 15 s, 60°C for 4 min, followed by cooling to 4°C until ready to purify.
9. Use Centri-Sep™ spin columns according to manufacturer's instruction to purify sequencing products and dry sample using a Speed-Vac. Samples are then submitted for sequencing using the dye-terminator sequencing method (Big Dye® Terminator Kit, see Note 8).

3.6. Isolation of Target Plasmid and Linearization

1. Grow *E. coli* clone showing an error-free insert in 50 mL of low salt LB containing 25 μg/mL zeocin at 37°C overnight with vigorous shaking. Extract plasmid according to the above extraction protocol (see steps 2–7 in Subheading 3.5). Increase volumes used to resuspend and lysis by a factor of 3.
2. Mix 20 μL of isolated plasmid (template), 5 μL of 10× buffer B+, 22 μL of water and 3 μL of *Pme*I to linearize plasmid by incubating overnight at 37°C.
3. Use PCR Purification Kit to purify and elute with water, check aliquots on 1% agarose gel, and estimate the size by comparing to the DNA mass ladder.

3.7. Preparation of Competent Cells and Transformation by Electroporation

1. Grow wild type *P. pastoris* X-33 (a Mut^+ His^+ prototropic strain) in 10 mL YPD medium in a 50 mL flask at 30°C with shaking at 280 rpm overnight.
2. Transfer 0.5 mL of this culture into 500 mL YPD medium in a 2.5 L flask and grow at 30°C and 280 rpm overnight until OD_{600} 1.3–1.5.
3. Centrifuge at 1,500 × *g* for 5 min at 4°C to pellet the cells, and then resuspend cells in 500 mL ice-cold water.
4. Centrifuge at 1,500 × *g* for 5 min at 4°C to pellet the cells, and then resuspend cells in 250 mL ice-cold water.
5. Centrifuge at 1,500 × *g* for 5 min at 4°C to pellet the cells, and then resuspend cells in 20 mL ice-cold 1 M sorbitol.

6. Centrifuge at 1,500 × *g* for 5 min at 4°C to pellet the cells, and then resuspend cells in 1 mL ice-cold 1 M sorbitol resulting in around 2 mL final volume. Keep competent cells on ice and use same day for electroporation to ensure high efficiency transformation. Using frozen competent cells, which are stored at −80°C, is not recommended.
7. Mix 80 μL of competent cells with 5–10 μg linearized plasmid (from step 3 in Subheading 3.6), transfer into an ice-cold electroporation cuvette (0.2 cm gap), then place the cuvette on ice for 5 min.
8. Pulse the cells with 1.5 kV charging voltage (use a capacitance of 25 μF and a resistance of 200 Ω, or use protocol according to manufacturer's instructions for *S. cerevisiae*) and immediately add 1 mL of ice-cold 1 M sorbitol.
9. Transfer content of cuvette into a sterile 15 mL tube and incubate at 30°C without shaking for 2 h.
10. Spread 10, 50, and 200 μL each of this solution on YPDS plates containing 100 μg/mL zeocin and incubate plates at 30°C until colonies are visible (usually 3–5 days).
11. Isolate 16 colonies on YPDS plates with 100 μg/mL zeocin (streak for single colony).

3.8. Screening for Protein Expression by Immunodetection (Dot-Blot)

1. Inoculate each zeocin-resistant clone (from step 11 in Subheading 3.7) aseptically using a sterile toothpick or 10 μL micropipet tip to 5 mL of BMGY (growth medium) in 50 mL sterile conical tubes with lids, and grow overnight at 30°C in a shaking incubator at 280 rpm.
2. Pellet cells (1,500 × *g*, 5 min) and resuspend in 5 mL of induction media (MM, MMY, BMM, or BMMY) in 50 mL conical tubes and grow at 28°C at 280 rpm overnight.
3. Centrifuge cells at 3,000 × *g* for 5 min and take 1 mL of supernatant and store on ice.
4. Spot 2 μL of each supernatant on nitrocellulose membrane and let air-dry. Block membrane with 5% fat-free milk in TBS buffer containing 0.1% Tween-20 at RT for 2 h.
5. Add first antibody (anti-c-*Myc*, 20 μL in 10 mL of 5% fat-free milk in TBS buffer containing 0.1% Tween-20) and incubate for 3 h at RT or at 4°C overnight.
6. Wash membrane 7 times (5 min each) with TBS buffer containing 0.1% Tween-20.
7. Add secondary antibody (anti-rabbit IgG peroxidase conjugated, 5 μL in 15 mL of 5% fat-free milk in TBS buffer containing 0.1% Tween-20) and incubate for 3 h at RT or at 4°C overnight.

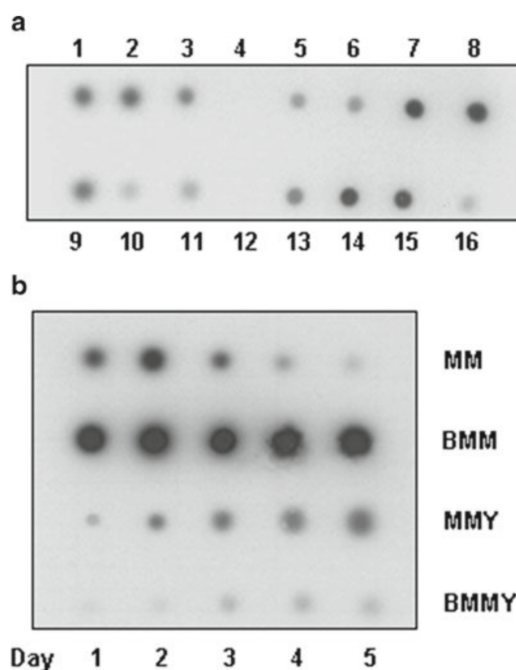


Fig. 2. Reprsentative immuno dot-blot detection of *Pichia*-expressed protein using the c-*Myc* epitope. (**a**) Protein expression levels for 16 clones from the same transformation are assessed after 24 h of induction to select the best expressing clone (i.e., clones 7, 8, 14, or 15). (**b**) Protein expression level of a selected clone grown with different induction media (MM, BMM, MMY, BMMY) over 5 days induction time (see Note 9). Differences in protein expression levels are clearly observed by intensities for different media.

8. Wash membrane 7 times (5 min each) with TBS buffer containing 0.1% Tween-20.
9. Transfer membrane into a ziplock bag and add 1–2 mL of SuperSignal West Pico Chemiluminescence mixture.
10. Wait for 5 min and visualize chemiluminescence using imaging system, LAS-4000, Fujifilm (Fig. 2) (see Note 9).

3.9. Protein Expression in Induction Media (Small-Scale Trial)

The transformants showing the highest protein expression are selected for small-scale expression and purification of recombinant proteins.

1. Inoculate a single colony aseptically using a sterile 10 μL micropipet tip to 1 mL of YPD medium in 5 mL sterile tubes, and grow at 30°C at 280 rpm overnight.
2. Transfer aseptically this 1 mL of YPD culture medium directly into 200 mL BMGY (growth) medium in 500 mL baffled flasks, and grow at 30°C at 280 rpm overnight.
3. Pellet the cells by centrifugation at 5,000×*g* for 5 min at 4°C in 250 mL sterile centrifuge bottles, and resuspend cell pellet in 200 mL induction medium (BMM, MM, MMY, or BMMY based on dot-blot analysis) in 500 mL flasks to induce expression.

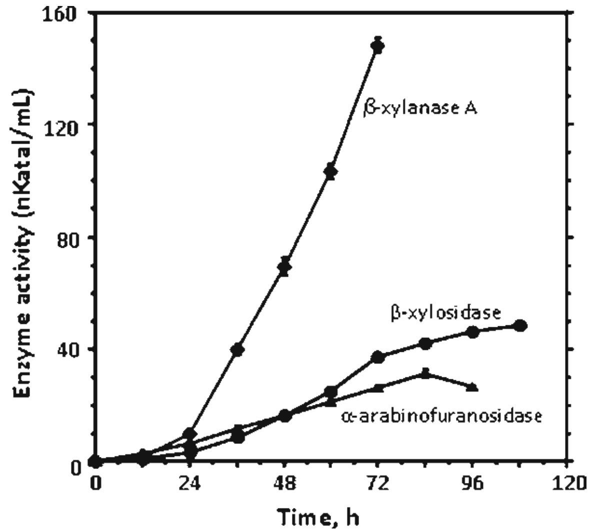


Fig. 3. Enzyme activity profiles for three *Aspergillus nidulans* xylanolytic hemicellulases expressed in *Pichia* cultures during induction in BMMY media. Representative enzymes are: β-xylanase A (AN3613.2), β-xylosidase (AN2359.2), and α-arabinofuranosidase (AN1571.2).

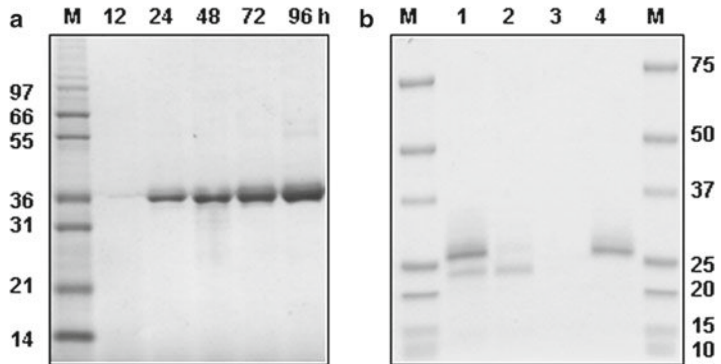


Fig. 4. Representative SDS-PAGE gels for expressed (secreted) and purified xylanolytic enzymes. (**a**) SDS-PAGE of culture supernatant (20 μL) expressing recombinant β-xylanase (AN1818.2) during induction (*lanes* indicate 12, 24, 48, 72, and 96 h of induction) in BMMY media. (**b**) SDS-PAGE profile of a typical enzyme purification using metal (Ni) chelation chromatography: 1, Supernatant concentrate. 2, Unbound column flow-through. 3, Buffer wash. 4, Eluted pure protein. *M* molecular weight markers (molecular masses are indicated in kDa).

4. Add 1 mL of 100% methanol (0.5% final concentration) to each flask and shake well and grow at 28°C at 280 rpm for 3–6 day.
5. Add 1 mL of 100% methanol to each flask every 24 h and pipet 1 mL aliquot every 12 h (0, 12, 24, 36 … and so on), preferably prior to methanol addition.
6. Analyze growth (OD_{600}) and protein expression levels [by the respective enzyme activities (Fig. 3) and SDS-PAGE (Fig. 4a)]

to determine the optimal time to harvest. Store the aliquot at 4°C until ready to analyze.

7. Once the culture reaches the maximum induction level (usually 3–6 days of induction), pellet the cells (centrifuge at 5,000 × *g* for 5 min at 4°C) in 250 mL centrifuge bottles.
8. Collect the supernatant and add sodium azide to a final concentration of 0.02%. Store at 4°C until use. Discard the pellet/cells according to safety protocol (see Note 10).

3.10. Purification of Recombinant Enzymes

3.10.1. Concentration of Protein by Ultrafiltration or Ammonium Sulfate Precipitation

1. Concentrate the supernatant (induction medium containing expressed enzymes, stored at 4°C) at least tenfold using an ultrafiltration device, such as an Amicon stirred cell, equipped with a 10-kDa NMWCO membrane, according to the manufacturer's instruction (see Note 11).
2. Alternatively precipitate protein with $(NH_4)_2SO_4$. For this add $(NH_4)_2SO_4$ to supernatant to 80% saturation (at 4°C) with constant stirring. After overnight incubation at 4°C, centrifuge the suspension (10,000 × *g* for 30 min at 4°C), collect the protein pellet, and dissolve in 10–20 mL of sodium acetate buffer (0.02 M, pH 5.0) (see Note 12).

3.10.2. Purification of Recombinant Enzymes Using Metal (Nickel) Chelation Column

1. Wash the His-Bind Resin (Novagen) (see Note 13) column (0.7 × 5 cm) extensively with water.
2. Charge with nickel by flushing the column with 10 mL of 400 mM $NiSO_4$.
3. Wash and equilibrate the column with 10 mL of the starting buffer (Buffer A: 20 mM Tris–HCl pH 7.9 and 500 mM NaCl) containing 10 mM imidazole (see Note 14).
4. Adjust the pH of the concentrated supernatant to pH 7.9 with NaOH (see Note 14) and load on to the Ni^{2+}-charged His-Bind column at 4°C.
5. Wash the column with 10 mL of Buffer B (Buffer A containing 50 mM imidazole).
6. Elute the bound His-tagged protein with 10 mL of Buffer C (Buffer A containing 250 mM imidazole).
7. Repeatedly wash the eluted fraction with 0.02 M acetate buffer pH 5.0 in order to remove imidazole and concentrate using Amicon stirred-cell ultrafiltration units. Further concentration to small volumes can be achieved by 10-kDa NMWCO Centricon centrifugal ultrafiltration devices.
8. Sample aliquot is assayed for protein content and enzyme activities and stored in the freezer. Purity of the protein is assessed by SDS-PAGE (Fig. 4b).

3.11. Analyses of Growth and Protein Expression (SDS-PAGE, Protein Content, and Enzyme Activities)

Growth is monitored by measuring absorbance of the culture media at 600 nm.

3.11.1. Determine Culture Growth by Absorbance (Optical Density) at 600 nm

1. Pipet 100 μL culture medium to clean glass test tube and add 900 μL fresh induction medium. Mix by pipetting up and down.
2. Blank spectrophotometer with induction medium then read absorbance of suspended cell solution at 600 nm.

3.11.2. SDS-PAGE Analysis (15)

Protein expression is monitored by SDS-PAGE. We use the NuPAGE Bis-Tris gel electrophoresis system by Invitrogen, but the traditional Laemli buffer system is suitable.

1. To 13 μL of media supernatant of different time points, or suitably diluted purified samples, add 5 μL of 4× LDS sample loading buffer and 2 μL of 10× sample reducing agent and incubate at 70°C for 10 min.
2. Load 10–20 μL of these samples to 10 or 12% Bis-Tris gels (1 mm thickness), and perform electrophoresis at 200 V (constant) for 55 min using MOPS buffer system, according to the manufacturer's instruction (15).
3. Stain the gels by SimplyBlue Safestain™ to visualize the proteins (see Note 15).
4. Transfer the gel to ziplock bag for scanning (Fig. 4) and store at 4°C.

3.11.3. Protein Estimation by Lowry's Method (16)

1. Pipet 10 μL of pre-mixed standards (10–100 μg BSA) or samples into different glass test tubes and make up the final volume to 200 μL using deionized water.
2. Add 1 mL of Lowry reagent (mix Lowry reagent A and B; 49:1, v/v) into these tubes, shake well, and incubate for 10 min at RT.
3. Add 0.1 mL of F-C reagent (diluted 1:1 with deionized water) into the tubes and shake to mix. Allow the test tubes to stand at RT for 45 min.
4. Blank spectrophotometer with water-reagent blank and read absorbance at 650 nm.

3.11.4. Micro-assays for Glycosidase Activities (α-/β-Arabinofuranosidases, α-/β-Xylosidases)

1. Pipet 14 μL of *p*NP substrate (*p*NP-α arabinofuranoside, *p*NP-β-arabinofuranosidase, *p*NP-α-xylosidase, or *p*NP-β-xylosidase) prepared as 3 mg/mL solution into a microcentrifuge tube.
2. Add 5 μL of 0.2 M sodium acetate buffer pH 5.0 and 1 μL of enzyme solution (culture supernatant).
3. Incubate the tubes at 37°C for exactly 15 min.

4. Stop the reaction immediately by adding 980 μL of 0.2 M Na_2CO_3.
5. Blank spectrophotometer with water-reagent blank and read absorbance at 420 nm.

3.11.5. Micro-assays for Glycanase Activities (β-Xylanases)

1. Pipet 10 μL of different 10 mg/mL substrate solutions (larch wood xylan).
2. Add 5 μL buffer 0.2 M sodium acetate buffer pH 5.0, then 5 μL of suitably diluted enzyme solution (culture supernatant) into these tubes.
3. Incubate the tubes at 37°C for exactly 20 min.
4. Stop the reaction immediately by adding 980 μL BCA working reagent (reagent A, reagent B, and water, 2:2:1 (v/v/v)).
5. Incubate the tubes at 80°C for 30 min. Cool to RT for about 20 min.
6. Blank spectrophotometer with water-reagent blank and read absorbance at 560 nm.

4. Notes

1. Different carbon sources are used to induce transcription of specific mRNAs in *A. nidulans*. We tried several other carbohydrate sources (e.g., carboxymethyl cellulose, oat spelt xylan, tamarind xyloglucan), but we observed better transcription of mRNA in the above-mentioned media.
2. Use sterile, disposable, certified RNase-free centrifuge tubes and pipet tips, automatic pipettors dedicated for RNA preparation, and wear disposable gloves in order to avoid contamination with RNase from labware and human skin. Avoid polycarbonate tubes for RNA precipitation as they are not resistant to isopropanol. Add solvents at RT to minimize co-precipitation of salts. The centrifugation steps should be performed at 4°C to prevent warming of the sample.
3. The expression vector (pPICZα C) is a shuttle vector of 3.6 kb, which propagates both in *E. coli* and *P. pastoris*. Important features of the plasmids include (1) the promoter of AOX1 gene, as the homologous sequence with the yeast genome for recombination, (2) the transcriptional termination sequence of AOX1 gene for efficient processing and polyadenylation of mRNAs, (3) the secretion signal sequence from the *S. cerevisiae* α-factor prepro-peptide sequence, which leads to secretion of the recombinant protein into external medium, (4) the MCS with 10 unique restriction sites for the insertion of the foreign gene

between the two AOX regions, (5) the C-terminal (c-) *Myc* epitope and polyhistidine tag downstream of MCS for detection and purification of the recombinant protein, (6) the pUC *ori* for maintenance and replication in bacterial hosts, (7) the unique restriction sites *Bgl*II, 5′ to the AOX1 promoter sequence, and *Bam*HI, 3′ to AOX1TT sequence, facilitate the generation of in vitro multimers. Additionally, restriction enzymes that cut one time in 5′ AOX1 region (*Pme*I, or *Sac*I or *BstI*) can be used to linearize the plasmid, and finally (8) the single, small, dominant selectable marker gene, (*Sh ble* gene) that confers resistance to Zeocin, which functions as selectable marker in both bacteria and yeast. The expression is driven and controlled by the EM7 promoter (for *E. coli*), TEF1 promoter (for yeast), and CYC1TT transcription termination regions in the plasmid. The gene of interest is inserted in the yeast genome by a single cross-over type insertion via shared sequences (5′ AOX1 region). Pichia Expression Kit (Invitrogen) contains 20 μg (40 μL) of pPICZα C plasmid. For long-term storage, transform TOP10 *E. coli* strain with pPICZα C (according to Subheading 3.4).

4. In general, the melting temperature (T_m) of the primers should be between 55 and 65°C. The simple formula for estimating T_m is T_m (°C) = 2 × (A + T) + 4 × (G + C). For best results, both primers used in PCR should have a melting temperature within 5°C of each other. The primers should not form hairpin loops, have no tendency to dimerize, and ideally lack a secondary binding site within the template. Further considerations involve the introduction of a GC clamp at the 3′, where the sequence ends in either G or C, but should not have more than 3 G or C in the last five nucleotides. Poly G, poly C, poly A, poly T as well as polypyrimidine (T & C) and polypurine (A & G) should also be avoided since this could cause either non-specific annealing or lower the efficiency of amplification. Ideally, the primer should consist of a random mix of nucleotides with a GC content of 45–55%.

5. When T_m is not known with certainty, a Touchdown PCR is used to optimize yields of amplified DNA, because choosing annealing temperatures too low or too high can lead to primer-dimer formation, non-specific products, or reduced yield due to poor primer annealing. Touchdown PCR uses a cycling program with varying annealing temperatures. The annealing temperature in the initial cycle is chosen to be 5–10°C above the T_m of the primers. This favors only accumulation of duplexes where primer–template complementation is highest. In subsequent cycles, the annealing temperature is decreased in steps of 1°C/cycle until the temperature is around 2–5°C below the

T_m of the primers. In this way touchdown PCR enhances the specificity of the initial primer–template duplex formation and hence the specificity of the final PCR product.

6. To cast 1% agarose gels, dissolve 1 g agarose (electrophoresis grade) in 100 mL of 1× TAE running buffer, and boil the solution until all agarose has dissolved completely. Though the solution appears clear, boil for another 30 s and let it cool for approximately 60–70°C. Add 5 μL of ethidium bromide (10 mg/mL) and mix thoroughly. Pour to the gel tray according to the manufacturer's guide. DNA electrophoresis is run at a voltage of 140 V for 45–60 min or 150 V for 30–45 min.
7. Colony PCR is used to check the presence of cloned insert in the transformants. By using a primer binding inside the vector sequence (5′AOX1), only inserts having the right orientation will show the expected size on the agarose gel. A convenient way of performing colony PCR is to use sterile 10 μL pipet tips to touch the parent colony and transfer the cells onto new numbered YNB plates containing 25 μg/mL zeocin, followed by pipetting into a 20 μL PCR set-up mixture. For 24 reactions, we usually prepare a 500 μL PCR mixture containing the respective primer pair, Ready-to-use mixture, and Taq polymerase in the proportions mentioned above. The PCR mixture is aliquoted into wells of a 96-well plate placed on ice.
8. We have used a dye-terminator sequencing method (Big Dye Terminator Kit, ABI PRISM™ Big Dye™ dideoxy nucleotide termination sequencing, Applied Biosystems) as it was the protocol recommended by the core facility. Good sequencing for around 600–700 base pairs was obtained downstream of each primer. Sequence information of the first ~30 bp is inaccurate, so place primers 40–50 bp upstream from sequence of interest. For inserts around 1.0–1.2 kb only 5′AOX1 and 3′AOX1 primers are used as sequencing primers. Larger inserts require additional primers designed to anneal around 400 bp apart from these primers to cover entire sequence.
9. Immunodetection by dot-blot is used to monitor the protein expression level of each clone. Based on the protein expression levels of 16 different clones of the same transformation (same construct) determined after 24 h of induction, the best expressing clone is selected. It is grown in BMGY media overnight and the recovered cell pellet is suspended in different induction media (MM, BMM, MMY, and BMMY) for a 1–5-day induction period. The best induction medium is then determined for each clone (Fig. 2b). Note that BMMY may not always be the best medium for highest protein expression. Select the best media for expressing the protein of interest based on dot-blot results.

10. Inactivate/sterilize the cells by autoclaving or by chemical bleaching, as recommended by your institution's biosafety protocol.
11. Amicon stirred cell ultrafiltration units with 10- or 3-kDa NMWCO membranes (such as YM) also work fine for concentration. Do not use regenerated cellulose membrane for concentrating xylanases and cellulases, as they may degrade the membrane.
12. Aliquots (pellets and final supernatant) are set aside for determining the protein content and assay for enzyme activities. For long-term storage, enzyme solutions are made to 25% or 50% glycerol, and stored in 50 mL Falcon conical tubes at –20°C.
13. We have tried different metal chelation columns Ni charged His-Bind column (Novagen), Ni-NTA column (Qiagen), Ni charged IDA column (POROS-M®, Boehringer Mannheim), and His-link protein purification resin (Promega). All worked well for most of the His-tagged proteins. A typical purification protocol is presented (Fig. 4b). Resin can be regenerated following the manufacturer's instructions.
14. The pH of the buffer should be between pH 7.5 and pH 7.9 to keep the histidine (pKa is 6.0) in ionized form to bind chelated nickel ions. In buffered media (BMMY or BMM) a precipitate of phosphate salts may be formed after pH adjustment. Remove by centrifugation or filtration. All buffers tested for affinity binding supported good binding. These include Tris buffer (20 mM Tris–HCl pH 7.9, 500 mM NaCl containing 10 mM imidazole), phosphate buffer (50 mM potassium phosphate pH 7.9, 500 mM NaCl containing 10 mM imidazole), and HEPES buffer (100 mM HEPES–HCl, pH 7.5, 500 mM NaCl containing 10 mM imidazole). Binding can be performed by either flow column chromatography or by batch column chromatography, where the resin is pre-mixed with protein solution prior to column elution. If the protein is not binding to the column, reduce the imidazole concentration in wash buffer. Elution can be performed either by single-step elution with 100 mM, 250 mM, and 500 mM imidazole or with continuous gradient elution (from 10 to 750 mM imidazole) in starting buffer.
15. The microwave technique involves washing the gel with 100 mL water in a loosely closed container using a microwave (for ca. 1 min) and shaking at RT for 1 min, three times each. After the third washing, gels are stained with 30 mL of SimplyBlue Safestain by microwaving for 1 min and shaking for 10 min at RT, after which they are de-stained with 100 mL water for several hours. While microwaving gel in solution, switch power off and on as needed to avoid boiling and

damaging the gels. Gels can be stored at 4°C for short-term storage until they can be processed for peptide mass fingerprinting analysis by MALDI-TOF MS with trypsin digestion, as described in Chapter 2 of this book.

Acknowledgments

The authors thank Drs. Andrew J. Mort (Oklahoma State University) and Christopher R. Somerville (University of California, Berkeley) for their support and supervision of this study. This was funded by the U.S. Department of Energy, Basic Energy Sciences Division grant DE-FG02-03ER15444 (to A.J.M. and C.R.S.) and USDOE Office of Energy Efficiency and Renewable Energy, Mid-South/Southeast Biofuels Consortium, DE-FG36-08BO88036 (to B.J.S.) Postdoctoral fellowships for S.B. provided by the Josef-Schormüller-Gedächtnisstiftung and the Deutsche Forschungsgemeinschaft.

References

1. Higgins, D. R. and Cregg, J. M. (1998). Introduction to *Pichia pastoris,* in *Methods in Molecular Biology*: *Pichia* Protocols (Higgins, D. R. and Cregg, J. M., eds.), Humana, Totowa, NJ, Vol. **103**, pp. 1–15.
2. Cereghino, J. L. and Cregg, J. M. (2000). Heterologous protein expression in the methylotrophic yeast *Pichia pastoris. FEMS Microbiol Rev.* **24**, 45–66.
3. Cregg, J. M., Vedvick, T. S., and Raschke, W. C. (1993). Recent advances in expression of foreign genes in *Pichia pastoris. BioTechnol.* **11**, 905–910.
4. Schutter, K. D., Lin, Y-C., Tiels, P., Hecke, A. V., Glinka, S., Weber-Lehmann, J., Rouze, P., Peer, Y. V., and Callewaert, N. (2009). Genome sequence of the recombinant protein production host *Pichia pastoris. Nature Biotechnol.* **27**, 561–566.
5. Mort, A. J., Gao, S., Wu, X., and Prade, R. (2005). The need for pure enzymes that degrade cell wall polysaccharides and how to get them. *Plant Biosystems* **139**, 84–87.
6. Galagan, J. E., Calvo, S. E., Ma, L-J., Wortman, J. R., Batzoglou, S., Lee, S-I., et al., (2005). Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae. Nature* **438**, 1105–1115.
7. *Aspergillus* Sequencing Project. Broad Institute of MIT and Harvard (http://www.broad.mit.edu).
8. Bauer, S., Vasu, P., Persson, S., Mort, A. J., and Somerville, C. R. (2006). Development and application of a suite of polysaccharide-degrading enzymes for analyzing plant cell walls. *Proc. Natl. Acad. Sci. USA* **103,** 11417–11422.
9. Bauer, S., Vasu, P., Mort, A. J., and Somerville, C. R. (2005). Cloning, expression, and characterization of an oligoxyloglucan reducing end-specific xyloglucanobiohydrolase from *Aspergillus nidulans. Carbohydr. Res.* **340**, 2590–2597.
10. Vasu, P., Bauer, S., Somerville, C. R., and Mort, A. J. (2011). Characterization and substrate specificity of recombinant xylanolytic enzymes: Potential analytical tools for the study of plant cell wall xylans. (Revised for Biochem. J).
11. Sambrook, J. and Russell, D. W. (eds.) (2001). Molecular Cloning: A Laboratory Manual, 3rd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
12. Cregg, J. M. and Russell, K. A (1998). Transformation, in *Methods in Molecular Biology*: *Pichia* Protocols (Higgins, D. R. and Cregg, J. M., eds.), Humana, Totowa, NJ, Vol. **103**, pp. 27–39.
13. Invitrogen (2009). EasySelect™ Pichia Expression Kit – For expression of recombinant proteins using pPICZ and pPICZα in *Pichia pastoris*, User Manual, Version I (Cat. No. K1740-0).

14. Invitrogen (2008). pPICZ A, B, and C – Pichia expression vectors for selection on Zeocin™ and purification of recombinant proteins, User Manual, Version E (Cat. No. V190-20).
15. Invitrogen (2003). NuPAGE Technical Guide – General information and protocols for using the NuPAGE electrophoresis system. Instruction Manual, Version E (Cat. No. IM-1001).
16. Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J. (1951). Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* **193**, 265–275.

Chapter 22

A Thiamine-Regulatable Epitope-Tagged Protein Expression System in Fission Yeast

Tiina Tamm

Abstract

Schizosaccharomyces pombe, the fission yeast, has been a popular and useful model system for investigating the mechanisms of biological processes for a long time. To facilitate purification, localization, and functional analysis of gene products, a wide range of expression vectors have been developed. Several of these vectors utilize the inducible/repressible promoter systems and enable the episomal expression of proteins as fusion proteins with epitope tags attached to their N terminus or C terminus.

This chapter provides a detailed protocol for expression of the epitope-tagged proteins from thiamine-regulatable *nmt* promoter in fission yeast. The yeast culture conditions and procedures for yeast transformation, expression induction, preparation of whole-cell extracts, and analysis of epitope-tagged protein expression by Western blotting are described.

Key words: Regulatable promoter, *nmt1* promoter, Epitope tagging, E2 epitope tag, Expression vector, Yeast transformation, *Schizosaccharomyces pombe*

1. Introduction

Transcriptional regulation is one of the main mechanisms by which cells control the gene expression. The ability to regulate transcription has also been widely utilized to study the function of genes. In addition, the modulation of gene expression enables the analysis of the gene function in specific time periods or ensures the expression of toxic genes. In the industry, the exploitation of the regulatable promoters enables the separation of the growth and production stages in manufacturing process of heterologous proteins.

Several inducible/repressible promoter systems have been developed for expressing proteins in fission yeast (see Table 1). As is shown in Table 1, they fall into two groups: those with induction

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_22, © Springer Science+Business Media, LLC 2012

Table 1
Regulatable promoter systems for fission yeast

Promoter, references	Regulation	Time for induction[a]	Available episomal vectors[b], references
Modified CaMV35S (1, 2)	Tetracycline and anhydrotetracycline inducible	6–12 h[c]	pSLF101 (*LEU2*) (3); pDUAL-tet (*ura4*⁺) (2)
ctr4⁺-*CYC1* (4)	Bathocuproinedisulfonate inducible; copper repressible	3 h	p*ctr4*⁺-X (*LEU2*, *ura4*⁺ or *his7*⁺) (4)
fbp1⁺ (5)	Glucose repressible	7.5 h	pCHY21 (*URA3*) (5)
hsp16⁺ (6)	Heat shock, cadmium, ethanol, and oxidative stress inducible	3 h[d]	pHIL (*LEU2*), pHIU (*URA3*) (6)
inv1⁺ (7)	Glucose repressible	1 h	pINV1 (*LEU2*) (7)
nmt1⁺ (full-strength) *nmt41* (intermediate) *nmt81* (weak) (8, 9)	Thiamine repressible	16 h[e]	pREP, pREPX (*LEU2* or *ura4*⁺) (3, 9); pDUAL (*ura4*⁺) (10)
nmt185 (11)	Thiamine repressible; repressed at 36°C, induced at 25°C	3 h[f]	pREP3/185 (*ura4*⁺) (11)
urg1⁺ (12)	Uracil inducible	0.5 h[g]	NA

NA not available
[a] Time needed for the maximum induction of reporter protein
[b] The available selection markers are shown in parentheses
[c] Induction time depends on the inducer type and used media
[d] Measured during heat-shock induction
[e] Determined by Northern blot analysis for *nmt1*⁺
[f] Measured during temperature shift from 36 to 25°C in thiamine-free media
[g] mRNA levels were determined using microarrays

times from 0.5 to 3 h and those with induction times longer than 6 h. Regulation of promoters varies also quite considerably: some of the promoters are inducible and others are repressible. These differences need to be considered in selecting promoter system and designing protein expression experiments.

The most important selection criterion of promoter system is the time needed for induction. The promoter systems with shorter induction times include promoters from *h*eat-*s*hock *p*rotein 16 (*hsp16*⁺), *inv*ertase (*inv1*⁺), and *u*racil-*r*egulatable *g*ene 1 (*urg1*⁺) genes and hybrid promoter between copper-signalling elements from *c*opper *tr*ansporter (*ctr4*⁺) promoter and the minimal promoter region of *Saccharomyces cerevisiae* *cy*tochrome *C* (*CYC1*) (4, 6, 7, 12). The modified plant viral cauliflower mosaic virus (CaMV) 35 S promoter, *f*ructose-1,6-*b*is*p*hosphatase (*fbp1*⁺), and *n*o *m*essage in *t*hiamine (*nmt1*⁺) gene promoters have longer induction times (1, 5, 8).

The promoter systems listed in Table 1 have been developed for use in the cultures in the exponential growth phase. Their use under other growth conditions may result in uncontrolled induction of the transcription. For example, the derepression of the promoter can occur during nitrogen starvation (e.g. *urg1*$^{+}$ promoter) and meiosis (e.g. *fbp1*$^{+}$, *urg1*$^{+}$ promoters) or when cells enter stationary phase (e.g. *fbp1*$^{+}$, *hsp16*$^{+}$ promoters) (5, 6, 13, 14). In some cases, the expression levels do not remain stable over a long period of time. Such an example is *inv1*$^{+}$ promoter system (7). Genes regulated by the *inv1*$^{+}$ promoter are fully induced within 1 h after the shift to sucrose-based medium and *inv1*$^{+}$ promoter is itself inactivated by rising glucose concentrations usually after 4 h of induction (sucrose is converted into glucose). Conversely, limiting glucose concentration in the medium leads to an uncontrolled induction of the promoter. This is usually the case when cells are grown on the agar plates and glucose is depleted locally around the colonies.

Some induction conditions, like elevated temperature, high copper levels, presence of Cu^{1+} chelator bathocuproinedisulfonate (BCS), and various stress conditions, trigger also a general cellular stress response and changes in a genome-wide gene expression (15, 16). This problem can be avoided by using *urg1*$^{+}$ promoter system (12). The genome-wide transcription analysis has indicated that the addition of uracil has little effect on global gene expression. However, the transcription from *urg1*$^{+}$ promoter is rapidly (within less than 30 min) activated and inactivated in response to the addition or removal of uracil, respectively. Unfortunately, there are no episomal vectors available for this system yet.

The *nmt1*$^{+}$ promoter was the first described regulatable promoter and it has remained the most widely used promoter till the present day (8). This strong promoter is repressible by addition of thiamine (vitamin B1) into a medium. Maximum repression is achieved at thiamine concentration of 15 μM. The promoter is induced by removing thiamine and the induction takes several generations, approximately 16–20 h. The slowness of induction is due to the slow decay of the intracellular pool of thiamine (17). Partial induction of the *nmt1*$^{+}$ promoter is achieved by using the intermediate amounts of thiamine (e.g. 0.05 μM) (18).

A few modified versions of the *nmt1* promoter have been constructed. Mutations in the TATA box region have resulted in a decrease of promoter strength (9). Variants *nmt41* and *nmt81* yield moderate or low-level expression of proteins, respectively (3, 9). A truncated version of the *nmt1* promoter – *nmt185* which is still repressed by thiamine – exhibits a temperature-dependent induction in the absence of thiamine (11).

Although thiamine causes over 100-fold repression of wild-type *nmt1*$^{+}$ promoter, it is unable to completely switch off transcription. Therefore, many genes expressed under the control of *nmt* promoters can complement chromosomal mutants even in the presence of thiamine.

The inducible/repressible promoter systems are frequently used in combination with epitope tags. Epitope tagging is a method of choice if antibodies are not available for a given protein. Using recombinant DNA techniques, an epitope sequence for a specific monoclonal antibody is fused to a target gene. The expressed fusion protein can then be characterized, purified, and localized using a commercially available antibody. Several expression vectors with different regulatable promoters have been constructed for N-terminal or C-terminal protein tagging with various epitopes (see Table 2). More information about fission yeast plasmids can be found in ref. 24 and in the Forsburg Laboratory Web site (http://www-rcf.usc.edu/~forsburg/).

Two limitations must be taken into consideration when using regulatable epitope-tagged protein expression system. Firstly, the replacement of endogenous promoter with inducible/repressible one may change the regulation of transcription and/or expression levels of the gene of interest. Secondly, the N-terminal or C-terminal epitope tag may have effect on the functionality of the expressed protein. Therefore, it is recommended to test the functionality of the epitope-tagged protein by complementation experiment.

2. Materials

2.1. Fission Yeast Transformation

1. Fission yeast strains (see Note 1): For the experiments described, wild-type (TYSP72; genotype: *h⁺ leu1.32 ura4.d18*, source: laboratory stock) and temperature-sensitive *sad1.1* (IH274; genotype: *h⁻ sad1.1 leu1.32 ura4.d18*, source: 25) strains were used.
2. Fission yeast expression plasmids (see Note 2): pREP41-E2Nsad1, pREP41-sad1E2C. These plasmids contain *nmt41* promoter (medium level of expression) and terminator, *sad1⁺* coding region in-frame with three copies of E2a epitope tag (pREP41-E2Nsad1 encodes N-terminally E2a-tagged Sad1, pREP41-sad1E2C encodes C-terminally E2a-tagged Sad1), *LEU2* yeast selection marker and fission yeast *ars1⁺* sequence. The construction of these plasmids is described in ref. 19.
3. Yeast extract with supplements (YES5) (see Note 3): 0.5% (w/v) Bacto™ yeast extract, 3% (w/v) glucose, 225 mg/L of adenine, histidine, leucine, uracile, and lysine hydrochloride.
4. 0.1 M lithium acetate, pH 4.9. Adjust pH with acetic acid and sterilize by autoclaving.
5. 50% (w/v) polyethyleneglycol (PEG) 4000 in 0.1 M lithium acetate, pH 4.9.
6. Sterile H_2O.

Table 2
Vectors for epitope-tagged protein expression in fission yeast

Epitope	Position	Promoter[a]	Episomal vector	References
E2Tag of bovine papillomavirus E2	N-term 3×E2a	*nmt41*	pREP	(19)
	C-term 3×E2a	*nmt41* *nmt81*	pREP	(19)
FLAG	N-term FLAG 6×His	*nmt1*$^{+}$	pREPX	(20)
	N-term 2×FLAG 6×His	*nmt1*$^{+}$ *nmt41* *nmt81*	pDUAL	(10)
	C-term FLAG 6×His	*nmt1*$^{+}$	pREPX	(20)
	C-term 2×FLAG 6×His	*nmt1*$^{+}$ *nmt41* *nmt81*	pDUAL	(10)
Hemagglutinin (HA) epitope of the influenza virus	N-term 3×HA	*nmt1*$^{+}$ *nmt41* *nmt81*	pREPX	(21)
		nmt41	pREP	(22)
	C-term 3×HA	*nmt1*$^{+}$ *nmt41* *nmt81*	pREPX	(21)
	C-term 8×His HA	*nmt1*$^{+}$	pREPX	(23)
	C-term 2×HA 6×His	*inv1*$^{+}$	pINV1	(7)
		ctr4-CYC1	p*ctr4*$^{+}$	(4)
c-myc (Myc) epitope of the human c-*myc* proto-oncogene	N-term 6×His 2×Myc	*nmt81*	pREP	(22)
	C-term 12×Myc	*nmt1*$^{+}$	pREPX	(23)
		inv1^{+}	pINV1	(7)
		ctr4-CYC1	p*ctr4*$^{+}$	(4)
Pk epitope of the SV5 P protein	N-term 3×Pk	*nmt41* *nmt81*	pREP	(22)
	C-term 3×Pk	*nmt1*$^{+}$	pREPX	(23)
		nmt41 *nmt81*	pREP	(22)

[a]The strength for *nmt* promoters: *nmt1*$^{+}$ – wild type; *nmt41* – intermediate; *nmt81* – weak

7. Salt stock solution (50×): 0.26 M $MgCl_2 \cdot 6H_2O$, 4.99 mM $CaCl_2 \cdot 2H_2O$, 0.67 M KCl, 14.1 mM Na_2SO_4. Autoclave and store at room temperature (RT).
8. Vitamin stock solution (1,000×): 4.2 mM pantothenic acid, 81.2 mM nicotinic acid, 55.5 mM inositol, 40.8 μM biotin. Store at 4°C.
9. Mineral stock solution (10,000×): 80.9 mM boric acid, 23.7 mM $MnSO_4$, 13.9 mM $ZnSO_4 \cdot 7H_2O$, 7.4 mM $FeCl_2 \cdot 6H_2O$, 2.47 mM molybdic acid, 6.02 mM KI, 1.6 mM $CuSO_4 \cdot 5H_2O$, 47.6 mM citric acid. Store frozen or at 4°C.
10. Edinburgh minimal medium (EMM) (see Note 3): 14.7 mM potassium hydrogen phthalate, 15.5 mM Na_2HPO_4, 93.5 mM NH_4Cl, 2% (w/v) glucose, 20 mL/L salt stock, 1 mL/L vitamin stock, 0.1 mL/L mineral stock; add 225 mg/L supplements as required, except for the selection marker.
11. 4 mM thiamine hydrochloride solution in H_2O; filter sterilize and store in dark at 4°C.

2.2. Protein Expression from Thiamine-Repressible nmt Promoter

1. EMM leu$^-$: EMM without leucine (see Subheading 2.1, item 10).
2. 4 mM thiamine hydrochloride (see Subheading 2.1, item 11).
3. Sterile H_2O.

2.3. Whole-Cell Extract Preparation by TCA Precipitation

1. STOP buffer: 10 mM EDTA, pH 8.0, 1 mM NaN_3, 154 mM NaCl, 50 mM NaF.
2. 20% (v/v) trichloro acetic acid (TCA) (see Note 4).
3. 5% (v/v) TCA.
4. Glass beads (see Note 5).
5. Sodium dodecyl sulphate (SDS) sample buffer: 60 mM Tris–HCl, pH 6.8, 2% (w/v) SDS (see Note 6), 10% (v/v) glycerol, 0.005% (w/v) bromophenol blue. Add 5% (v/v) β-mercaptoethanol just prior to use.
6. 1 M Tris–HCl, pH 8.0.

2.4. Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis

1. 30% acrylamide/bis stock solution: 29% (w/v) acrylamide, 1% (w/v) bisacrylamide (see Note 7).
2. 4× separating buffer: 1.5 M Tris–HCl, pH 8.8, and 0.4% (w/v) SDS.
3. 4× stacking buffer: 1 M Tris–HCl, pH 6.8, and 0.4% (w/v) SDS.
4. 10% (w/v) ammonium persulphate (APS) solution.
5. *N*,*N*,*N*′,*N*′-Tetramethylethylenediamine (TEMED).
6. Running buffer: 25 mM Tris base, 192 mM glycine, and 0.1% (w/v) SDS.

2.5. Protein Transfer to Nitrocellulose Membrane

1. Transfer buffer: 25 mM Tris base, 192 mM glycine, 0.037% (w/v) SDS, 20% (v/v) methanol. Do not adjust pH.
2. Nitrocellulose membrane.
3. 3MM chromatography paper.
4. Ponceau-S red staining solution: 0.2% (w/v) Ponceau S and 3% (v/v) TCA.

2.6. Western Blotting

1. Phosphate-buffered saline (PBS) 10× stock: 1.4 M NaCl, 27 mM KCl, 100 mM Na_2HPO_4, and 18 mM KH_2PO_4. Sterilize by autoclaving and store at RT. Prepare working solution (1×) by dilution of one part of 10× stock with nine parts of H_2O.
2. Blocking solution: 5% (w/v) non-fat milk in PBS.
3. 5E11 E2Tag monoclonal antibody (e.g. Abcam).
4. Anti-α-tubulin monoclonal antibody TAT1 (see Note 8).
5. Goat anti-mouse alkaline phosphatase-conjugated antibody (Sigma).
6. Alkaline phosphatase (AP) buffer: 100 mM Tris–HCl, pH 9.5, 100 mM NaCl, and 5 mM $MgCl_2$.
7. 50 mg/mL 4-Nitroblue tetrazolium chloride (NBT) stock in 70% dimethylformamide (DMF). Prepare 66-μL aliquots and store at −20°C.
8. 50 mg/mL 5-Bromo-4-chloro-3-indolyl-phosphate (BCIP) 4-toluidine salt stock in 100% DMF. Prepare 33-μL aliquots and store at −20°C.

3. Methods

The protocol for epitope-tagged protein expression from the thiamine-regulatable *nmt* promoter can be outlined as comprising three steps: (1) transformation of the expression plasmid into yeast cells to create an "expression strain"; (2) induction of the expression from the thiamine-regulatable *nmt* promoter; (3) analysis of the epitope-tagged protein expression (e.g. complementation assay, Western blotting).

Two fission yeast transformation procedures are described below. The longer procedure uses the liquid cell culture and is more efficient, yielding 10^4 colonies/μg of DNA. Such a protocol is used for plasmid transformation, DNA integration by homologous recombination, and transformation of the libraries. The shorter transformation protocol starts with colonies growing on the plate and it works well only for plasmid transformation, yielding around 10^3 colonies/μg of DNA.

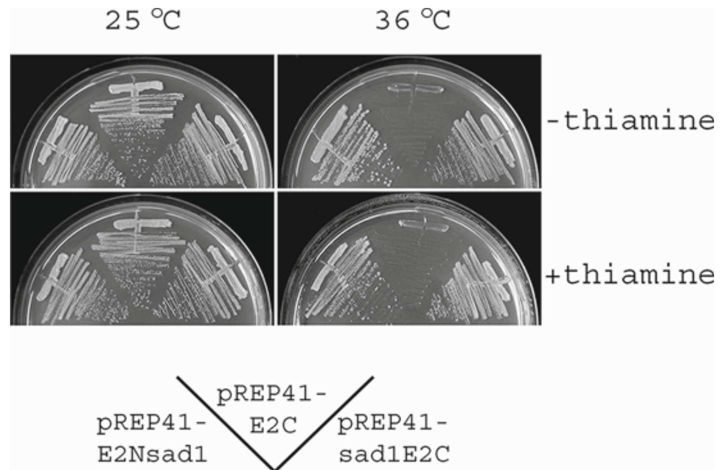


Fig. 1. Complementation assay. Cells of *sad1.1* (strain IH274) containing empty vector (pREP41-E2C), overexpressing N-terminally E2a-tagged Sad1 (pREP41-E2Nsad1) or C-terminally E2a-tagged Sad1 (pREP41-sad1E2C) were grown on EMM leu⁻ plates in the absence (–thiamine) or presence (+thiamine) of 15 μM thiamine. Pictures were taken after 3 and 5 days of growth at 36 and 25°C, respectively.

The expression of E2a-tagged Sad1 is presented as an example. To demonstrate that the E2a tagging of Sad1 is not affecting the protein functionality, the expression plasmids (pREP41-E2Nsad1 encodes N-terminally E2a-taged Sad1; pREP41-sad1E2C encodes C-terminally E2a-tagged Sad1) were introduced into the temperature-sensitive (*ts*) *sad1.1* mutant (strain IH274; 25). Both plasmids succeeded in rescuing the *ts* phenotype of *sad1.1* (see Fig. 1). The complementation occurred also in the presence of thiamine, indicating that repressed *nmt41* promoter allowed a low-level expression of E2a-tagged Sad1.

The results of induction/shut-off experiment are shown in Fig. 2. Wild-type cells (strain TYSP72) carrying the expression plasmid pREP41-sad1E2C were grown in the EMM leu⁻ medium containing 15 μM of thiamine. For induction experiment, the thiamine was removed from the medium. The samples were collected before and after induction, and analyzed by Western blotting. For shut-off experiment, the 15 μM of thiamine was added to the culture after 20 h of expression while collected samples were processed for Western blot analysis at the indicated time points. The promoter shut-off experiment showed that the degradation of E2a-tagged Sad1 is slow and the tagged protein is stable for at least 8 h.

3.1. Fission Yeast Transformation

This yeast transformation protocol represents a modified version of the one described in ref. 26.

3.1.1. Yeast Transformation Using Cell Culture

1. Set up a 10 mL YES5 pre-culture from single colony growing on YES5 plates and grow overnight without shaking at permissive temperature (see Note 9).

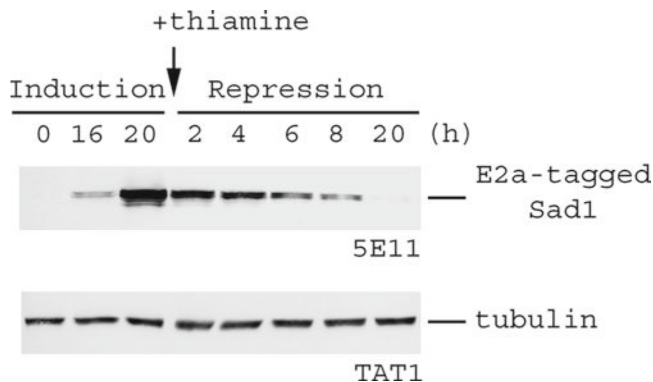


Fig. 2. Western blot analysis of the expression of E2a-tagged Sad1. Wild-type cells (strain TYSP72) carrying the plasmid for C-terminally E2a-tagged Sad1 expression (pREP41-sad1E2C) were grown at 25°C in the EMM leu$^-$ medium containing 15 μM of thiamine. The expression was induced by washing out the thiamine and samples were taken 0, 16, and 20 h after induction (lanes marked 0, 16, and 20, respectively). For the shut-off experiment, 15 μM of thiamine was added after 20 h of expression and samples were taken at indicated time points (lanes marked 2, 4, 6, 8, and 20, respectively). Total protein extracts were separated by 10% SDS-PAGE and transferred to nitrocellulose membrane. Immunoblotting was performed with the mouse E2Tag 5E11 antibody (*top*) or anti-α-tubulin antibody TAT1 as a loading control (*bottom*).

2. Inoculate an appropriate volume of cells from pre-culture into 50 mL of medium so that the cell density following the overnight growth is $0.5–1 \times 10^7$ cells/mL (see Note 10). Use the same growth temperature and medium as for the pre-culture. Grow cells overnight with shaking.
3. Harvest the cells by centrifugation at $1,200 \times g$ for 3 min.
4. Wash the cells in 30 mL of sterile H_2O and in 10 mL of 0.1 M lithium acetate, pH 4.9.
5. Re-suspend the cell pellet in 0.1 M lithium acetate to give a final concentration of 1×10^9 cells/mL. Incubate at permissive temperature for 1 h.
6. Use 100 μL of competent cells for each transformation. Add 1–2 μg of plasmid pREP41-sad1E2C DNA in a volume of 10 μL or less. Add 2.33 vol. of 50% PEG 4000. Mix gently and incubate for 1 h at permissive temperature.
7. Heat shock the sample at 42°C for 10 min.
8. Pellet the cells by centrifugation at $380 \times g$ for 2 min. Optionally wash the cells once in 1 mL of sterile H_2O.
9. Re-suspend the cell pellet in 150 μL of sterile H_2O and plate onto selective EMM leu$^-$ plates supplemented with 15 μM of thiamine.
10. Wrap the plates with parafilm to prevent their drying out and incubate for 3–4 days at permissive temperature.

3.1.2. Yeast Transformation Using Colonies from Plate

1. Pick three to four medium-sized colonies from fresh YES5 plates (less than 1 week old) and re-suspend in 150 μL of sterile H_2O.
2. Pellet the cells by centrifugation at 850 × *g* for 5 min.
3. Wash the cells with 100 μL of 0.1 M lithium acetate, pH 4.9.
4. Pellet the cells by centrifugation.
5. Re-suspend the cell pellet in 100 μL of 0.1 M lithium acetate, pH 4.9.
6. Add 1 μg of plasmid pREP41-sad1E2C DNA and 290 μL of 50% PEG 4000. Mix well and incubate at permissive temperature at least for 30 min.
7. Heat shock the transformation mixture at 42°C for 12 min.
8. Pellet the cells by centrifugation at 380 × *g* for 2 min.
9. Re-suspend the cell pellet in 150 μL of sterile H_2O and plate onto selective EMM leu$^-$ plates supplemented with 15 μM of thiamine.
10. Wrap the plates with parafilm to prevent their drying out and incubate for 3–4 days at permissive temperature.

3.2. Protein Expression from Thiamine-Repressible nmt Promoter

1. Prepare a pre-culture of TYSP72 cells carrying the plasmid pREP41-sad1E2C by transferring a single colony from EMM selective plate to 10 mL of EMM without leucine and supplemented with 15 μM of thiamine. Grow for 1–2 days without shaking at 25°C until cells are in an early stationary phase.
2. Inoculate an appropriate volume of cells from pre-culture into 50 mL of EMM leu$^-$ supplemented with 15 μM of thiamine so that the cell density following the overnight growth in a shaking incubator is approximately 5×10^6 cells/mL (see Note 10). Use the same growth temperature as for the pre-culture.
3. For expression induction, wash cells in 40 mL of sterile H_2O for five times (see Note 11). Inoculate an appropriate volume of cells so that the cell density following the 16-h growth at 25°C is approximately 2.5×10^6 cells/mL. Use EMM leu$^-$ media without thiamine.
4. Collect the time points at 16, 20, and 25 h after induction.
5. For shut-off experiment, add 15 μM of thiamine and collect the time points at 0, 2, 4, 6, 8, and 20 h after repression.

3.3. Whole-Cell Extract Preparation by TCA Precipitation (see Note 12)

The procedure is based on the one described in ref. 27.

1. Pellet the cells (5×10^7 cells/sample) by centrifugation at 2,800 × *g* for 3 min. Wash the cell pellet in 1 mL of STOP buffer and transfer the cell suspension to a 1.5-mL screw-cap tube.
2. Pellet the cells by centrifugation at 18,000 × *g* for 10 s, remove the supernatant, and snap freeze the pellet in liquid nitrogen.

It is possible to store the cell pellet in −80°C or proceed with the whole-cell extract preparation immediately.

3. Add 200 μL of 20% TCA to the cell pellet. Do not re-suspend.
4. Add 500 μL of pre-washed glass beads to each tube.
5. Break cells in a RiboLyser (Hybaid) for 40 s or by vortexing for 2 min using the multi-tube holder.
6. Add 500 μL of 50% TCA.
7. Punch the hole in the bottom of screw-cap tube, place the tube on the top of a new 1.5-mL Eppendorf tube, and spin at 490 × *g* for 2 min to separate the lysate from glass beads.
8. Pellet the proteins by centrifugation at 18,000 × *g* for 5 min at 4°C.
9. Re-suspend the pellet in 70 μL of SDS sample buffer, and add 30 μL of 1 M Tris–HCl, pH 8.0, to balance the acidic pH.
10. Denature the proteins by heating the samples at 95°C for 5 min.
11. Use 3–10 μL of the sample per lane for expression analysis.
12. Snap freeze the remaining sample in liquid nitrogen and store at −80°C.

3.4. Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis

The given recipe is calculated for a 0.75–1.0-mm-thick minigel using 10 × 10-cm glass plates but can be easily modified for other gel formats.

1. Prepare a 10% separating gel solution by mixing 3.6 mL of 30% acrylamide/bis stock, 2.7 mL of 4× separating buffer, and 4.5 mL of H_2O (see Note 13).
2. Add 120 μL of 10% APS and 12 μL of TEMED, mix, and pour the separating part of the gel. Leave the space for stacking part and overlay with 1 mL of distilled H_2O to achieve even gel formation. Allow the gel to polymerize for 30 min and pour off the overlaid H_2O.
3. Prepare the stacking gel solution by mixing 1 mL of 30% acrylamide/bis stock, 1.75 mL of 4× stacking buffer, and 4.2 mL of H_2O.
4. Add 70 μL of 10% APS and 7 μL of TEMED, mix, pour the stacking part of the gel, and insert the comb. Allow the gel to polymerize for 30 min.
5. Assemble the gel in the electrophoresis apparatus, and fill the upper and lower chambers of the unit with running buffer. Using the syringe or Pasteur pipette, wash the wells with running buffer before loading.
6. Load 3–10 μL of protein sample (see Subheading 3.3, step 11) into each well. Include the protein molecular weight marker if desired (see Note 14). Connect the gel unit with the power

supply and run at 80 V until the bromophenol blue dye passes the stacking part of the gel and continue at 150 V until the dye front reaches at the bottom of the gel.

3.5. Protein Transfer to Nitrocellulose Membrane

1. Prepare six sheets of 3MM Whatman filter paper and one nitrocellulose membrane of sizes similar to that of the gel and wet in a transfer buffer.
2. After electrophoresis, remove the gel from the glass plates and trim away the stacking part of the gel. Equilibrate the separating part of the gel in the transfer buffer for 10 min.
3. Place three pieces of filter paper on the anode plate of the semi-dry blotting apparatus; avoid bubbles (see Note 15).
4. Place the wetted membrane on the top of the filter papers. Remove air bubbles if needed.
5. Lay down the gel on the top of the membrane. Eliminate all air bubbles.
6. Place three sheets of the filter paper on the top of the gel; avoid bubbles.
7. Close the blotting apparatus by placing the cathode electrode plate on the top of the stack. Transfer the proteins to the membrane at constant voltage (9 V) for 40 min.
8. At the end of the transfer, discard the filter paper and the gel. Wash the membrane with distilled H_2O.
9. The proteins transferred to the membrane can be visualized by staining with Ponceau-S red. Immerse the membrane in staining solution for 2–4 min (see Note 16). Wash the membrane with distilled H_2O until bands become visible. When using the unstained protein molecular weight marker, mark the positions of the marker bands. The staining can be washed off with distilled H_2O.

3.6. Western Blotting

1. Incubate the membrane in 30–50 mL of blocking solution for at least 1 h at RT on a rocking platform.
2. Wash the membrane once with PBS. Add the E2Tag monoclonal primary antibody 5E11 (1:1,000 dilution in PBS with 1% (w/v) non-fat milk) or anti-α-tubulin monoclonal antibody TAT1 (1:500 dilution in PBS with 1% (w/v) non-fat milk) (see Note 17). Incubate at RT for 2 h or at 4°C overnight on a rocking platform.
3. Wash the membrane four times for 10 min with approx. 50 mL of PBS.
4. Add goat anti-mouse alkaline phosphatase-conjugated secondary antibody (1:10,000 dilution in PBS). Incubate for at least 1 h at RT on a rocking platform.

5. Wash the membrane four times for 10 min with approx. 50 mL of PBS and once for 2 min with 10 mL of AP buffer.
6. Prepare the alkaline phosphatase developing solution by mixing 10 mL of AP buffer, 66 μL of NBT stock, and 33 μL of BCIP stock. Add the developing solution to the membrane and incubate in dark for 10–30 min (see Note 18).
7. Stop the reaction by washing the membrane several times with distilled H_2O. An example of results is shown in Fig. 2.

4. Notes

1. Commonly used fission yeast strains are available at the Yeast Genetic Resource Center of the National BioResource Project (Yeast) (http://yeast.lab.nig.ac.jp/nig/index_en.html) or the Yeast Genetics Stock Culture Center of the American Type Culture Collection, ATCC (http://www.lgcstandards-atcc.org/).
2. Fission yeast plasmids used in transformation experiments are shuttle vectors and consist of a bacterial origin of replication and selectable marker, a yeast selectable marker, and an autonomously replicating sequence (*ars1* and 2 μ *ori* are mainly used). The most frequently used yeast selectable markers are *ura4*$^+$ gene from fission yeast and the *LEU2* gene from budding yeast *S. cerevisiae* complementing *ura4* and *leu1* recessive auxotrophic mutations, respectively. The *ars1* sequence is also required for high frequency of transformation. There are no low-copy centromeric plasmids available because the fission yeast centromere region is too large to be included into a bacterial shuttle vector. Other features in the plasmids may include various promoters and fusion or tagging sequences. More information about plasmids that have been developed for molecular studies in fission yeast is available in ref. 24.
3. For solid medium, add 2% (w/v) of agar. A litre of medium is sufficient for approximately 30 plates.
4. To prepare 100% (w/v) TCA solution, add 227 mL of distilled H_2O to previously unopened bottle containing 500 g of crystalline TCA. Mix until dissolved and keep in dark.
5. Use acid-washed glass beads. Some companies (e.g. Sigma) sell already washed beads.
6. Wear mask when weighing SDS powder.
7. Wear gloves when handling acrylamide. Unpolymerized acrylamide is neurotoxic.
8. TAT1 antibody was generously provided by Prof. Keith Gull (University of Oxford, UK). The preparation of TAT1

monoclonal antibody recognizing specifically α-tubulin is described in ref. 28.

9. The permissive temperatures for wild-type and temperature-sensitive *sad1.1* strains are 32°C and 25°C, respectively.
10. The following formula can be used to calculate the correct volume of pre-culture used for inoculation:

$$\text{Volume of cells inoculated} = \frac{\text{Required density of culture}}{\text{Density of preculture}} \times \frac{\text{Volume of culture}}{2^n}$$

n shows the expected number of generations. If the pre-culture is in stationary phase (cells are non-dividing and density is more than 1×10^7 cells/mL), extend time by one generation time for cells to re-enter the cell cycle following inoculation. The typical mean generation time for wild-type cells is 3 h at 25°C and 2 h 10 min at 32°C when grown in YES5, and 4 h at 25°C and 2 h 30 min at 32°C when grown in EMM. Mutant strains may grow at different rates.

11. For efficient removal of thiamine, wait for 5 min between each washing step.
12. All subsequent steps should be carried out with ice-cold solution at 0–4°C.
13. The different acrylamide/bisacrylamide concentration of the separating gel can be used by varying the amount of 30% acrylamide/bis stock in separating gel solution.
14. Prestained protein molecular weight marker allows monitoring the protein migration during the SDS-polyacrylamide gel electrophoresis (PAGE) and the protein transfer onto membrane.
15. Air bubbles between the gel, filter, and filter papers can be removed by rolling a glass rod across the surface.
16. The Ponceau-S red stain can be reused several times.
17. These antibody dilutions can be reused several times. Collect the primary antibody solution and keep at –20°C.
18. The colour reaction is normally completed within 30 min, but can take up to 1 h or even overnight. Check regularly to avoid the overdeveloping and increasing of the background staining.

Acknowledgments

The author would like to thank Iain Hagan and members of the Hagan laboratory for encouragement, advice, and support and members of the Sedman's group for discussions. This work was supported by Tartu University (baseline funding grant).

References

1. Faryar, K., and Gatz, C. (1992) Construction of a tetracycline-inducible promoter in *Schizosaccharomyces pombe*. *Curr Genet.* **21**, 345–349.
2. Erler, A., Maresca, M., Fu, J., et al., (2006) Recombineering reagents for improved inducible expression and selection marker re-use in *Schizosaccharomyces pombe*. *Yeast* **23**, 813–823.
3. Forsburg, S.L. (1993) Comparison of *Schizosaccharomyces pombe* expression systems. *Nucleic Acids Res.* **21**, 2955–2956.
4. Bellemare, D.R., Sanschagrin, M., Beaudoin, J., et al., (2001) A novel copper-regulated promoter system for expression of heterologous proteins in *Schizosaccharomyces pombe*. *Gene* **273**, 191–198.
5. Hoffman, C.S., and Winston, F. (1989) A transcriptionally regulated expression vector for the fission yeast *Schizosaccharomyces pombe*. *Gene* **84**, 473–479.
6. Fujita, Y., Tohda, H., Giga-Hama, Y., et al (2006) Heat shock-inducible expression vectors for use in *Schizosaccharomyces pombe*. *FEMS Yeast Res.* **6**, 883–887.
7. Iacovoni, J.S., Russell, P., and Gaits, F. (1999) A new inducible protein expression system in fission yeast based on the glucose-repressed *inv1* promoter. *Gene* **232**, 53–58.
8. Maundrell, K. (1990) *nmt1* of fission yeast. A highly transcribed gene completely repressed by thiamine. *J. Biol. Chem.* **265**, 10857–10864.
9. Basi, G., Schmid, E., and Maundrell, K. (1993) TATA box mutations in the *Schizosaccharomyces pombe nmt1* promoter affect transcription efficiency but not the transcription start point or thiamine repressibility. *Gene* **123**, 131–136.
10. Matsuyama, A., Shirai, A., Yashiroda, Y., et al., (2004) pDUAL, a multipurpose, multicopy vector capable of chromosomal integration in fission yeast. *Yeast* **21**, 1289–1305.
11. Kumar, R., and Singh, J. (2006) A truncated derivative of *nmt1* promoter exhibits temperature-dependent induction of gene expression in *Schizosaccharomyces pombe*. *Yeast* **23**, 55–65.
12. Watt, S., Mata, J., López-Maury, L., et al., (2008) *urg1*: a uracil-regulatable promoter system for fission yeast with short induction and repression times. *PLoS One* **3**, e1428.
13. Mata, J., Lyne, R., Burns, G., et al., (2002) The transcriptional program of meiosis and sporulation in fission yeast. *Nature Genet.* **32**, 143–147.
14. Mata, J., and Bähler, J. (2006) Global roles of Ste11p, cell type, and pheromone in the control of gene expression during early sexual differentiation in fission yeast. *Proc. Natl. Acad. Sci. USA* **103**, 15517–15522.
15. Chen, D., Toone, W.M., Mata, J., et al., (2003) Global transcriptional responses of fission yeast to environmental stress. *Mol. Biol. Cell* **14**, 214–229.
16. Rustici, G., van Bakel, H., Lackner, D.H., et al., (2007) Global transcriptional responses of fission and budding yeast to changes in copper and iron levels: a comparative study. *Genome Biol.* **8**, R73.
17. Tommasino, M., and Maundrell, K. (1991) Uptake of thiamine by *Schizosaccharomyces pombe* and its effect as a transcriptional regulator of thiamine-sensitive genes. *Curr. Genet.* **20**, 63–66.
18. Javerzat, J.P., Cranston, G., and Allshire, R.C. (1996) Fission yeast genes which disrupt mitotic chromosome segregation when overexpressed. *Nucleic Acids Res.* **24**, 4676–4683.
19. Tamm, T. (2009) Plasmids with E2 epitope tags: tagging modules for N- and C-terminal PCR-based gene targeting in both budding and fission yeast, and inducible expression vectors for fission yeast. *Yeast* **26**, 55–66.
20. Huang, Y., Hamada, M., Patel, J., et al., (2001) Construction of FLAG and histidine tagging vectors for *Schizosaccharomyces pombe*. *Yeast* **18**, 463–468.
21. Forsburg, S. L., and Sherman, D. A. (1997) General purpose tagging vectors for fission yeast. *Gene* **191,** 191–195.
22. Craven, R.A., Griffiths, D.J., Sheldrick, K.S., et al., (1998) Vectors for the expression of tagged proteins in *Schizosaccharomyces pombe*. *Gene* **221**, 59–68.
23. Forsburg Laboratory website (http://www-rcf.usc.edu/~forsburg/), unpublished.
24. Siam, R., Dolan, W. P., and Forsburg, L. (2004) Choosing and using *Schizosaccharomyces pombe* plasmids. *Methods* **33,** 189–198.
25. Hagan, I., and Yanagida, M. (1995) The product of the spindle formation gene *sad1*+ associates with the fission yeast spindle pole body and is essential for viability. *J. Cell Biol.* **129**, 1033–1047.

26. Okazaki, K., Okazaki, N., Kume, K., et al., (1990) High-frequency transformation method and library transducing vectors for cloning mammalian cDNAs by trans-complementation of *Schizosaccharomyces pombe*. *Nucleic Acids Res.* **18,** 6485–6489.
27. MacIver, F.H., Tanaka, K., Robertson, A.M., et al., (2003) Physical and functional interactions between polo kinase and the spindle pole component Cut12 regulate mitotic commitment in *S. pombe*. *Genes Dev.* **17,** 1507–1523.
28. Woods, A., Sherwin, T., Sasse, R., et al., (1989) Definition of individual components within the cytoskeleton of *Trypanosoma burucei* by a library of monoclonal antibodies. *J. Cell Sci.* **93,** 491–500.

Chapter 23

Heterologous Gene Expression by Chromosomal Integration in Fission Yeast

Akihisa Matsuyama and Minoru Yoshida

Abstract

Thanks to the convenience and flexibility of the multicopy plasmid-based approach for heterologous gene expression, this technique has long been used for biological studies, especially in prokaryotes and lower eukaryotes. For better understanding of biological mechanisms, however, there are increasing demands on the experimental technologies enabling fine-tuned expression of introduced heterologous genes or serving conditions that are closer to the physiological conditions. For this purpose, the use of direct tagging of a chromosomal gene has been gradually increasing, although the use conditions of this approach are relatively limited compared to plasmid-based methods. Expression of a cloned gene using chromosomal integration has a property intermediate between multicopy plasmid-based method and direct tagging of an endogenous gene. Here, we describe the principle and methods of introduction of a cloned gene into the targeting loci of the chromosome in fission yeast.

Key words: Plasmid, Integration, Homologous recombination, *Schizosaccharomyces pombe*

1. Introduction

Expression of a cloned gene using plasmids is a traditional method. Plasmids are very powerful and flexible, since they have many options by adding, deleting, replacing, or combining functional pieces. Therefore, they are still one of first choices in a variety of biological researches, even when direct tagging of an endogenous gene has been gradually increasing. However, despite their ease of use, they are sometimes inappropriate for some kinds of experiments. The expression levels of cloned genes vary from cells to cells mainly due to the numerical instability of plasmids in the cell. Therefore, a cloned gene is overexpressed in a subset of cells, even when a weak promoter is used. This feature is one of demerits,

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_23, © Springer Science+Business Media, LLC 2012

especially for genes that are sensitive to the expression levels. For example, some gene products are known to mislocalize when overexpressed. These proteins might localize to several compartments in the cell, making it very difficult to figure out which is the true localization. From this standpoint, expression of a cloned gene from a chromosomal locus has substantial merit since the copy number is restricted to only one in every cell, allowing uniform expression.

For chromosomal integration of a cloned gene in the fission yeast *Schizosaccharomyces pombe*, several vectors have been developed. The first generation vectors for such purpose were derived from commonly used plasmid vectors, pREP-series vectors, by deleting *ars1*, an autonomously replicating sequence (1). Because the resulting integration vectors cannot replicate by themselves in a cell, only transformants in which the introduced vector is integrated into the chromosomes can grow in the selective media. To enhance the integration efficiency, plasmids are linearized by digestion with restriction endonucleases before introduction into cells. However, since the integration into the chromosome occurs accidentally and cannot be regulated, it is at risk of integration within a gene or into an adjacent site which may impair or affect the function of an endogenous gene close to the landing site. Especially, since a significant proportion of yeast genome is occupied by functional genes unlike human, this feature is not always negligible.

To overcome the disadvantages of the early versions of integration vectors, a new set of vectors have been developed (2, 3). In the current version, the integration site can be controlled by adding parts of marker genes to the vectors. Here, we describe the structural features of these vectors and their use in heterologous expression of cloned genes.

2. Materials

1. Plasmid DNA preparation.
2. *E. coli*-competent cells: We usually use the DH5α strain, but other strains that are generally used are also possible. Many strains are commercially available.
3. LB (+Amp) solid medium: 1% Trypton (Becton, Dickinson and Company), 0.5% yeast extract, 1% NaCl, 0.2% glucose, 1.5% agar. Adjust pH to 7.2 with a 2 N NaOH solution. Premix powder is also commercially available (e.g., Becton, Dickinson and Company). When the temperature drops to around 65°C after autoclaving, add ampicillin at 50 μg/mL, mix immediately, and pour the liquid into Petri dishes. After gelation, store the plates at 4°C.

4. 2× YT (+Amp) liquid medium: 1% yeast extract, 1.6% peptone (Becton, Dickinson and Company), 0.5% NaCl. Adjust pH to 7.0 with a 2 N NaOH solution. Premix powder is also commercially available (e.g., Becton, Dickinson and Company). Sterilize by autoclaving. Ampicillin should be added at 50 μg/mL after autoclaving. Store at 4°C after addition of ampicillin.
5. Cell resuspension solution: 50 mM Tris–HCl, pH 7.5, 10 mM EDTA, and 100 μg/mL RNase A.
6. Cell lysis solution: 0.2 N NaOH and 1% SDS.
7. Neutralization solution: 7.5 M ammonium acetate (see Note 1).
8. 2-Propanol.
9. 70% (v/v) ethanol.
10. TE buffer: 10 mM Tris–HCl, pH 8.0, and 1 mM EDTA.
11. Agarose gel for DNA electrophoresis: Agarose is suspended in TAE buffer (40 mM Tris–acetate, pH 8.3, and 1 mM EDTA) and dissolved by heating in a microwave oven. When the temperature drops to around 65°C, pour it into trays and allow for gelation (see Note 2). Agarose is used generally at a concentration of 0.8–1.2% (w/v).
12. Ethidium bromide solution is supplied from many manufacturers (e.g., AMRESCO). A commonly used final concentration of ethidium bromide is 0.5 μg/mL. This chemical is a carcinogen, thereby requiring careful handling.
13. 10× H buffer: 500 mM Tris–HCl, pH 7.5, 100 mM $MgCl_2$, 10 mM dithiothreitol (DTT), 1 M NaCl. Buffers for restriction enzyme digestion are usually supplied together with enzymes from manufacturers.
14. 0.1% (v/v) Triton X-100 solution.
15. 0.1% (w/v) bovine serum albumin (BSA) solution: This is required only for certain restriction enzyme digestion, and is usually supplied together with enzymes from manufacturers.
16. Host strain: For integration using pDUAL-series vectors, strains having the *leu1-32* mutation must be used. For use of other vectors, strains having an intact targeted marker gene should be selected. For example, pDUAL2 can be used for only host strains having the intact *leu1*$^+$ gene. In another instance, pHIS3K can be used for strains having the intact *his3*$^+$ gene. Thus, a marker gene should be selected depending on the genotype of the host strain into which the gene is to be introduced.
17. YE liquid medium: 0.5% yeast extract, 2% glucose, 50 μg/mL adenine. To make 1 L of this medium, these materials are added to water, dissolved well, and then autoclaved at 120°C for 15 min. The concentration of glucose is often increased to 3%.

18. 0.1 M lithium acetate solution, pH 5.0: Adjust pH with HCl. Sterilize by autoclaving or filtration, and store at room temperature (RT).
19. 50% PEG solution: Prepare with water just before use and sterilize by filtration. The concentration can be reduced to at least 40%. It is sometimes hard to dissolve completely. Heating is effective for dissolving. Sterilization by autoclaving is also possible.
20. Dimethylsulfoxide (DMSO).
21. SD solid medium: 0.67% Yeast nitrogen base w/o amino acids (Becton, Dickinson and Company), 2% glucose, and 2% agar. When the temperature drops to around 65°C after autoclaving, pour the liquid medium into Petri dishes to allow gelation. Supplements, such as amino acids, should be added at the final concentration of 50 μg/mL if needed.
22. YE solid medium containing antibiotics: Prepare as is the case for YE liquid medium, except adding agar at the final concentration of 2% (w/v). When the temperature drops to around 65°C after autoclaving, add a drug solution at the final concentration of 100 μg/mL and pour the liquid into dishes to allow gelation.
23. Polymerase for amplification: MightyAmp DNA Polymerase (TaKaRa) is our recommendation. Other polymerases are also suitable; however, it is needed to select ones with extremely high potential for amplification. High fidelity is not required. Buffers and dNTPs are usually supplied together with polymerases from manufacturers.
24. Universal primers for PCR: For checking integration at the *leu1* locus, use leu1-R: 5′-GGTCATAAAGTTGAACGGAT GTCG-3′ and ADHterm-F: 5′-CTCTTATTGACCACACCT CTACC-3′. For the *lys1*, *his3*, and *arg1* loci, use 5′-GTGATG TGTCTGGGAAAGGCAGAG-3′,5′-AGCATCCAAAGCTAA ACGAGA GG-3′, and 5′-TCTTACTGGAGACAATGGTG GC-3′, respectively, instead of leu1-R. ADHterm-F can be used regardless of the loci.
25. YET solid medium: Prepare as is the case for YE liquid medium, except adding agar and thiamine (vitamin B1) at the final concentration of 2% (w/v) and 15 μM, respectively. Unless using the *nmt1* promoter to express heterologous genes, thiamine is not required.
26. MM solid medium: To make 1 L of this medium, the materials described below are added to water, mixed well, and then autoclaved (120°C, 15 min): 3 g $C_8H_5KO_4$, 2.2 g Na_2HPO_4, 5 g NH_4Cl, 20 g glucose, 20 mL 50× salt stock (260 mM $MgCl_2$, 5 mM $CaCl_2$, 670 mM KCl, and 14.1 mM Na_2SO_4), 1 mL of 1,000× vitamin stock (4.2 mM pantothenic acid,

81.2 mM nicotinic acid, 55.5 mM *myo*-inositol, and 40.8 mM biotin; filter sterilized and stored at 4°C), and 1 mL 10,000× mineral stock (80.9 mM H_3BO_3, 23.7 mM $MnSO_4$, 13.9 mM $ZnSO_4$, 7.4 mM $FeCl_3$, 2.47 mM $(NH_4)_6Mo_7O_{24}$, 6.02 mM KI, 1.6 mM $CuSO_4$, and 47.6 mM citric acid; store at 4°C after sterilization by filtration). Supplements, such as amino acids, should be added at the final concentration of 50 μg/mL if needed.

27. 0.7 N sodium hydroxide (NaOH) solution.
28. 1× SDS-PAGE sample buffer: 50 mM Tris–HCl, pH 6.8, 100 mM DTT, 2% SDS, 0.1% bromophenol blue, and 10% glycerol.

3. Methods

There are two main types of integration vectors: pDUAL type (Fig. 1) and pDUAL2 type (Fig. 2). Whereas pDUAL-series vectors are designed for only strains having the *leu1-32* mutation, pDUAL2-type vectors can be used in the wild-type background (2). Although all the vectors described herein (Tables 1 and 2) are designed for chromosomal integration, some of them can be maintained in the fission yeast cells as extrachromosomal multicopy plasmids since they have the *ars1-ura4⁺* fragment which enables autonomous replication of the plasmids. To use them as multicopy plasmids, *ura4⁺* is used as a positive selection marker (4). However, regardless of whether they possess the *ars1-ura4⁺* fragment, these plasmids

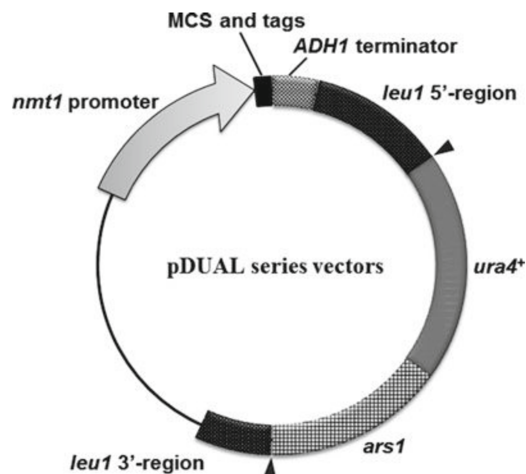


Fig. 1. Schematic representation of the construction of the pDUAL-series plasmids. *Arrowheads* indicate restriction sites (*Not*I, *Apa*I, and *Sac*II) required for the elimination of the *ars1-ura4⁺* fragment, which allows targeted integration. ORFs are inserted into the multiple cloning site (MCS). The schematic diagrams are not at scale.

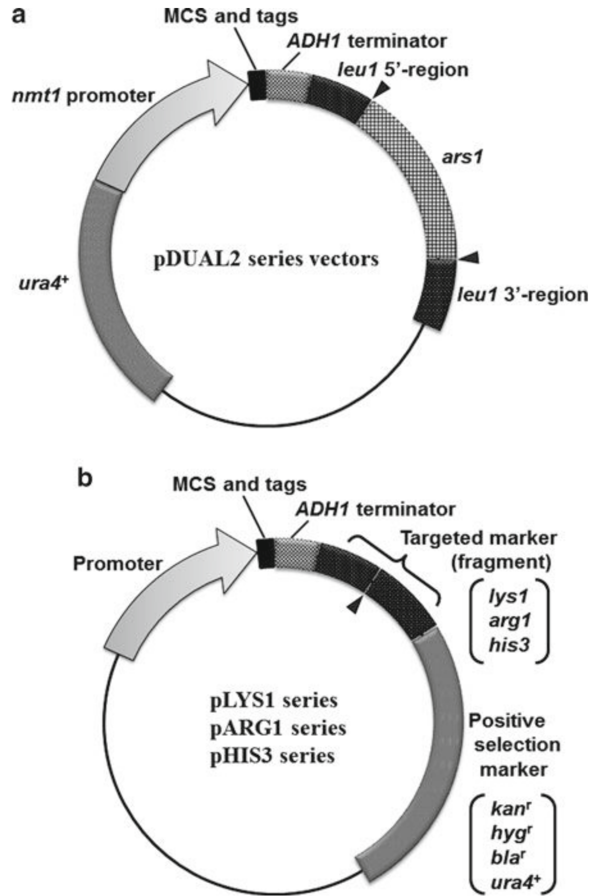


Fig. 2. The construction of the pDUAL2-series and related plasmids. (**a**) The construction of pDUAL2. *Arrowheads* indicate restriction sites (*Not*I, *Apa*I, and *Sac*II), which are required for the elimination of the *ars1* element to generate the *leu1*-targeting fragment. MCS includes unique restriction sites that allow insertion of an ORF. (**b**) The construction of vectors targeted to *lys1*, *arg1*, and *his3*. *Arrowheads* indicate a restriction site for *Not*I to split the marker fragment. The *ars1*-*ura4*+ fragment is inserted at this site in some vectors. For these vectors, restriction sites for *Apa*I and *Sac*II exist in addition to the *Not*I site. The schematic diagrams are not at scale.

can be easily converted to fragments for chromosomal integration by digestion of them with a certain restriction endonuclease before transformation of the yeast cells. In both situations, the resultant fragments, lacking the autonomously replicating sequence *ars1* (5), are designed for targeting into the particular chromosomal loci by homologous recombination (Figs. 3 and 4).

To select only transformants (the so-called positive selection), pDUAL-series vectors have a fragment of the *leu1* gene (6), comprising the promoter region and the open reading frame lacking only its 3′-terminal region (Fig. 1). This fragment itself is not functional due to lack of the 3′-terminal region. However, single-crossover homologous recombination at the chromosomal *leu1*

Table 1
pDUAL vector series. These are part of vectors developed. For other available vectors, *see* http://www.riken.jp/SPD/Plasmids.html

Plasmid name	Promoter	Tag	Tag position	Multicopy marker	Integration marker
pDUAL	–	–	–	*ura4*$^+$	*leu1*$^+$
pDUAL-Pef1a	*ef1a-c*	–	–	*ura4*$^+$	*leu1*$^+$
pDUAL-Ptub1	*tub1*	–	–	*ura4*$^+$	*leu1*$^+$
pDUAL-HFF1	*nmt1*	His_6-$FLAG_2$	5′	*ura4*$^+$	*leu1*$^+$
pDUAL-HFF21	*tub1*	His_6-$FLAG_2$	5′	*ura4*$^+$	*leu1*$^+$
pDUAL-HFF31	*cam1*	His_6-$FLAG_2$	5′	*ura4*$^+$	*leu1*$^+$
pDUAL-HFF41	*nmt1*[a]	His_6-$FLAG_2$	5′	*ura4*$^+$	*leu1*$^+$
pDUAL-HFF51	*tif51*	His_6-$FLAG_2$	5′	*ura4*$^+$	*leu1*$^+$
pDUAL-HFF61	*ef1a-c*	His_6-$FLAG_2$	5′	*ura4*$^+$	*leu1*$^+$
pDUAL-HFF81	*nmt1*[b]	His_6-$FLAG_2$	5′	*ura4*$^+$	*leu1*$^+$
pDUAL-HFG1	*nmt1*	His_6-FLAG-GFP(S65C)	5′	*ura4*$^+$	*leu1*$^+$
pDUAL-HFG21	*tub1*	His_6-FLAG-GFP(S65C)	5′	*ura4*$^+$	*leu1*$^+$
pDUAL-HFG31	*cam1*	His_6-FLAG-GFP(S65C)	5′	*ura4*$^+$	*leu1*$^+$
pDUAL-HFG41	*nmt1*[a]	His_6-FLAG-GFP(S65C)	5′	*ura4*$^+$	*leu1*$^+$
pDUAL-HFG51	*tif51*	His_6-FLAG-GFP(S65C)	5′	*ura4*$^+$	*leu1*$^+$
pDUAL-HFG61	*ef1a-c*	His_6-FLAG-GFP(S65C)	5′	*ura4*$^+$	*leu1*$^+$
pDUAL-HFG81	*nmt1*[b]	His_6-FLAG-GFP(S65C)	5′	*ura4*$^+$	*leu1*$^+$
pDUAL-FFH1	*nmt1*	$FLAG_2$-His_6	3′	*ura4*$^+$	*leu1*$^+$
pDUAL-FFH21	*tub1*	$FLAG_2$-His_6	3′	*ura4*$^+$	*leu1*$^+$
pDUAL-FFH31	*cam1*	$FLAG_2$-His_6	3′	*ura4*$^+$	*leu1*$^+$
pDUAL-FFH41	*nmt1*[a]	$FLAG_2$-His_6	3′	*ura4*$^+$	*leu1*$^+$
pDUAL-FFH51	*tif51*	$FLAG_2$-His_6	3′	*ura4*$^+$	*leu1*$^+$
pDUAL-FFH61	*ef1a-c*	$FLAG_2$-His_6	3′	*ura4*$^+$	*leu1*$^+$
pDUAL-FFH81	*nmt1*[b]	$FLAG_2$-His_6	3′	*ura4*$^+$	*leu1*$^+$
pDUAL-GFH1	*nmt1*	GFP(S65C)-FLAG-His_6	3′	*ura4*$^+$	*leu1*$^+$

(continued)

Table 1 (continued)

Plasmid name	Promoter	Tag	Tag position	Multicopy marker	Integration marker
pDUAL-GFH21	*tub1*	GFP(S65C)-FLAG-His_6	3′	*ura4*$^+$	*leu1*$^+$
pDUAL-GFH31	*cam1*	GFP(S65C)-FLAG-His_6	3′	*ura4*$^+$	*leu1*$^+$
pDUAL-GFH41	*nmt1*[a]	GFP(S65C)-FLAG-His_6	3′	*ura4*$^+$	*leu1*$^+$
pDUAL-GFH51	*tif51*	GFP(S65C)-FLAG-His_6	3′	*ura4*$^+$	*leu1*$^+$
pDUAL-GFH61	*ef1a-c*	GFP(S65C)-FLAG-His_6	3′	*ura4*$^+$	*leu1*$^+$
pDUAL-GFH81	*nmt1*[b]	GFP(S65C)-FLAG-His_6	3′	*ura4*$^+$	*leu1*$^+$
pDUAL-YFH1	*nmt1*	YFP-FLAG-His_6	3′	*ura4*$^+$	*leu1*$^+$

[a]Medium-strength *nmt1* promoter
[b]Low-strength *nmt1* promoter

locus containing the *leu1-32* mutation (a single base substitution of G137 with A, resulting in an amino acid substitution from Gly46 to Glu of the Leu1 protein) (2) generates a functional *leu1*$^+$ gene (Fig. 3). Therefore, the correct integrants are expected to become prototrophic for leucine. To induce single-crossover homologous recombination, a 2-for-1 split on the targeting *leu1* fragment, which results in the linearization of the plasmid, should be needed. For linearization of pDUAL-series vectors containing the *ars1-ura4*$^+$ cassette, the restriction enzyme recognition sites for *Not*I, *Apa*I, and *Sac*II exist at both ends of the *ars1-ura4*$^+$ cassette. Digestion with one of three restriction enzymes releases the *ars1-ura4*$^+$ fragment from these plasmids, resulting in the generation of the same targeting fragment that is produced by digestion of pDUAL.

To use pDUAL-series vectors, the host strain should have the *leu1-32* mutation. Although the *leu1-32* mutation is very popular in researches using fission yeast, strains do not always have this mutation. Therefore, to utilize the wild-type strain as the host, other series of vectors that enable both targeted chromosomal integration and selection of proper transformants have been developed. pDUAL2-series vectors have been initially designed for this purpose (Fig. 2a), and then additional three vectors named pLYS1-, pARG1-, and pHIS3 series have been developed using the similar strategy (Fig. 2b). Four chromosomal loci containing genes *leu1*$^+$, *lys1*$^+$, *arg1*$^+$, and *his3*$^+$, each encoding a component of amino acid

Table 2
pDUAL2-series and other pDUAL2-type vectors. These are part of vectors developed. Other vectors are listed in http://www.riken.jp/SPD/Plasmids.html

Plasmid name	Promoter	Tag	Tag position	Multicopy marker	Targeted locus	Integration marker
pDUAL2	–	–	–	$ura4^+$	$leu1^+$	$ura4^+$
pDUAL2-HFF1	*nmt1*	His_6-$FLAG_2$	5′	$ura4^+$	$leu1^+$	$ura4^+$
pDUAL2-HFF21	*tub1*	His_6-$FLAG_2$	5′	$ura4^+$	$leu1^+$	$ura4^+$
pDUAL2-HFF41	*nmt1*[a]	His_6-$FLAG_2$	5′	$ura4^+$	$leu1^+$	$ura4^+$
pDUAL2-HFF61	*ef1a-c*	His_6-$FLAG_2$	5′	$ura4^+$	$leu1^+$	$ura4^+$
pDUAL2-HFF81	*nmt1*[b]	His_6-$FLAG_2$	5′	$ura4^+$	$leu1^+$	$ura4^+$
pDUAL2-HFG1	*nmt1*	His_6-FLAG-GFP(S65C)	5′	$ura4^+$	$leu1^+$	$ura4^+$
pDUAL2-HFG41	*nmt1*[a]	His_6-FLAG-GFP(S65C)	5′	$ura4^+$	$leu1^+$	$ura4^+$
pDUAL2-HFG81	*nmt1*[b]	His_6-FLAG-GFP(S65C)	5′	$ura4^+$	$leu1^+$	$ura4^+$
pDUAL2-FFH1	*nmt1*	FLAG2-His6	3′	$ura4^+$	$leu1^+$	$ura4^+$
pDUAL2-FFH21	*tub1*	FLAG2-His6	3′	$ura4^+$	$leu1^+$	$ura4^+$
pDUAL2-FFH41	*nmt1*[a]	FLAG2-His6	3′	$ura4^+$	$leu1^+$	$ura4^+$
pDUAL2-FFH61	*ef1a-c*	FLAG2-His6	3′	$ura4^+$	$leu1^+$	$ura4^+$
pDUAL2-FFH81	*nmt1*[b]	FLAG2-His6	3′	$ura4^+$	$leu1^+$	$ura4^+$
pDUAL2-GFH1	*nmt1*	GFP(S65C)-FLAG-His_6	3′	$ura4^+$	$leu1^+$	$ura4^+$
pDUAL2-GFH41	*nmt1*[a]	GFP(S65C)-FLAG-His_6	3′	$ura4^+$	$leu1^+$	$ura4^+$
pDUAL2-GFH81	*nmt1*[b]	GFP(S65C)-FLAG-His_6	3′	$ura4^+$	$leu1^+$	$ura4^+$
pLYS1U	–	–	–		$lys1^+$	$ura4^+$

(continued)

Table 2 (continued)

Plasmid name	Promoter	Tag	Tag position	Multicopy marker	Targeted locus	Integration marker
pLYS1U-Pnmt1	*nmt1*	–	–		$lys1^+$	$ura4^+$
pLYS1U-Ptub1	*tub1*	–	–		$lys1^+$	$ura4^+$
pLYS1K	–	–	–		$lys1^+$	kan^r
pLYS1H	–	–	–		$lys1^+$	hyg^r
pLYS1B	–	–	–		$lys1^+$	bla^r
pHIS3U	–	–	–		$his3^+$	$ura4^+$
pHIS3U-Pnmt1	*nmt1*	–	–		$his3^+$	$ura4^+$
pHIS3K	–	–	–		$his3^+$	kan^r
pHIS3K-Ptub1	*tub1*	–	–		$his3^+$	kan^r
pHIS3H	–	–	–		$his3^+$	hyg^r
pHIS3B	–	–	–		$his3^+$	bla^r
pARG1U	–	–	–		$arg1^+$	$ura4^+$
pARG1K	–	–	–		$arg1^+$	kan^r
pARG1H	–	–	–		$arg1^+$	hyg^r
pARG1B	–	–	–		$arg1^+$	bla^r
pLYS1UF	–	–	–	$ura4^+$	$lys1^+$	$ura4^+$
pLYS1KF	–	–	–	$ura4^+$	$lys1^+$	kan^r
pLYS1HF	–	–	–	$ura4^+$	$lys1^+$	hyg^r

(continued)

pLYS1BF	–	–	–	*ura4*$^+$	*lys1*$^+$	*bla*^{r}
pHIS3UF	–	–	–	*ura4*$^+$	*his3*$^+$	*ura4*$^+$
pHIS3KF	–	–	–	*ura4*$^+$	*his3*$^+$	*kan*^{r}
pHIS3HF	–	–	–	*ura4*$^+$	*his3*$^+$	*hyg*^{r}
pHIS3BF	–	–	–	*ura4*$^+$	*his3*$^+$	*bla*^{r}
pARG1UF	–	–	–	*ura4*$^+$	*arg1*$^+$	*ura4*$^+$
pARG1KF	–	–	–	*ura4*$^+$	*arg1*$^+$	*kan*^{r}
pARG1HF	–	–	–	*ura4*$^+$	*arg1*$^+$	*hyg*^{r}
pARG1BF	–	–	–	*ura4*$^+$	*arg1*$^+$	*bla*^{r}

kan^r G418-resistant gene, *hyg*^{r} hygromycin-resistant gene, *bla*^{r} blasticidin-resistant gene
[a]Medium-strength *nmt1* promoter
[b]Low-strength *nmt1* promoter

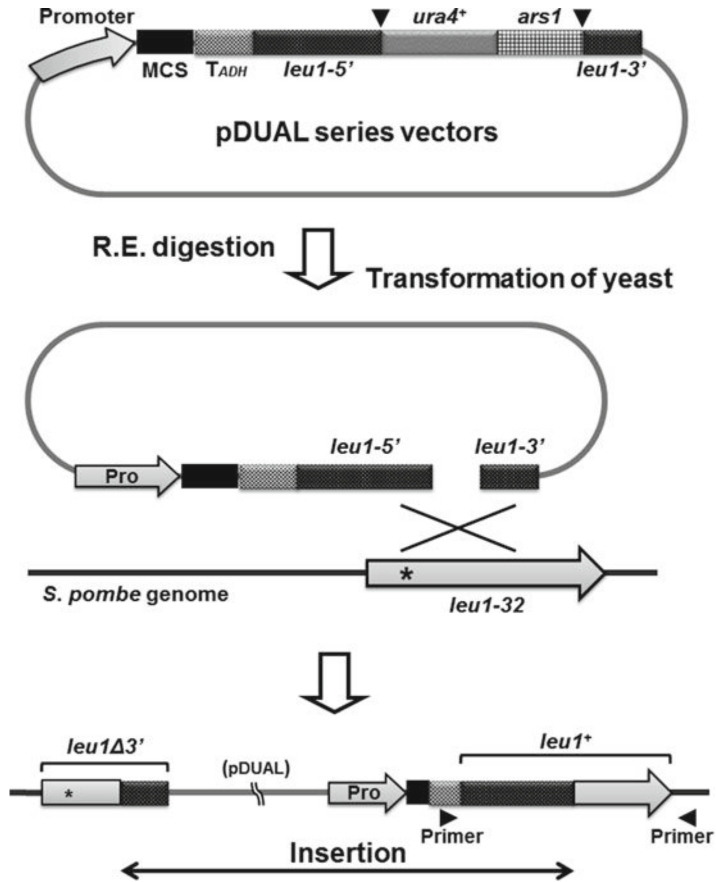


Fig. 3. Strategy for chromosomal integration of the pDUAL-derived fragment. Restriction enzymes (R.E.) cleave both ends of the *ars1-ura4*+ fragment, which releases this fragment and makes the plasmid DNA linear. The resultant linear fragment is targeted for the endogenous *leu1* locus. After homologous recombination with the chromosomal *leu1-32* allele, the functional *leu1* gene is generated in addition to the incomplete *leu1* fragment. Therefore, correct integrants are expected to show leucine prototrophy. The regions that match to the primers sequences for colony PCR are indicated by *inverted arrowheads*. An *asterisk* indicates the position of the *leu1-32* mutation. The schematic diagrams are not at scale.

synthesis, are targeted by these vectors. In contrast to the *leu1* fragment in the pDUAL-series vectors, each targeting marker fragment in these vectors is incomplete at both 5′- and 3′-ends (Fig. 4). Therefore, when these fragments experience homologous recombination in cells, each endogenous target gene is expected to be split in two, both of which lose their function (Fig. 4). For linearization of these vectors that do not have the *ars1-ura4*+ cassette, the restriction enzyme recognition site for *Not*I exists at the middle of the targeting marker fragments. The integrants would become auxotrophic for leucine, histidine, lysine, or arginine which allows selection of proper integrants. However, since this strategy itself is negative selection that cannot choose only transformants,

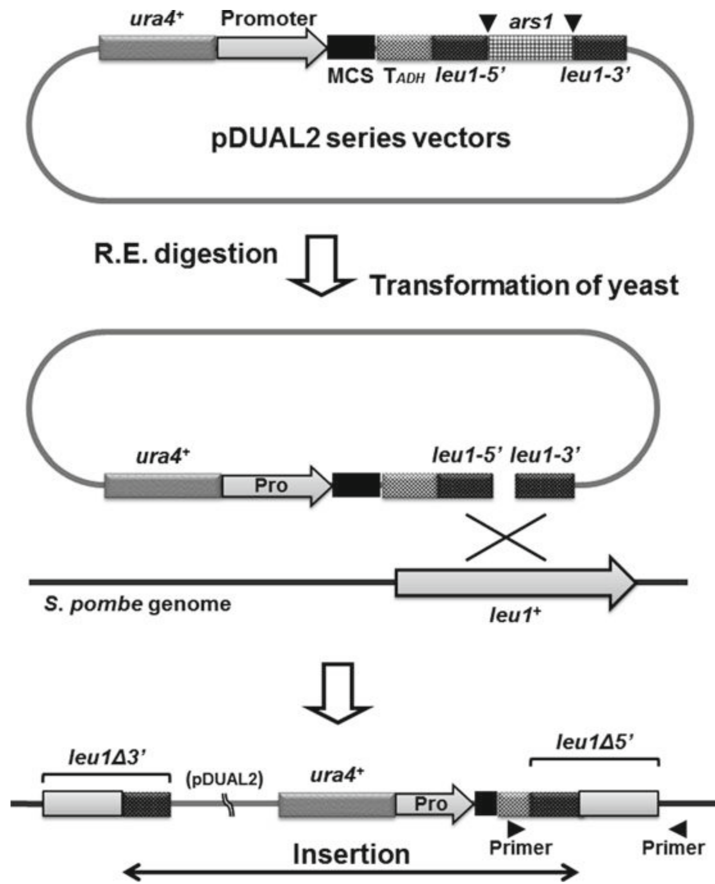


Fig. 4. Strategy for chromosomal integration using the pDUAL2-type plasmids. Plasmids can be linearized at the middle of the integration marker fragments by restriction enzyme (R.E.) digestion. For plasmids containing the *ars1* sequence, this fragment is inserted at the middle of the integration marker and can be cut off by restriction enzyme digestion. When introduced in a yeast cell, the linearized fragment is expected to be integrated at the chromosomal marker gene locus by single-crossover homologous recombination. If each linearized fragment is successfully integrated, the endogenous marker gene is split into two fragments, both of which lose their function. Therefore, transformants are expected to become auxotrophic for a certain amino acid, whereas they show drug resistance or nucleotide prototrophy according to the markers used. Correct integration can be checked by PCR using primers indicated by *inverted arrowheads*. The diagrams are not at scale. The principle of chromosomal integration using other vectors, such as pLYS1U, is essentially the same.

another strategy to select only transformants (positive selection) should be required at first. For this purpose, kanMX6 (G418) (7), hphMX4 (hygromycin B) (8), bsd (blasticidin S) (9) cassettes, or the fission yeast *ura4*+ gene are installed in another site of the vector (Fig. 3). Therefore, only transformants can be selected initially on medium containing an appropriate drug or lacking uracil before checking for their amino acid requirement. For linearization of the plasmids, the restriction enzyme recognition sites are designed to split the targeting fragments as is the case of pDUAL-series vectors.

3.1. Plasmid DNA Preparation

There are many protocols for alkali method to isolate plasmid DNA from *E. coli*. For the protocol specialized for handling a large number of samples using a 96-well plate, see ref. 10.

1. Culture *E. coli* transformants in 0.8–1.5 mL of 2× YT (+ Amp) liquid medium with a test tube at 37°C overnight. Shake or rotate the tube vigorously.
2. Transfer cell culture to a 1.5-mL microtube and centrifuge at 8,400 × *g* for 1 min. Discard medium by aspiration.
3. Suspend cells with 100 μL of resuspension solution and vortex vigorously until cells are well-suspended.
4. Add 100 μL of cell lysis solution and mix mildly until the solution becomes clear.
5. Add 100 μL of neutralization solution and mix vigorously (see Note 4).
6. Centrifuge at 18,800 × *g* for 10 min.
7. Transfer the supernatants to a new 1.5-mL microtube.
8. Add 300 μL of 2-propanol. Mix well.
9. Centrifuge at 18,800 × *g* for 5 min, and then discard the supernatant by aspiration.
10. Rinse the tube with 120 μL of 70% ethanol. Centrifuge at 18,800 × *g* for 2 min and aspirate the supernatant.
11. Dry up the tube.
12. Dissolve DNA with 50 μL of TE buffer or deionized water.
13. Digest the plasmid DNA with an appropriate restriction enzyme. Use 1 μL of the DNA solution for a 10-μL-scale reaction.
14. Check the DNA by agarose gel electrophoresis. Load 2–10 μL of the reaction mixture after addition of 1 μL of loading buffer.

3.2. Linearization of Plasmid DNAs by Restriction Enzyme Digestion

1. Assemble the following recipe: 4 μL plasmid DNA (usually, approximately 0.1 μg/μL), 1.2 μL 10× H buffer, 1.2 μL 0.1% Triton X-100 solution, 1.2 μL 0.1% BSA, 3.9 μL deionized water, 0.5 μL restriction endonuclease (usually, *Not*I). Incubate the tube at the appropriate temperature (usually, at 37°C) for at least 30 min.
2. Add 1 μL 10× loading buffer to stop the reaction. This is optional (see Note 5).
3. Check the DNA by agarose gel electrophoresis. This is also optional. Usually, it is not needed.

3.3. Transformation of *S. pombe* Cells

1. Culture the host strain overnight in appropriate liquid medium, such as YE, at 30°C. The optimal growth temperature depends on the host strain.

2. Either exponentially growing or just-saturated cells cultured are harvested by centrifugation (2,100×*g*, 10 s, at room temperature) and resuspended in 50 μL of 0.1 M lithium acetate solution, pH 5.0.
3. Sample DNA (see Note 6), together with salmon sperm DNA and 150 μL of 50% (v/v) polyethylene glycol (PEG 4000), is added immediately to the cell suspension, and the mix was incubated at RT for at least 30 min.
4. After addition of 20 μL of DMSO, cells are heated at 42°C for 15 min.
5. Spin down the cells (2,100×*g*, 10 s, at RT), discard the supernatant, and resuspend the cells in 50 μL of sterilized water.
6. Plate the cells on appropriate selective medium (e.g., SD supplemented with lysine and lacking uracil when using pLYS1U) (see Note 7).
7. Check the amino acid or nucleotide requirement of the transformants by streaking them on appropriate medium (such as SD lacking arginine) (see Note 8).

3.4. Colony PCR for Confirmation of Chromosomal Integration

In any cases of chromosomal integration described here, correct integration of the targeting fragment into the desired locus can be analyzed by polymerase chain reaction (PCR) in addition to check by auxotrophy of the transformant. However, since the targeting efficiency is extremely high, this check is almost always omissible.

1. Culture several candidate transformants on an appropriate solid medium (generally, YE can be used).
2. For eight samples, assemble the following recipe: 80 μL 2× buffer, 16 μL dNTPs (2.5 mM each of dGTP, dATP, dTTP, and dCTP), 8 μL 10 μM primers (see Note 9), 55 μL deionized water, and 1 μL MightyAmp DNA polymerase (2.5 U/μL).
3. Dispense 20 μL of the premix solution to each well of an 8-well tube.
4. Suspend cells in each well of an 8-well tube by taking a scraping from cells on solid medium using toothpicks. Then, cap the tube (spin down the samples quickly if needed).
5. The conditions for the first PCR are as follows: 95°C for 3 min, 35 cycles of 95°C for 15 s, 52°C for 15 s, and 72°C for XX s, and 72°C for 3 min, where XX is determined according to the length of an expected product (~45 s/kbp).
6. Check 2 μL of the resultant solution by agarose gel electrophoresis (see Note 10).

3.5. Preparation of Protein Samples for Checking the Expression of an Introduced Gene

If an appropriate tag (such as HA and FLAG) is fused to the cloned gene, expression of its product could be detected by Western blotting following SDS-PAGE (see Note 11). Since protocols for SDS-PAGE and Western blotting are described in many literatures, we provide a quick protocol to prepare whole-cell lysate samples for SDS-PAGE.

1. Streak the integrants on YET or SD medium, and incubate them overnight at an appropriate temperature (generally, 30°C).
2. Restreak a part of the precultured colonies on MM medium, and incubate them for an appropriate time (see Note 12).
3. Harvest cells into a microtube prefilled with 70 μL of MM liquid medium using toothpicks.
4. Add 70 μL of a 0.7 N NaOH solution, mix well, and incubate for 3 min at RT.
5. Centrifuge at 2,100 × *g* for 1 min, and then discard the supernatant by aspiration.
6. Suspend the pellets in 50 μL of 1× SDS-PAGE sample buffer.
7. Incubate the samples at 100°C for 5 min.
8. Load 8–10 μL of the sample on a polyacrylamide gel.

4. Notes

1. Since ammonium acetate is highly hygroscopic, add 90 mL of deionized water per 100 g of ammonium acetate directly to the reagent bottle.
2. For storage of agarose gel, remove combs and trays, submerge the gel in 1× TAE buffer, and store at 4°C. Turn the gel upside down to prevent sedimentation of crystallized salt in the well, which affects negatively on electrophoresis.
3. It is also possible to use 5 M sodium acetate, pH 4.8, in place of ammonium acetate in this step. However, the debris resulted from the addition of ammonium acetate would become highly compact so that the transfer of the supernatant becomes easier.
4. The 10× loading buffer sometimes contains SDS at a low concentration. However, it does not interfere with transformation of yeast cells. Rather, it can reduce the risk of contamination of bacteria or fungi into the DNA solution. We, therefore, recommend adding this reagent.
5. We usually use 4 μL of the solution of digested (linearized) DNA prepared as described in Subheading 3.2 for each transformation. However, this amount can be changed according to the purpose of each experiment.

6. For colony selection on antibiotic-containing plates, cells should be plated initially on nonselective medium (YE) to allow expression of the drug-resistance gene. One or two days later, when a lawn of cells becomes visible, cells are replica plated using sterile velvet or filter paper onto the drug-containing medium. In case of selection by amino acid requirement, cells can be plated directly onto selective medium.
7. This process is not required for integration using pDUAL-series plasmids.
8. Primers that can be used are not limited to those described in Subheading 2, item 4. A pair of primers, one of which anneals to the sequence within the plasmid and the other of which anneals to the region just outside the inserting site and has no identical sequence in the plasmid, can be used. However, the primers described in Subheading 2, item 4, can be used as universal primers for confirmation regardless of sequences to be cloned, since the region used for PCR check is completely independent of the multiple cloning site (MCS).
9. A positive band can be obtained only when the introduced fragment is integrated into the proper locus. When no band is seen in the colony PCR, it is often hard to know whether the correct integrants are absent or it is simply the failure in PCR, since the amplification efficiency is not so high. However, since the integration efficiency of the plasmids described herein is very high, we recommend performing PCR using genomic DNA purified from candidates as template. For purification of genomic DNA from fission yeast, see ref. 10.
10. When using a series of vectors for chromosomal integration, it is to be noted that not all the transformants are proper integrants. According to the previous report (11), this phenomenon occurs due to gene conversion. Although it was reported that gene conversion occurs in nearly 30% of the transformants, we estimate that this frequency is less than 10% in the pDUAL systems. However, we recommend checking expression of the tagged proteins before use in the experiment. It is enough to check two transformants in most cases.
11. We usually streak cells on solid medium as three lines in a 2×2-cm area. The number of cells obtained from this area after a 20–22-h incubation is about 1×10^8, which corresponds to the number of exponentially growing cells (late log phase) collected from a 10 mL liquid culture.
12. In fission yeast, the powerful thiamine-repressible *nmt1* promoter has been the most widely used and modified. Two versions of the *nmt1* promoter have been additively employed (denoted as *nmt1** and *nmt1***, which show medium and low transcriptional activities, respectively, both maintaining repressibility

by thiamine) (12). To use the *nmt1* promoter and its derivatives, transcriptional induction of a downstream cloned gene occurs about 14 h after cells are transferred to medium lacking thiamine. An appropriate culture time should be determined considering this property. Other promoters can be selected depending on the introduced gene and the purpose of the experiment.

Acknowledgments

This work was supported by CREST Research Project, Japan Science and Technology Agency.

References

1. Maundrell K. (1993) Thiamine-repressible expression vectors pREP and pRIP for fission yeast. *Gene* **123**, 127–130.
2. Matsuyama A, Shirai A, Yashiroda Y, et al. (2004) pDUAL, a multipurpose, multicopy vector capable of chromosomal integration in fission yeast. *Yeast* **21**, 1289–1305.
3. Matsuyama A, Shirai A, Yoshida M. (2008) A novel series of vectors for chromosomal integration in fission yeast. *Biochem. Biophys. Res. Commun.* **374**, 315–319.
4. Bach ML. (1987) Cloning and expression of the OMP decarboxylase gene *URA4* from *Schizosaccharomyces pombe*. *Curr. Genet.* **12**, 527–534.
5. Maundrell K, Hutchison A, Shall S. (1988) Sequence analysis of ARS elements in fission yeast. *EMBO J.* 7, 2203–2209.
6. Kikuchi Y, Kitazawa Y, Shimatake H, Yamamoto M. (1988) The primary structure of the *leu1*+ gene of *Schizosaccharomyces pombe*. *Curr. Genet.* **14**, 375–379.
7. Bähler J, Wu JQ, Longtine MS, et al. (1998) Heterologous modules for efficient and versatile PCR-based gene targeting in *Schizosaccharomyces pombe*. *Yeast* **14**, 943–951.
8. Goldstein AL, McCusker JH. (1999) Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. *Yeast* **15**, 1541–1553.
9. Kimura M, Kamakura T, Tao QZ, et al. (1994) Cloning of the blasticidin S deaminase gene (BSD) from Aspergillus terreus and its use as a selectable marker for *Schizosaccharomyces pombe* and *Pyricularia oryzae*. *Mol. Gen. Genet.* **242**, 121–129.
10. Matsuyama A, Yoshida M. (2009) Systematic cloning of an ORFeome using the Gateway system. *Methods Mol. Biol.* **577**, 11–24.
11. Keeney JB, Boeke JD. (1994) Efficient targeted integration at *leu1-32* and *ura4-294* in *Schizosaccharomyces pombe*. *Genetics* **136**, 849–856.
12. Basi G, Schmid E, Maundrell K. (1993) TATA box mutations in the *Schizosaccharomyces pombe* *nmt1* promoter affect transcription efficiency but not the transcription start point or thiamine repressibility. *Gene* **123**, 131–136.

Chapter 24

Genetic Engineering of Industrial Strains of *Saccharomyces cerevisiae*

Sylvie Le Borgne

Abstract

Genetic engineering has been successfully applied to *Saccharomyces cerevisiae* laboratory strains for different purposes: extension of substrate range, improvement of productivity and yield, elimination of by-products, improvement of process performance and cellular properties, and extension of product range. The potential of genetically engineered yeasts for the massive production of biofuels as bioethanol and other nonfuel products from renewable resources as lignocellulosic biomass hydrolysates has been recognized. For such applications, robust industrial strains of *S. cerevisiae* have to be used. Here, some relevant genetic and genomic characteristics of industrial strains are discussed in relation to the problematic of the genetic engineering of such strains. General molecular tools applicable to the manipulation of *S. cerevisiae* industrial strains are presented and examples of genetically engineered industrial strains developed for the production of bioethanol from lignocellulosic biomass are given.

Key words: *Saccharomyces cerevisiae*, Industrial strain, Adaptation, Ethanol production, Pentose fermentation, Chromosomal integration, Lignocellulosic biomass

1. Introduction

The yeast *Saccharomyces cerevisiae* has long been used in the food industry in baking, brewing, and winemaking as well as in industrial fermentations for the production of bioethanol from sugar and starch feedstocks (1, 2). Branduardi et al. (3) have reviewed the potential of metabolically engineered yeasts for industrial applications, including the production of biofuels and nonfuel petroleum-derived products from renewable resources as an alternative to the use of oil (Fig. 1). Nonfuel products include glycerol, pyruvate, and organic acids that are substances directly utilized in the pharmaceutical or chemical industry or used as building blocks

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_24, © Springer Science+Business Media, LLC 2012

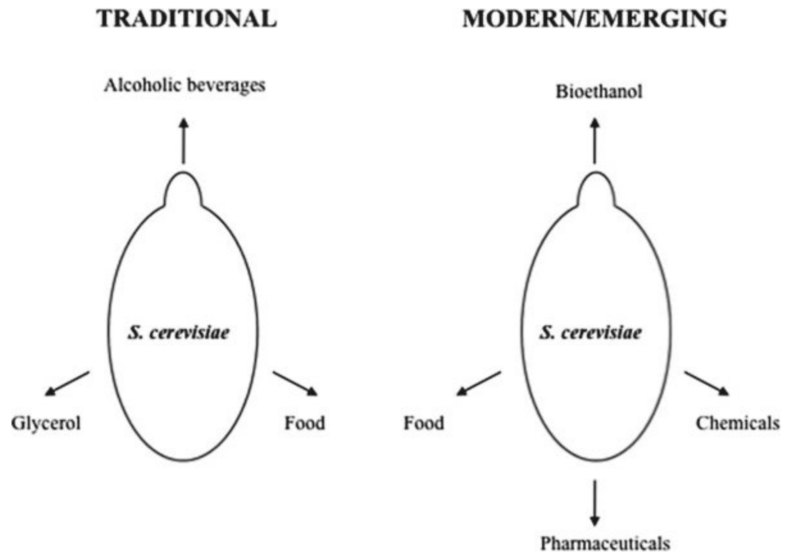


Fig. 1. Traditional, modern, and emerging industrial applications of the yeast *S. cerevisiae*. Adapted from ref. 1.

for further chemical or enzymatic syntheses (4, 5). An example of biofuel is the production of bioethanol from sugars derived from biomass hydrolysates (6).

S. cerevisiae is highly tolerant to low pH values and high sugar and ethanol concentrations lowering the risk of contamination in industrial fermentations; it is also able to grow anaerobically and is quite resistant to high osmotic pressure, oxidative stress, and inhibitors present in biomass hydrolysates (7, 8). There is a vast knowledge concerning its genetics, biochemistry, physiology, and fermentation technologies. This yeast has been safely used since ancient times in winemaking, brewing, and baking. It is classified as generally regarded as safe (GRAS) by the US Food and Drug Administration (FDA) and does not produce toxins or oncogenic and viral DNA. These features confer to *S. cerevisiae* a general robustness and potential applicability for industrial applications.

In contrast to classical methods of genetic strain improvement, such as selection, mutagenesis, and mating, genetic engineering consists in the targeted manipulation of the cell's genetic information to improve the production of a metabolite naturally formed by an organism in a low concentration or provide the organism with the ability to utilize atypical substrates allowing to extend the spectrum of usable industrial media, in particular renewable substrates as lignocellulosic biomass for example, or form metabolites not naturally produced. The genetic engineering of *S. cerevisiae* has mainly been performed with well-known laboratory strains. For industrial applications, the use of *S. cerevisiae* strains with an industrial background is necessary since industrial strains have adapted

to stress conditions found in industrial processes. However, several interesting technical challenges arise concerning the genetic engineering of industrial strains in order to obtain yeast strains useful for commercial applications. This chapter addresses some general issues and problematic about the genetic engineering of industrial strains of *S. cerevisiae* for the production of bulk chemicals as biofuel and nonfuel products.

2. Importance of Using Industrial Strains of *S. cerevisiae*

Hahn-Hägerdal et al. (6) have pointed out the importance of using industrial yeast strains for the production of bioethanol from biomass residues. In general, practically all industrial biotechnological processes expose cells to simultaneous or sequential combinations of stressful conditions. So, desirable phenotype for industrial organisms is resistance to multiple stresses or, more generally, environmental robustness in addition to adequate yield and productivity. Industrial *S. cerevisiae* strains are highly specialized organisms, which have adapted to different stress conditions as those found in industrial fermentations. This process known as "domestication" involves genetic, molecular, and physiological mechanisms of adaptation and has been extensively reviewed for wine and brewery yeasts (9, 10). The ability of such strains to adapt to varying environmental conditions and physiological challenges governs their usefulness and applicability. In general, it has been shown that commercial yeasts are more tolerant to stress conditions than laboratory strains. Concerning distillers' yeasts used for the production of fuel ethanol, Silva-Filho et al. (11) have demonstrated, by PCR fingerprinting of yeast samples from several Brazilian distilleries, that indigenous strains are more adapted to the industrial process than commercial yeasts. In fact, these indigenous yeasts tend to become dominant in the fermentative process even if commercial baker's yeasts were used as starter cultures in a proportion of 500 g of indigenous yeast in 1 ton of commercial yeast (12).

During industrial biomass production, conditioning, propagation, and fermentation, yeast cells are subject to a number of stresses as illustrated in Fig. 2.

Minimizing fresh water usage and recirculation of process water streams leads to more concentrated solutions with increased ionic strength. The high sugar content in fermentation media also contributes to hyperosmotic stress. Yeast cell recycling and washing in acid solutions also impose a selective pressure on cells. It has been shown that strains isolated from an industrial winery showed higher values for characters typically subjected to selective pressure, such as maximum capability to produce ethanol, fermentation rate, and SO_2 resistance (13). So, continual selective

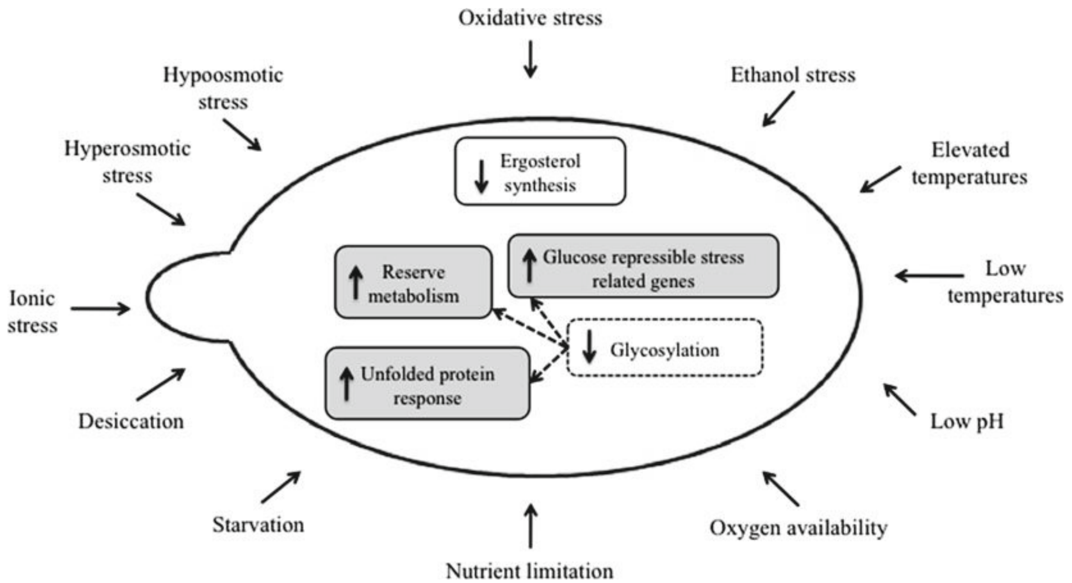


Fig. 2. Environmental stress and transcriptional response of yeast during industrial bioethanol fermentation. Adapted from ref. 14. The *black arrows* indicate upregulation, i.e., transcriptional activation (*up arrow*) or downregulation, i.e., transcriptional repression (*down arrow*).

pressure imposed on cells in industrial environments has led to the adaptation and selection of strains with higher stress tolerance and improved fermentative behavior (12).

The molecular and physiological response of an organism to changes in the environment is referred to as "stress response." The two major stress response pathways in *S. cerevisiae* are the heat-shock response and the general stress response activated by a number of environmental stresses, including oxidative, pH, heat, osmotic stresses, and nitrogen starvation (10). The general stress response is believed to be an evolutionary adaptation allowing yeast to respond to adverse environmental conditions in a nonspecific manner. Among the genes activated by stress in *S. cerevisiae* are the *HSP* genes, encoding heat-shock proteins most of which act as chaperones (9). The molecular basis of the technological properties of industrial yeast strains is still largely unknown. One possibility is that the adaptation of these strains is dependent on specific expression profiles of their genomes. The comparison of gene expression in industrial and nonindustrial strains could lead to the identification of genes involved in the adaptation to industrial environments. The transcriptional responses of *S. cerevisiae* during bioethanol industrial fermentation was evaluated by DNA microarray analysis showing that strong transcriptional changes occurred during both continuous and fed-batch fermentation processes (Fig. 2) (14). Repression of glycosylation was observed and upregulation of genes involved in the unfolded protein response, reserve metabolism, and glucose-repressible genes was

thought to be a consequence of protein glycosylation deficiency. On the other hand, many of the overexpressed genes in a sherry flor yeast strain, compared with a laboratory strain, were found to be located within amplified regions in the genome, suggesting that changes in gene expression were due to an increase in DNA copy number and not due to differences in regulation (9).

3. Genomic Features of Industrial Strains of *S. cerevisiae*

The genetics of industrial *Saccharomyces* strains are more complex than that of laboratory reference or academic strains (1). Industrial strains are mainly diploid, aneuploid, or even polyploid. Most of the studies on the characteristics of the genomes of industrial strains have been performed with strains related to the production of alcoholic beverages. This knowledge has been recently reviewed (15).

Codon et al. (16) measured the DNA content and ploidy of 17 industrial *Saccharomyces* strains, including bakers' (11 strains), wine (2 strains), brewers' (2 strains), and distillers' (2 strains) yeasts and compared them to two well-known laboratory strains. All industrial strains had variable DNA contents ranging from 1.3 to 3*n* (*n* being the DNA content of an haploid strain). The nonentire ploidy numbers indicated partial amplifications of certain chromosomal regions and the presence of extra sets of chromosomes. Electrophoretic chromosomal patterns indicated strong polymorphisms between strains. However, all industrial strains possessed a high degree of DNA homology with laboratory yeasts. Hybridization studies indicated that all tested genes were located on the same chromosomes both in laboratory and industrial strains indicating that intrachromosomal arrangements had taken place in industrial strains. The *SUC* gene family includes six loci *SUC1*, *2*, *3*, *4*, *5*, and *7*. These genes encode the invertase enzyme that splits sucrose, the major carbon source in molasses and sugar cane juice, into glucose and fructose. All of these genes, except *SUC2*, have telomeric location (near chromosome ends), highly homologous sequences, and are dispersed on different chromosomes. The different industrial yeasts tested presented multiple copies of different *SUC* genes in their genomes while laboratory strains only carried the *SUC2* gene. This suggests that the amplification of *SUC* genes could be an adaptive mechanism conferring better fitness for fermentation on sucrose juices and molasses.

Infante et al. (17) have used the DNA microarray technology to compare the 16 chromosomes of two yeasts strains isolated from sherry wine aging fermentation that differ on their ethanol and acetaldehyde resistance. Gross chromosomal rearrangements were responsible for the amplification of 116 genomic regions that

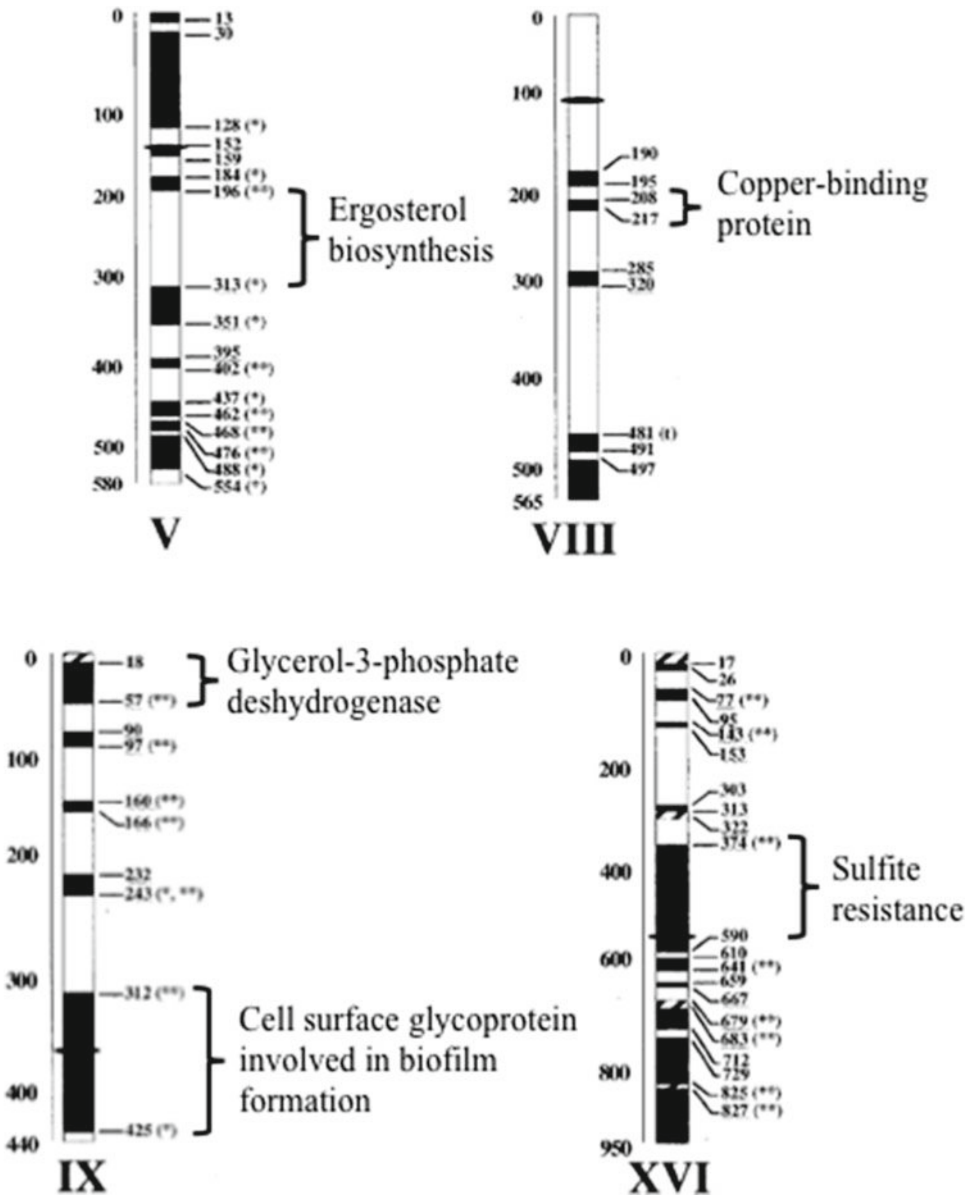


Fig. 3. Genomic rearrangements in the chromosomes of a sherry wine yeast. *Solid bars* indicate amplified regions. *Roman numbers* indicate the chromosome number. The location of genes conferring specific traits is indicated. Adapted from ref. 17.

comprise 38% of their genomes. For one strain, the amplified regions contained genes that were clearly related to specific characteristics as biofilm formation, sulfite and copper resistance, ergosterol biosynthesis, and assimilation of glycerol, a major carbon source in sherry wine (Fig. 3). For the other strain, the amplified regions were related to "unknown biological processes."

Concerning yeasts used to produce fuel ethanol, Lucena et al. (18) have shown that yeasts cells adapt for survival through chromosome rearrangements that can be produced in laboratory media under simulated industrial conditions. Such rearrangements are produced only after 435 generations and are specific to parental genetic backgrounds. The comparison between karyotypes of original and evolved strains showed that the observed chromosome variations originated from the parental strains and not from contaminant yeasts.

Microarray karyotyping and pulse field gel electrophoresis studies have shown that industrial strains of *S. cerevisiae* used in the production of fuel ethanol have amplifications of telomeric genes for the biosynthesis of vitamins B6 and B1 (19). These amplifications allow them to grow better in high sugar environments, where the demand of these vitamins is high and their bioavailability low. Thiamine pyrophosphate is the cofactor of the enzyme pyruvate decarboxylase. Such strains of *S. cerevisiae* have been selected during the industrial production of fuel ethanol and rapidly replace commercial baker's yeast used as "starter" strains in a matter of weeks. They ferment cane juice or diluted molasses containing up to 200 g/L of total sugar, and produce high ethanol concentrations up to 12% (v/v) with 90–92% of the maximum theoretical yield. The fermentation lasts 6–10 h, allowing two to three fermentations per day. At the end of each fermentation, the cells are centrifuged, washed in dilute sulfuric acid, and recycled back for the next fermentation during 6–9 months. The fuel strains appear to cluster together and be more similar to baker's yeast than to brewing or wine strains. Interestingly, none of the fuel strains showed amplification of the *SUC2* gene, indicating that invertase activity is not limiting sucrose fermentation during industrial ethanol production by these yeasts. These authors propose that such a genetic adaptation can also be useful in the development and engineering of yeast strains for the efficient utilization of biomass-derived sugars.

The genome of a haploid derivative of the PE-2 wild isolate currently used to produce bioethanol in 30% of the Brazilian distilleries was sequenced and analyzed (20). This strain possessed a high degree of DNA homology with other sequenced laboratory yeasts, but presented extensive structural differences at the periphery of its chromosomes. These regions typically contain genes that participate in alternative carbon source and vitamin metabolisms, ion and amino acid transport, flocculation, and other processes not essential for viability but related to ecological niche adaptation. Here, again, genes involved in vitamin B6 metabolism were amplified. According to theses authors, the development of genetically engineered yeast strains for industrial applications should be based on strains naturally adapted to industrial conditions; any other strain would be rapidly washed out of the process.

From these studies, it is clear that industrial strains contain traits that enable them to perform well under a particular set of conditions.

4. Genetically Engineered Industrial Yeasts for the Production of Chemicals and Biofuels

4.1. Molecular Tools for the Genetic Engineering of Industrial Strains

Classical genetic techniques have been successfully used for years to improve the performance of production organisms. Such techniques are not site directed. The ability to perform directed genetic changes by genetic engineering techniques allows to add or remove specific features from strains allowing to design new cell factories or improve the existing ones (21).

Laboratory strains of *S. cerevisiae* usually exist as stable haploids, exhibit good mating ability, readily take up exogenous DNA, and contain convenient selectable (auxotrophic) markers. Industrial strains lack many of these properties (1, 22). Such strains are typically prototrophic, diploid, and even polyploid, often heterozygous and homothallic. These characteristics restrict the use of traditional selection systems based on the complementation of recessive auxotrophic markers generally used in yeast genetic engineering. The use of classical genetic techniques is also limited because these strains tend to sporulate poorly, switch their mating type (mediated by the *HO* gene), and produce few viable spores.

In haploid laboratory strains, single genes can be disrupted and the resulting mutants easily selected and studied. This is difficult to achieve in diploid or polyploid and heterozygous strains since multiple copies of a same gene are present and all the copies present on different chromosomes have to be inactivated.

There are several ways of amplifying genes introduced by transformation in order to increase the amount of gene product. One possibility is to insert the gene of interest in a multicopy plasmid vector. Most plasmids are unstable and lost under nonselective conditions. Chromosomal integration of genetic material is a method to avoid the inherent instability of most plasmid vectors. Genomic conserved regions as ribosomal DNA (rDNA) sequences can be used as target for chromosomal integration of genes in industrial yeast strains. A yeast integration plasmid, pIARL28, containing an rDNA sequence as homologous recombination site and the bacterial *kanamycin* gene (*kan* from Tn903) that confers resistance to geneticin for selection of the integrants has been described (23). The rDNA sequence of this plasmid exists at about 140 copies in a yeast cell; integration at the rDNA genes occurs, therefore, very efficiently. The undesirable geneticin resistance sequence can be excised from the transformants by repetitive culture under nonselective conditions. So this plasmid is a useful vector for gene transfer into industrial *Saccharomyces* strains.

In laboratory strains, plasmids containing auxotrophic markers can be introduced and selected (24). To avoid the use of antibiotic resistance markers, the successful generation of auxotrophic mutants in industrial diploid strains of *S. cerevisiae* has been described (25). His⁻, Met⁻, Lys⁻, Trp⁻, Leu⁻, Arg⁻, and Ura⁻ auxotrophic mutants of five sake strains were obtained after UV mutagenesis and screening by conventional replica plating on minimal and rich media. HIS3 was used as a selectable marker for the insertion of a yeast overexpression promoter upstream of the gene encoding the alcohol acetyltransferase by one-step gene replacement in a *his3* mutant, allowing the production of a larger amount of isoamyl acetate, a banana-like flavor in Sake.

Due to obstacles concerning the acceptance of antibiotic selection markers on the industrial scale, the ability to eliminate or recycle markers is very important. The Cre–loxP system is usually used for marker cassette excision purposes in a variety of organisms, enabling marker recycling and elimination during the construction of multiply deletant strains or for industrial applications. Carter y Delneri (26) have developed dominant antibiotic resistance marker cassettes flanked with mutated loxP sites (loxLE and lox2272). The loxP cassettes contain resistance for geneticin (*kanMX* gene), nourseothricin (*natNT2*), or hygromycin (*hphNT1*) as selectable markers. These cassettes can be present in the yeast genome together with the widely used loxP–marker gene–loxP cassettes. The probability of interaction with and between the mutated loxP sites and consequent recombination between the cassettes, or undesirable chromosomal rearrangements between lox sequences, is minimized. A phleomycin-resistant Cre-expressing vector to excise multiple markers simultaneously has also been developed. Such a toolkit should be particularly useful for the genetic engineering of industrial strain bypassing the need of auxotrophic markers and the presence of residual heterologous antibiotic-resistance genes.

4.2. Examples of Genetically Engineered Industrial Yeasts

Many examples of genetic engineering of *S. cerevisiae* have appeared in the two last decades (Table 1), most of them developed in laboratory strains (27, 28).

The development of industrial pentose-fermenting strains of *S. cerevisiae* has been described in the literature and is currently a very active area of research due to the future needs to produce large quantities of biofuels and other chemicals from renewable sources. Pentoses are 5-carbon sugars (C5) which are not naturally fermented by *S. cerevisiae*. They represent up to 25% of the total carbohydrates present in lignocellulosic raw materials (29).

Initially, the genetic engineering of industrial strains had been limited to the introduction of xylose utilization pathways; since *S. cerevisiae* is unable to ferment pentoses, further improvements were obtained by nondirected adaptation strategies as reviewed by Hahn-Hägerdahl et al. (30). Industrial strains supplemented with

Table 1
Genetic engineering of *S. cerevisiae*

Target	Example	Type of strain
Extension of substrate range	Transformation of *S. cerevisiae* with xylose reductase and xylitol deshydrogenase from *Pichia stipitis* or with bacterial xylose isomerase. This enables *S. cerevisiae* to use pentose, a sugar abundant in lignocellulosic biomass	Laboratory and industrial
Improvements of productivity and yield, and elimination of by-products	A *GPD2* (glycerol dehydrogenase) mutant grown under anaerobic conditions, had a higher ethanol yield of 8% in addition to a 40% reduction of the glycerol yield. This is an example of redirecting the carbon flux	Laboratory
Improvement of process performance	*FLO1* encodes a cell surface protein that plays a direct role in the flocculation process. The *FLO1* gene has successfully been integrated into the genome of a nonflocculent brewer's yeast strain and a stable constitutive flocculating strain was obtained	Industrial
Improvements of cellular properties	*MIG1* encodes a zinc finger protein that mediates glucose repression by controlling the expression of *SUC2* (encoding the invertase). Deletion of *MIG1* has allowed a ninefold increase in *SUC2* expression of cells grown on glucose	Laboratory
Extension of product range	Heterologous expression of *LDH* (lactate dehydrogenase) allowed the NADH-dependent reduction of pyruvate to lactate	Laboratory

Adapted from ref. 21 with input from refs. 27, 28

the "yeast pathway" for xylose utilization and overexpressing the endogenous *XK* successfully cofermented xylose and glucose. These genes were generally integrated into the yeast chromosome (in the HIS3 locus, for example) and selected by their ability to grow on xylose. In some cases, the strains were transformed with multicopy plasmids overexpressing these genes and selected by geneticin resistance. The "yeast pathway" comprises two enzymes from *Pichia stipitis*, the XR and the XDH (Fig. 4). It can be redox neutral if the XR is linked to NADH instead of NADPH. The "bacterial pathway" uses the enzyme xylose isomerase (XI) to directly convert D-xylose to D-xylulose (Fig. 4).

Other strategies for the chromosomal integration of the "yeast pathway" genes included the use of the pAUR101 shuttle vector which is a chromosomal integrating vector for *S. cerevisiae* containing the aureobasidin A-resistance gene for integration at the AUR1-C allele (31, 32). The industrial strains tested included bakery yeast, shochu yeasts, wine yeasts, and industrial alcohol fermentation yeasts, including flocculent yeast with high xylulose-fermenting ability. Both xylose consumption and ethanol production remarkably increased in the latter.

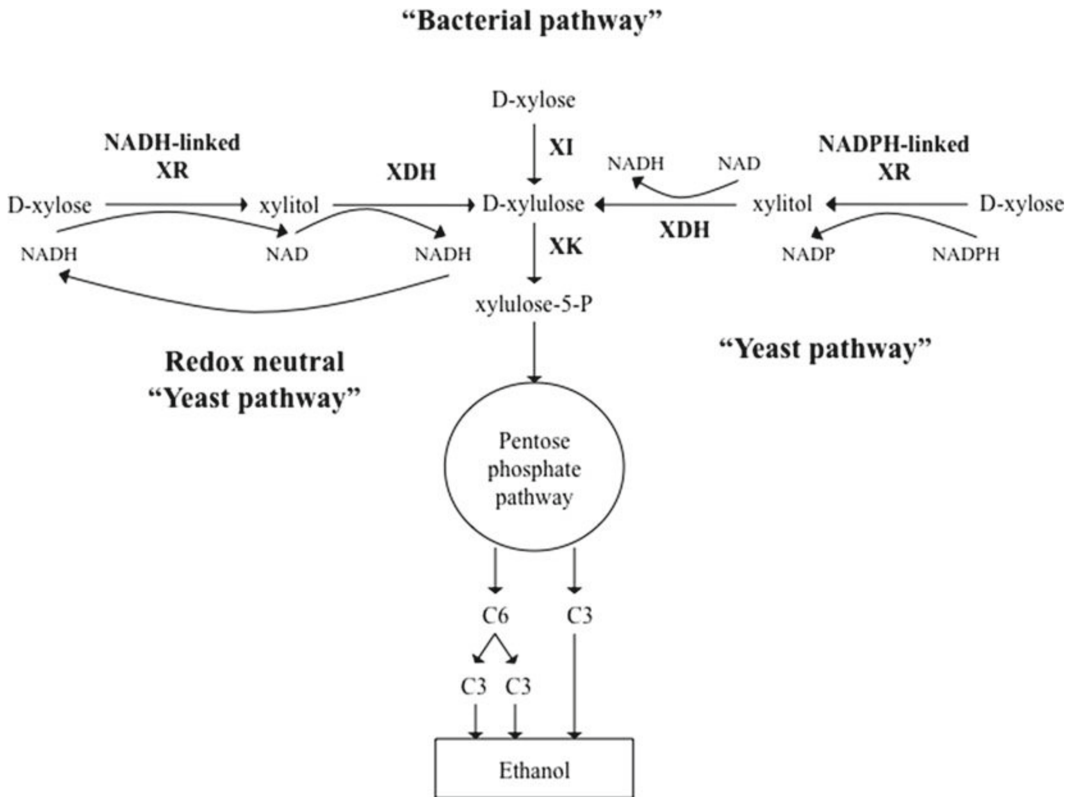


Fig. 4. Pathways for the conversion of xylose to ethanol. *XI* xylose isomerase, *XK* xylulose kinase, *XR* xylose reductase, *XDH* xylitol deshydrogenase. Adapted from ref. 41.

Karhumaa et al. (33) demonstrated the importance of increasing the metabolic steps downstream of xylulose for efficient xylose utilization in *S. cerevisiae*. For this, the genes encoding the XK (*XKS1*) and enzymes of the pentose phosphate pathway (*TAL1* encoding the TAL transaldolase, *TKL1* encoding the TKL transketolase, *RKI1* encoding the ribose-5-phosphate ketol-isomerase RKI, and *RPE1* encoding the RPE ribulose 5-phosphate epimerase) were placed under the control of stronger promoters for overexpression and chromosomally integrated. Chromosomal integration was performed sequentially using a zeocin-resistance marker for selection and the loxP–cre system for marker excision. However, this work was only performed in laboratory strains.

Genome-wide transcription analysis revealed that the expression of an ORF of unknown function (*YLR042c*) was reduced 6.0-fold in strains with improved xylose utilization compared with their respective parental strain; deletion of this ORF improved aerobic growth on xylose (34). Deletion of *YLR042c* improved ethanolic xylose fermentation in several recombinant strains of baker's yeast. However, the magnitude of the effect was strain dependent (35).

The obtained results suggested that YLR042c influences the assimilation of carbon sources other than glucose and has a role at the membrane level when cell damage occurs. The increased xylose flow into the cell resulted in increased xylitol formation, rather than increased ethanol formation. These authors concluded that the transfer of this property to industrial strains might not be so attractive.

Bera et al. (36) have reported industrial strains of *S. cerevisiae* engineered for simultaneous xylose and arabinose utilization. The new strains were based on the previously developed *S. cerevisiae* yeast 424A(LNH-ST) with multiple copies of the *XR*, *XDH*, and *XK* stably integrated in the chromosome. The new strain was constructed by overexpressing fungal genes for L-arabinose utilization in plasmids with hygromycin resistance selection. An ethanol production, about 72.5% the theoretical yield, was achieved from sugar mixtures containing glucose, galactose, mannose, xylose, and arabinose. An inconvenience of this system is that the fermentation medium contained hygromycin. Other industrial strains cofermenting xylose and arabinose were reported (37, 39). In this case, bacterial genes for arabinose utilization were chromosomally integrated in rDNA genes. Evolution of the genetically engineered strains in the presence of xylose and arabinose as sole carbon sources allowed to obtain strains with increased consumption rate of xylose and arabinose under aerobic and anaerobic conditions and improved anaerobic ethanol production at the expense of xylitol and glycerol; however, arabinose was almost stoichiometrically converted to arabitol. These improvements were attributed to mutations in the arabinose genes, duplication of the xylose utilization genes, and increased transport capacity.

Concerning the introduction of "bacterial pathways" for xylose utilization in industrial strains, Brat et al. (40) have described the functional expression of the *Clostridium phytofermentans* XI with high activity. The codon usage of the corresponding gene was adapted to that of highly expressed glycolytic genes of *S. cerevisiae* and chromosomally integrated into the *FAA2* gene encoding the long-chain fatty acid CoA ligase 2 using *kanMX* as the selection marker and geneticin in the selection and fermentation media. This enzyme was less inhibited by xylitol, a side product during xylose fermentation. This study is encouraging for further improvement of xylose fermentation in industrial yeast strains.

On the other hand, Karhumaa et al. (38) compared the performance of strains containing the bacterial and the yeast pathways, concluding that the yeast xylose utilization pathway resulted in faster ethanol production. However, the comparison was performed between two laboratory strains transformed with plasmids, one containing the *Pyromyces* XI gene and the other the *P. stipitis XR y XDH* genes; the performance of industrial strains containing these pathways was not compared and is difficult to predict.

5. Final Remarks

S. cerevisiae is extensively used in the industry in baking, brewing, winemaking, and bioethanol production from conventional carbohydrates (starch and sucrose). The interest for producing bioethanol and other interesting compounds from renewable resources has grown. The process robustness of *S. cerevisiae* makes this yeast a potential platform for the development of genetically engineered yeasts for these new applications. However, this requires the development of stable genetically engineered strains which is not an easy task due to the peculiar characteristics of industrial strains as aneuploidy, polyploidy, frequent chromosomal rearrangements, difficulty of using auxotrophic markers, and problematic of using antibiotic-resistance markers at a commercial scale. Examples of genetically engineered strains for the production of ethanol from lignocellulosic biomass have been described; however, contrary to baking, brewing, and wine applications, there is no knowledge concerning the genetic stability and performance of such strains under real process conditions.

References

1. Walker, G.M. (1998) Yeast technology. In: *Yeast physiology and biotechnology*. p. 265–320. John Wiley and Sons Ltd, Chichester, England.
2. Bai, F.W., Anderson, W.A. and Moo-Young, M. (2008) Ethanol fermentation technologies from sugar and starch feedstocks. *Biotechnol Adv.* **26(1)**:89–105.
3. Branduardi, P., Smeraldi, C. and Porro, D. (2008) Metabolically engineered yeasts: “Potential” industrial applications. *J Mol Microbiol Biotechnol.* **15(1)**:31–40.
4. Werpy, T. and Petersen, G. (2004) Top Value added chemicals from biomass. Vol. 1 Results of screening for potential candidates from sugars and synthesis gas.
5. BREW (2006) Medium and long-term opportunities and risks of the biotechnological production of bulk chemicals from renewable resources – final report. Prepared under the European Commissions GROWTH program, Utrecht.
6. Hahn-Hägerdal, B., Galbe, M., Gorwa-Grauslund, M.F., Lidén, G. and Zacchi, G. (2006) Bio-ethanol-the fuel of tomorrow from the residues of today. *Trends Biotechnol.* **24(12)**: 549–556.
7. Gibson, B.R., Lawrence, S.J., Leclaire, J.P., Powell, C.D. and Smart, K.A. (2007) Yeast responses to stresses associated with industrial brewery handling. *FEMS Microbiol Rev.* **31(5)**: 535–569.
8. Klinke, H.B., Thomsen, A.B. and Ahring, B.K. (2004) Inhibition of ethanol-producing yeast and bacteria by degradation products produced during pre-treatment of biomass. *Appl Microbiol Biotechnol.* **66(1)**:10–26.
9. Querol, A., Fernández-Espinar, M.T., del Olmo, M. and Barrio, E. (2003) Adaptative evolution of yeast. *Int J Food Microbiol.* **86**: 3–10.
10. Gibson, B.R., Lawrence, S.J., Leclaire, J.P., Powell, C.D. and Smart, K.A. (2007) Yeast responses to stresses associated with industrial brewery handling. *FEMS Microbiol Rev.* **31(5)**: 535–569.
11. Silva-Filho, E.A., dos Santos, S.K.B., Resende, A.M., de Moraes, J.O.F., Morais Jr, M.A. and Simoes, D.A. (2005) Yeast population dynamics of industrial fuel ethanol fermentation process assessed by PCR-fingerprinting. *Antonie van Leeuwenkoek* **88**:13–23.
12. Basso, L.C., de Amorim, H.V., de Oliveira, A.J. and Lopes, M.L. (2008) Yeast selection for fuel ethanol production in Brazil. *FEMS Yeast Res.* **8(7)**:1155–1163.
13. Cocolin, L., Pepe, V., Comitini, F., Comi, G. and Ciani, M. (2004) Enological and genetic

traits of *Saccharomyces cerevisiae* isolated from former and modern wineries. *FEMS Yeast Res.* **5(3)**:237–245.

14. Li, B.Z., Cheng, J.S., Qiao, B. and Yuan, Y.J. (2010) Genome-wide transcriptional analysis of *Saccharomyces cerevisiae* during industrial bioethanol fermentation. *J Ind Microbiol Biotechnol.* **37(1)**:43–55.
15. Querol, A. and Bond, U. (2009) The complex and dynamic genomes of industrial yeasts. *FEMS Microbiol Lett.* **293(1)**:1–10.
16. Codón, A.C., Benítez, T. and Korhola, M. (1998) Chromosomal polymorphism and adaptation to specific industrial environments of *Saccharomyces* strains. *Appl Microbiol Biotechnol.* **49(2)**:154–163.
17. Infante, J.J., Dombek, K.M., Rebordinos, L., Cantoral, J.M. and Young, E.T. (2003) Genome-wide amplifications caused by chromosomal rearrangements play a major role in the adaptive evolution of natural yeast. *Genetics* **165(4)**:1745–59.
18. Lucena, B.T., Silva-Filho, E.A., Coimbra, M.R., Morais, J.O., Simões, D.A. and Morais, M.A. Jr. (2007) Chromosome instability in industrial strains of *Saccharomyces cerevisiae* batch cultivated under laboratory conditions. *Genet Mol Res.* **6(4)**:1072–1084.
19. Stambuk, B.U., Dunn, B., Alves, S.L. Jr, Duval, E.H., Sherlock, G. (2009) Industrial fuel ethanol yeasts contain adaptive copy number changes in genes involved in vitamin B1 and B6 biosynthesis. *Genome Res.* **19(12)**:2271–2278.
20. Argueso, J.L., Carazzolle, M.F., Mieczkowski, P.A., Duarte, F.M., Netto, O.V., Missawa, S.K., Galzerani, F., Costa, G.G., Vidal, R.O., Noronha, M.F., Dominska, M., Andrietta, M.G., Andrietta, S.R., Cunha, A.F., Gomes, L.H., Tavares, F.C., Alcarde, A.R., Dietrich, F.S., McCusker, J.H., Petes, T.D. and Pereira, G.A. (2009) Genome structure of a *Saccharomyces cerevisiae* strain widely used in bioethanol production. *Genome Res.* **19(12)**:2258–2270.
21. Bro, C., Regenberg, B. and Nielsen, J. (2003) Yeast functional genomics and metabolic engineering: past, present and future. Topics in Current Genetics, Vol. 2, J.H. de Winde (Ed): Functional genetics of Industrial Yeasts. Springer Verlag, Berlin, Heidelberg.
22. Spencer, J.F. and Spencer, D.M. (1983) Genetic improvement of industrial yeasts. *Annu Rev Microbiol.* **37**:121–142.
23. Fujii, T., Kondo, K., Shimizu, F., Sone, H., Tanaka, J. and Inoue, T. (1990) Application of a ribosomal DNA integration vector in the construction of a brewer's yeast having alpha-acetolactate decarboxylase activity. *Appl Environ Microbiol.* **56(4)**:997–1003.
24. Mumberg, D., Müller, R. and Funk, M. (1995) Yeasts vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* **156(1)**:119–122.
25. Hashimoto, S., Ogura, M., Aritomi, K., Hoshida, H., Nishizawa, Y. and Akada, R. (2005) Isolation of auxotrophic mutants of diploid industrial yeast strains after UV mutagenesis. *Appl Environ Microbiol.* **71(1)**:312–319.
26. Carter, Z. and Delneri, D. (2010) New generation of loxP-mutated deletion cassettes for the genetic manipulation of yeast natural isolates. *Yeast* **27(9)**:765–775.
27. Ostergaard, S., Olsson, L. and Nielsen, J. (2000) Metabolic engineering of *Saccharomyces cerevisiae. Microbiol Mol Biol Rev.* **64(1)**: 34–50.
28. Nevoigt, E. (2008) Progress in metabolic engineering of *Saccharomyces cerevisiae. Microbiol Mol Biol Rev.* **72(3)**:379–412.
29. Lee, J. (1997) Biological conversion of lignocellulosic biomass to ethanol. *J Biotechnol.* **56**:1–24.
30. Hahn-Hägerdal, B., Karhumaa, K., Fonseca, C., Spencer-Martins, I. and Gorwa-Grauslund, M.F. (2007) Towards industrial pentose-fermenting yeast strains. *Appl Microbiol Biotechnol.* **74(5)**:937–953.
31. Matsushika, A., Inoue, H., Watanabe, S., Kodaki, T., Makino, K. and Sawayama, S. (2009a) Efficient bioethanol production by a recombinant flocculent *Saccharomyces cerevisiae* strain with a genome-integrated $NADP^+$-dependent xylitol dehydrogenase gene. *Appl Environ Microbiol.* **75(11)**:3818–3822.
32. Matsushika, A., Inoue, H., Murakami, K., Takimura, O. and Sawayama, S. (2009b) Bioethanol production performance of five recombinant strains of laboratory and industrial xylose-fermenting *Saccharomyces cerevisiae. Bioresour Technol.* **100(8)**:2392–2398.
33. Karhumaa, K., Hahn-Hägerdal, B. and Gorwa-Grauslund, M.F. (2005) Investigation of limiting metabolic steps in the utilization of xylose by recombinant *Saccharomyces cerevisiae* using metabolic engineering. *Yeast* **22(5)**:359–368.
34. Bengtsson, O., Jeppsson, M., Sonderegger, M., Parachin, N.S., Sauer, U., Hahn-Hägerdal, B. and Gorwa-Grauslund, M.F. (2008) Identification of common traits in improved xylose-growing *Saccharomyces cerevisiae* for inverse metabolic engineering. *Yeast* **25(11)**: 835–847.
35. Parachin, N.S., Bengtsson, O., Hahn-Hägerdal, B. and Gorwa-Grauslund, M.F. (2010) The deletion of YLR042c improves ethanolic xylose fermentation by recombinant *Saccharomyces cerevisiae. Yeast* **27(9)**:741–751.

36. Bera, A.K., Sedlak, M., Khan, A. and Ho, N.W. (2010) Establishment of L-arabinose fermentation in glucose/xylose co-fermenting recombinant *Saccharomyces cerevisiae* 424A(LNH-ST) by genetic engineering. *Appl Microbiol Biotechnol.* **87(5)**:1803–1811.
37. Karhumaa, K., Wiedemann, B., Hahn-Hägerdal, B., Boles, E. and Gorwa-Grauslund, M.F. (2006) Co-utilization of L-arabinose and D-xylose by laboratory and industrial *Saccharomyces cerevisiae* strains. *Microb Cell Fact.* **5**:18.
38. Karhumaa, K., Garcia Sanchez, R., Hahn-Hägerdal, B. and Gorwa-Grauslund, M.F. (2007) Comparison of the xylose reductase-xylitol dehydrogenase and the xylose isomerase pathways for xylose fermentation by recombinant *Saccharomyces cerevisiae*. *Microb Cell Fact.* **6**:5.
39. GGarcia Sanchez, R., Karhumaa, K., Fonseca, C., Sánchez Nogué, V., João RM Almeida, J.R.M., Larsson, C.U., Bengtsson, O., Bettiga, M., Hahn-Hägerdal, B. and Gorwa-Grauslund, M.F. (2010) Improved xylose and arabinose utilization by an industrial recombinant *Saccharomyces cerevisiae* strain using evolutionary engineering. *Biotechnol Biofuels* **3**:13.
40. Brat, D., Boles, E. and Wiedemann, B. (2009) Functional expression of a bacterial xylose isomerase in *Saccharomyces cerevisiae. Appl. Environ. Microbiol.* **75(8)**: 2304–2311.
41. Kuyper, M., Harhangi, H.R., Stave, A.K., Winkler, A.A., Jetten, M.S., de Laat, W.T., den Ridder, J.J., Op den Camp, H.J., van Dijken, J.P. and Pronk, J.T. (2003) High-level functional expression of a fungal xylose isomerase: The key to efficient ethanolic fermentation of xylose by *Saccharomyces cerevisiae*? *FEMS Yeast Res.* **4(1)**:69–78.

Part V

Plants and Plant Cells

Chapter 25

Recombinant Protein Production in Plants: Challenges and Solutions

Elizabeth E. Hood and Deborah V. Vicuna Requesens

Abstract

In a recent presentation at the 2010 International Association for Plant Biotechnology meeting, Dr. Richard Flavell (Ceres, Malibu, CA, USA) motivated the plant community to act quickly and with purpose to move a multitude of traits into crop plants to improve their productivity. Current progress toward understanding of plants is too slow and will not achieve our communal goal of doubling agricultural productivity by 2050. Major breakthroughs are necessary! Thus, high-throughput methods that couple gene identification and phenotype observations are required to put potential products into the hands of plant breeders to make varieties with good agronomic characteristics that will be approved by the regulatory agencies. These first improved crops must be on the market in the next 10 years, according to Flavell, in order to begin to meet our doubled productivity goals in 30 years. Because it takes approximately 10 years to produce a characterized variety from an identified gene and move it through product development and regulatory approval, we must begin now. Presumably, by employing the techniques in the following chapters, we can do that.

Key words: Plant-based products, Recombinant proteins, Protein expression, Protein accumulation, Intellectual property, Regulatory affairs

1. Introduction

In our ever-changing and more crowded world in which the population is predicted to be approximately ten billion by 2050, all technologies are needed to maximize production of our food, feed, fiber, and fuel requirements. Moreover, these technologies and products need to come from a renewable bio-based industry to protect the vitality of the environment. Our current dependence on nonrenewable raw materials, such as fossil oil, coal, and natural gas, is not sustainable. Not only are the supplies being depleted, their use contributes to climate change and pollution.

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_25, © Springer Science+Business Media, LLC 2012

Renewable resources are biological – they can grow and reproduce. The basis for all those resources is of course the plant kingdom with energy input from the sun. Thus, maximizing plant variety (for myriad products) and productivity (for maximal biomass) is critical to maintenance of long-term quality of life.

A host of plant-based technologies can contribute to improved plant production to meet these challenges, including improved production practices and improved plant varieties generated through plant breeding, tissue culture, and biotechnology. Although breeding and tissue culture are techniques that have been accepted for many years as valid methods for crop improvement, plant biotechnology is faced with many challenges, not the least of which is social acceptance of the products from genetic engineering. Moreover, to allay these social fears, federal regulatory agencies have instituted multiple layers of control and containment governing genetically engineered plant release into the environment. An additional challenge to using plant genetic engineering or biotechnology involves the intellectual property landscape. Many useful processes, genes, regulatory regions, and methods have been patented and require licenses in order to use them for commercial purposes. Many academic researchers use these technologies without licenses to identify new phenotypes useful in agricultural applications. However, identifying an application or useful trait is only the beginning of being able to apply that finding to a product that can be used for food, feed, fiber, or fuel. Thus, whoever moves the new concept into a product has to be focused on all the above issues.

In this chapter, we focus on the production of protein products in plants using biotechnology that can be used for vaccines, pharmaceuticals, and industrial products and processes. This is a young technology with much room for significant advances. Many of the technologies that have been developed to maximize protein expression in plants have been developed by groups who focus on these types of applications. Nevertheless, the technologies developed by these groups also can be applied to protein expression for metabolic engineering of pathways or agronomic trait alterations.

Currently, thanks to the advances in molecular biology over the past 20 years, we have the amazing ability of producing proteins of interest while crossing the species barrier. Plant molecular farming has allowed us to produce recombinant pharmaceutical molecules and vaccines, enzymes with industrial applications, and other metabolites in plants. Potentially, the production of recombinant protein in plants presents us with great advantages over other systems, linked in particular to the possibility of large-scale production and the potential of low costs of production (1). Unlike microorganism systems, they have the advantage of eukaryotic posttranslational modifications, most importantly glycosylation (2). Recombinant proteins in plants present lower risk of pathogen contamination; high levels of expression; direct delivery of the

product to the recipient animal, person, or industrial process; lack of necessity for purification in many cases; and, in the case of industrial enzymes, volumes are anticipated to be very large at a low capital investment (3).

We discuss challenges and limitations associated with expressing foreign genes in plants, intellectual property issues, and regulatory considerations. This is not intended to be a "how-to" chapter, but to set the stage for the following chapters that address the methods for generating new plants through biotechnology.

2. Expression Systems

When trying to maximize the expression of a recombinant protein in plants, several aspects must be taken into consideration. The plant system to be used depends on the approach to generate the transgenic plants and the tissue type chosen for regeneration. There are three main types of plant systems: (1) biolistics or *Agrobacterium*-mediated transformation which is used to integrate recombinant gene sequences into the host plant's nuclear genome; (2) the biolistics approach used to generate transplastomic plants and incorporate recombinant sequences into the plastid genome; and (3) inoculation of plant tissue with recombinant plant viral vectors. In the third case, the sequences are transiently expressed in infected plant tissue, usually leaves (4).

Once the plant system has been chosen, many other factors need to be considered, such as the type of tissue to be transformed, targeting of the recombinant protein, posttranslational processing, and the potential of contamination with harmful secondary metabolites. Nonetheless, the choice of host species and production system is largely influenced by the production costs inherent to the recovery of the recombinant protein. Numerous factors must be taken into account, such as the setup costs of the transformation and plant regeneration system and the actual time necessary for this process; costs associated with scaling up and maintaining the transgenic plants; biomass yield of the chosen plant; processing costs related to extraction, purification, and characterization; costs of storage and containment; and of course costs of deregulating the transgenic crop and its commercialization, which are discussed in this chapter (5).

3. Strategies to Increase Expression of Recombinant Proteins

Many strategies and technical considerations are used to maximize the production of recombinant proteins in plants. Here, we discuss the possible approaches to obtain higher expression and higher production.

3.1. Molecular Approaches to Increase Gene Expression

3.1.1. Transcription

One of the most studied approaches used to boost expression of foreign proteins in stably transformed plants is to increase the levels of transcription. To obtain high levels of transcription, the most important element is the promoter used to drive the expression of the transgene. Many promoters have been tested and they can be of plant origin, obtained from plant pathogens or synthesized. The most widely used promoter is the CaMV 35S promoter from the cauliflower mosaic virus (6). This viral promoter is a strong constitutive promoter and the most popular choice for dicotyledonous plants. Even though this promoter has a lower activity in monocotyledonous plants, it is often used to drive expression of selectable marker genes. Alternatives, such as the maize ubiquitin-1 promoter (7) or the *Agrobacterium tumefaciens* nopaline synthase (*nos*) promoter (8), are usually better suited to obtain higher expression in monocots. In order to obtain expression in a particular tissue of the plant or during a specific developmental stage, tissue-specific promoters have been developed. Examples of this type of promoter are the maize zein, rice and wheat glutenin, and pea legumin gene promoters (1). The advantage of this type of promoter resides in the restriction of expression of the transgene to certain areas, increasing the stability of the protein, and avoiding the detrimental effects of accumulation of the recombinant protein in vegetative plant tissues which can be toxic to the host plant. It is important to understand that the same promoter can result in distinct, tissue-dependent accumulation patterns within different species and thus, in order to maximize levels of recombinant protein in plants we must choose a promoter that suits the specific system (9).

3.1.2. Transcript Stability

To ensure high-level expression of foreign proteins in plants, we must assure stability of the message once it is transcribed. Several strategies can be considered to increase transcript stability.

Polyadenylation sites directly affect message stability. mRNAs contain long 3′ untranslated regions comprising hundreds of nucleotides. They undergo cleavage and polyadenylation at polyadenylation sites. These untranslated regions contain regulatory elements affecting mRNA stability or translation efficiency, and thus we could choose alternate polyadenylation sites in plant transformation vectors to affect the final expression of transgenes (9, 10). Specific sequences from plants and plant viruses, such as the CaMV 35S terminator and the potato proteinase inhibitor II (Pin II) terminator, can be utilized to process messages without destabilizing them (2).

The presence of introns in a gene construct is believed to increase accumulation of mRNA in plants, particularly grasses. They occur naturally in eukaryotic genomic DNA and must be spliced from a pre-mRNA before leaving the nucleus (11, 12). The introduction of introns in recombinant DNA used for plant transformation can enhance the stability of the RNA and in turn

increase the level of recombinant protein produced. In monocots, some of the introns known to stimulate expression are the maize *Adh1*, *Ubi1*, *Sh1*, and actin and the rice *sal*T, *act1*, and *tpi* genes, among others. In dicots, examples of introns known to elevate expression include the potato *ST-LS1* gene and the *Arabidopsis UBQ3*, *PAT1*, and *EF-1*α genes. An important fact is that different introns affect expression in different ways. In plants, introns affect mRNA accumulation posttranscriptionally. They facilitate maturation and enhance the stability of new transcripts (11). Studies indicate that introns increase mRNA accumulation by enhancing production rather than by reducing degradation (13). When preparing a plant construct, introns must be located within transcribed sequences to elevate mRNA accumulation. Nonetheless, when introns are placed in the 3′ untranslated regions, their ability to stimulate expression is greatly reduced. When located more than 50 nucleotides downstream of the coding sequences, introns might trigger RNA decay by allowing the stop codon to be recognized as premature and thus negate their positive effects (14). Also, introns are less effective at increasing expression when near the start of a gene or when placed near the 3′ end of a transcript (13). A noted exception is the maize ubiquitin intron that is just downstream of a short, noncoding exon.

Additionally, to boost transcript stability, one must be aware of sites that promote RNA decay. RNA decay can be enhanced by specific recognition sites that are thought to correspond to binding sites. To improve accumulation of recombinant protein in transgenic plants, it may be necessary to ensure that these specific sites are removed when they exist in the gene of interest (12).

3.1.3. Replication

With the more conventional plant transformation systems mentioned above, the best way to achieve high-level expression is to couple the transgene with a suitable promoter. However, highly expressed transgenes are susceptible to posttranscriptional gene silencing, reducing the accumulation of the mRNAs. New methods, such as plant viral vectors, constitute a viable alternative (15). Even though the classical system for expressing recombinant proteins in plants uses stable genetic transformation, this alternative method of plant virus-based vectors achieves high expression of foreign proteins in transient transformations (16). In viral systems, recombinant proteins are encoded from an engineered vector that offers high-level expression by RNA-directed RNA amplification using viral replicases. Viral genomes are small and easy to manipulate, and the infection of plants with this type of vector is simpler and faster than stable transformation methods (16). Nonetheless, this system has some difficulties and presents some challenges. There is a size limitation on the foreign protein expressed, the foreign gene is not heritable, the genes inserted into viral genomes can sometimes be spliced out, and there are concerns regarding the potential

spread of the modified viruses in the environment. Ongoing research addresses some of these issues by using the viral replication machinery to boost expression. The potential of this type of system to achieve high accumulation of recombinant proteins is enormous (15, 17, 18).

3.1.4. Translation

Numerous approaches for improving recombinant protein production deal with strategies designed to boost translation. When trying to increase the translation efficiency in a recombinant system, codon optimization should be a priority. Different organisms have particular codon usage. Variations exist even between monocot and dicot plants. The nucleotide sequence of the transgene must be then modified to optimize its codon preference to suit the host plant without changing the amino acid sequence (19).

Along with optimizing the codons of transgenes, certain plant and plant viral 5 nontranslated sequences, such as the tobacco mosaic virus and the potato virus leader sequences, have been used to boost translation initiation (19). It is necessary to modify any sequence located immediately next to the translation start site to fit the consensus initiation sequence, which can vary from species to species (20). Furthermore, we should try to avoid predicted mRNA secondary structures, since they might impede translation or result in premature termination of transcripts (4).

An alternative strategy for boosting protein expression by increasing translation is to construct a translational fusion between the recombinant protein of interest and a second protein. This might increase expression from the mRNA by stabilizing the fused protein or improving the solubility and folding of the protein (21).

3.2. Genetics

Increasing the transgene copy number is a technique that has been applied to boost expression of transgenes in genetically engineered plants. One way to achieve this is to self-pollinate single-transgenic homozygous plants or crossing high-expressing transgenic events with each other to increase copy number and expression levels (4). However, transgenes present in multiple copies are more likely to be silenced. Silencing in plants occurs in two common instances: (1) at a transcriptional level, where promoter inactivation or methylation may occur; (2) at a posttranscriptional level, when specific destruction of transcripts may occur due to the presence of homologous, double-stranded RNA. Silencing is strongly correlated with the number of transgenes in vegetative tissue (20). Several strategies have been utilized to reduce this problem. Transgenic plants obtained by agrobacterial infection are generally less subject to silencing than those obtained through biolistics (22). Also, transgenic plants in which the new genes have been inserted into the chloroplast DNA, also called transplastomic plants, are not subject to silencing (23, 24). Additionally, the formation of antisense transcripts can be prevented by flanking the

transgenes with matrix attachment region sequences (25). Often, expressing a protein in seed rather than vegetative tissue could reduce this phenomenon (26).

A second genetic approach to boost expression and generate high protein accumulation is to make use of germplasm. Unique germplasm, such as high oil in maize, can be used to increase yield, influence germination, and maintain or improve transgene expression (26). Other germplasm includes transgenic lines expressing a cytokinin-synthesizing enzyme under the control of a senescence-inducible promoter, creating two normal-size embryos (27). Such germplasm has an increased embryo-to-endosperm ratio and is of great benefit for recombinant proteins expressed in the embryo.

3.3. Protein Accumulation and Stability

Low protein yield is an important limiting factor for the production of recombinant proteins in plants. In fact, high mRNA levels do not guarantee high levels of protein accumulation. One of the most critical factors influencing the accumulation of heterologous proteins is the presence of proteases in the host plant. They degrade incorrectly folded or abnormal proteins. Several techniques have been developed to avoid or reduce the effect of proteases on the stability of recombinant proteins and they include expressing proteins in specific organelles, directing expression to a tissue or at a particular time, coexpression with protease inhibitors, or use of protein fusions (4).

Recombinant proteins can be targeted to several subcellular compartments in the plant cell, including the mitochondria, the chloroplast, and the secretory pathway. Without a signal sequence added to the transgene, the recombinant protein accumulates in the cytoplasm, where it could be exposed to proteases. Several signal sequences have been identified and used that can target the recombinant protein to membrane-bound organelles, where the effect of proteolytic degradation can be minimized. Sequences, like KDEL (Lys-Asp-Glu-Leu) and HDEL (His-Asp-Glu-Leu), are used to retain the protein in the endoplasmic reticulum (ER), where few proteases are found and a protective environment for the recombinant protein is provided. Also, the presence of chaperone proteins in the ER provides help to fold and assemble foreign proteins (20, 28). An additional benefit of targeting foreign proteins to the ER is the ability to obtain posttranslational modifications of recombinant proteins, especially glycosylation (1). Alternatively, when seeds are the target organ, protein storage vacuoles (PSVs) are a suitable subcellular compartment (29). They are specific to plants and are the main storage site for accumulation of protein in seeds.

Although subcellular location is important, the tissue where proteins are accumulated can play an essential role as well. Not only does it influence the accumulation and stability due to inherent conditions of the specific tissue, but also the product application

planned for the recombinant protein can be crucial when choosing the tissue. Proteins can be expressed in seeds or tubers or specifically in an edible tissue in the case of oral vaccines (30).

Accumulation of proteins can equally be enhanced by restricting the expression to a certain time or inducing with a certain treatment. The use of inducible promoters can limit the negative effects of foreign proteins on the plant host and can increase accumulation by focusing on a restricted time frame.

Diverse strategies have been developed to increase and maximize the yield of recombinant proteins in transgenic plants. These include choosing strong promoters, dealing with copy numbers, improving transcript stability and translational efficiency by optimizing the transgene sequence, and generally increasing protein accumulation by directing expression to a specific target tissue or utilizing the benefit of certain germplasm. Advances in these areas are required in the future to optimize the expression of recombinant proteins, and perhaps the combination of several of these strategies to increase the expression of a single protein results in better yields.

4. Intellectual Property

Wow! You have a new rice plant that exhibits 50% increased growth with a single gene added through genetic engineering. You want to file a patent on this phenomenon because this will really be of interest to the big agricultural companies, you believe. What should you do?

Navigating the intellectual property (IP) landscape is daunting to many small companies and academic crop developers. Publishing papers in which new phenotypes are observed can be done without being too concerned with IP and freedom to utilize the methods and DNAs that are patented by others. However, IP is such an important piece of the commercialization process that the topic cannot be ignored long term. Once commercialization is the goal, IP issues often control the path forward.

What are the mechanisms of protecting your plant-produced products? A number of mechanisms are available. Filing patents is only one form of protection, and is the most expensive and time consuming. So, why file a patent rather than a plant variety protection (PVP) application, publication, or keeping a trade secret? Patents are the most comprehensive form of technology protection to prevent others from infringing on your ability to earn money from your invention (31). A trade secret is only protective until and unless someone reveals it to the public. At that time, it is no longer protected. Once it is revealed, in the USA, the inventor has 1 year from disclosure to file a patent application. However, in

most other countries, a public disclosure negates the ability to file a patent on the technology.

In an academic setting, filing patents is often handled by a research office, a research foundation, or a technology transfer office. First, you would file an invention disclosure – a paper that describes why your invention is important and why the university should be interested. You have to think about why this might be important to some entity to produce the products or services your invention enables. How much might someone be enticed to pay for a licensing fee? It is often necessary to enlist the help of a marketing person to figure out how much the invention might be worth. Who will benefit from the invention? In the rice case, does the gene work in other plants? If so, which ones? Inventions with broad applicability generally have the greatest value.

In a small company, a patent filing is usually handled by a patent attorney on an ad hoc basis. However, the same questions apply to determine the value of the patent.

PVP prevents reproduction of a new variety that has key characteristics similar to your variety. However, your germplasm or variety can be used to create a new variety as long as the new variety does not exhibit the essential characteristics that are present in your variety (31). The PVP is a great vehicle to protect a biotechnology-derived plant that expresses your protein of interest in a particular way, even if a patent has not been filed.

Publication is a way to prevent others from patenting your work. A publication is considered a public disclosure of information and could then be used against the issuance of a patent on similar work. The public disclosure in this instance is termed “prior art” in the context of filing a patent. When the patent office reviews patent applications, they look in the literature, the database of patents that are published, and any other sources to determine if your idea might have been derived from prior work. Any prior art identified by the patent examiner is then cited as a reason for not to patent your invention. The burden is on the inventor to “prove” to the patent examiner that the invention is novel, not obvious, and has value.

What is freedom to operate (FTO) and why do you need it? FTO means that you have obtained licenses (i.e., permission) to use patented technology. Alternatively, one can use technology that is in the “public domain,” meaning published or nonpatented technology. Determining FTO is a mine field and is often a moving target. This is done through searching of published literature and patents, examining the relevant finds, and figuring out if there is a reason to license a particular technology. In an active field, such as plant biotechnology, it is important to engage the services of an attorney who has related experience to the field. This person would know the most common relevant patents and the best questions to ask about uses and types of technology. Sometimes, the best way to

attain FTO is to choose a different genetic regulatory element that works as well as the original one in your gene expression cassette. Nevertheless, a blocking technology may be the only way to do a particular experiment. In this latter case, a license is in order from the assignee/owner of the patent. Licenses can be structured in many ways, including, for example, an agreement that allows a small company to only pay if an income is obtained – i.e., pay a royalty on revenues earned. In this way, large up-front fees may be avoided. Whatever the case, knowing the competition prevents downstream heartache and headache.

Suppose you now have your filed patents and your FTO in place. You are ready to commercialize your product. But you receive a letter from a large biotechnology company claiming that it invented your technology first – that it deserves the patent! How do you ensure that you have supportive information for your patent challenges? The best way to ensure that you were the first to invent is through your laboratory notebooks. Keeping an accurate, detailed laboratory notebook is the most frustrating part of research for many people. However, it is the one best document to stand up to legal scrutiny. The important features of a laboratory notebook are (1) that it is bound with numbered pages so that they cannot be added or taken out; (2) that each page is dated and signed by the scientist and a witness; and (3) that any added data, such as photos or printouts, are permanently affixed to the page and signed across the edge. An accurate and complete notebook also makes for easy writing of patent applications because the data are easy to find.

A well-kept notebook with proper signatures and dates can corroborate date of conception of an idea and who was involved (inventors), provide proof of reduction to practice, and demonstrate diligence in moving the idea through to a patent. Moreover, the information must be complete and clear to a fellow scientist so that he/she could repeat the experiments and obtain similar results.

This discussion is meant to be a guide to thinking about different aspects of IP so that the scientist can think about what questions to ask a patent attorney. The authors are not attorneys, nor are they competent to offer legal advice. However, the issues discussed apply to any biotechnology-derived plant and can affect one's ability to reach commercialization.

5. Regulatory Issues

Do you want to commercialize a biotechnology product? If so, regulatory affairs are important. The regulatory barriers that face crop developers in commercializing biotechnology-derived crops are numerous. First, the regulatory framework is complex and

difficult to understand. A road map does not exist that would direct a developer in the process for achieving nonregulated status from the Animal and Plant Health Inspection Service (APHIS), achieving registration at the Environmental Protection Agency (EPA), or for voluntary consultation with the Food and Drug Administration (FDA). Collecting relevant and comprehensive data sets with the proper controls while also controlling associated costs is challenging. Standards of data collection and analysis are not set. Normal ranges to establish "substantial equivalence" (32) to which the transgenic crop is to be compared are accepted for commodity crops, but few if any specialty crops.

It is desirable to address these barriers and take advantage of the benefits of biotechnology along with the contributions of other technologies in agriculture. Providing food, feed, fiber, and fuels for the growing human population requires all technologies available to us. A new paradigm is needed for data collection and coordination to simplify and standardize the petition process for federal agencies – fulfilling data needs for crop assessment while ensuring safety of the product.

What is needed to achieve nonregulated status of biotechnology-derived crops? A petition must be submitted to the US Department of Agriculture (USDA) to be reviewed and assessed for safety of the product. Determining the data necessary to complete a petition can be confusing and daunting. The data to be collected include safety of the inserted gene product, molecular characterization, composition, allergenicity, toxicity, as well as field performance to assess disease resistance, yield, and insect resistance among others. Environmental assessments and/or impact are also required, the depth of which depends on the product.

New technologies are being developed that can be of use to the regulatory process. For example, meta-analysis, a common technique used to assess multiple clinical study data, can be used to assess diverse studies of transgenic plant characteristics. Comparison of multiple data sets may show that certain data are not necessary to collect because they never vary for a particular type of transgenic plant. Metagenomics, on the other hand, can be used to assess the impact of transgenic plants on mixed populations of organisms. Next-generation sequencing has been applied to a number of complex populations to understand changes in those populations during a change in the milieu (33–35).

Data and text mining may also become tools for understanding plant normality and complexity.

The Specialty Crop Regulatory Assistance (SCRA) program was formed in response to a workshop sponsored by the USDA that brought together scientists from across the nation to discuss the issues of facilitating the deregulation process specifically for small market specialty crops (http://www.specialtycropassistance.org/). However, the goal of this organization also includes

standardization of the application process and data collection. The SCRA is working with consultants and data collection organizations to establish a pipeline of transgenic plant products.

References

1. Ma, J. K.-C., Drake, P., and Christou, P. (2003) The production of recombinant pharmaceutical proteins in plants, *Nature Reviews Genetics 4*, 794–805.
2. Hood, E., Witcher, D., Maddock, S., Meyer, T., Baszczynski, C., and al, e. (1997) Commercial production of avinid from transgenetic maize: characterization of transformant, production, processing, extracting, and purification, *Molecular Breeding 3*, 291–306.
3. Howard, J., Nikolov, Z., and Hood, E. (2011) Enzyme production systems for biomass conversion, In *Plant Biomass Conversion* (Hood, E., Nelson, P., and Powell, R., Eds.), Wiley-Blackwell, Ames, IA, pp. 227–253.
4. Streatfield, S. (2007) Approaches to achieve high-level heterologous protein production in plants, *Plant Biotechnology Journal 5*, 2–15.
5. Schillberg, S., Twyman, R., and Fischer, R. (2005) Opportunities for recombinant antigen and antibody expression in transgenic plants-technology assessment, *Vaccine 23*, 1764–1769.
6. Odell, J., Nagy, F., and Chua, N. (1985) Identification of DNA sequences required for activity of the cauliflower mosaic virus 35 S promoter, *Nature 313*, 810–812.
7. Christensen, A., Sharrock, R., and Quail, P. (1992) Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation, *Plant Molecular Biology 18*, 810–812.
8. Shaw, C., Carter, G., Watson, M., and Shaw, C. (1984) A functional map of the nopaline synthase promoter, *Nucleic Acids Research 12*, 7831–7846.
9. Stoger, E., Sack, M., Perrin, Y., Vaquero, C., Torres, E., and et al., (2002) Practical considerations for pharmaceutical antibody production in different crop systems, *Molecular Breeding 9*, 149–158.
10. Beaudoing, E., and Gautheret, D. (2001) Identification of Alternate Polyadenylation Sites and Analysis of their Tissue Distribution Using EST data, *Genome Research 11*, 1520–1526.
11. Rose, A., and Beliakoff, J. (2000) Intron-mediated Enhancement of Gene Expression Independent of Unique Intron Sequences and Splicing, *Plant Physiology 122*, 535–542.
12. Kusnadi, A. R., Nikolov, Z. L., and Howard, J. A. (1997) Production of recombinant proteins in transgenic plants: Practical considerations, *Biotechnology and Bioengineering 56*, 473–484.
13. Rose, A. (2004) The effect of intron location on intron-mediated enhancement of gene expression in Arabidopsis, *The Plant Journal 40*, 744–751.
14. Maquat, L. E. (2004) Nonsense-mediated mRNA decay: Splicing, translation and mRNP dynamics, *Nature Reviews Molecular Cell Biology 5*, 89–99.
15. Mori, M., Fujihara, N., Mise, K., and Furusawa, I. (2001) Inducible high-level mRNA amplification system by viral replicase in transgenic plants, *Plant Journal 27*, 79–86.
16. Canizares, M. C., Nicholson, L., and Lomonossoff, G. P. (2005) Use of viral vectors for vaccine production in plants, *Immunology and Cell Biology 83*, 263–270.
17. Streatfield, S. J. (2005) Oral hepatitis B vaccine candidates produced and delivered in plant material, *Immunology and Cell Biology 83*, 257–262.
18. Hefferon, K. L., and Ying, F. (2004) Expression of a vaccine protein in a plant cell line using a geminivirus-based replicon system, *Vaccine 23*, 404–410.
19. Lau, O., and Sun, S. (2009) Plant seeds as bioreactors for recombinant protein production, *Biotechnology Advances 27*, 1015–1022.
20. Desai, P., Shrivastava, N., and Padh, H. (2010) Production of heterologous proteins in plants: Strategies for optimal expression, *Biotechology Advances 28*, 427–435.
21. Gil, F., Titarenko, E., Terrada, E., Arcalis, E., and Escribano, J. M. (2006) Successful oral prime-immunization with VP60 from rabbit haemorrhagic disease virus produced in transgenic plants using different fusion strategies, *Plant Biotechnology Journal 4*, 135–143.
22. Dai, Z., Hooker, B., Anderson, D., and Thomas, S. (2000) Improved plant-based production of E1 endoglucanase using potato: expression optimization and tissue targeting, *Molecular breeding 6*, 277–285.
23. Daniell, H., Kumar, S., and Dufourmantel, N. (2005) Breakthrough in chloroplast genetic engineering of agronomically important crops, *Trends in Biotechnology 23*, 238–245.

24. Doran, P. M. (2006) Foreign protein degradation and instability in plants and plant tissue cultures, *Trends in Biotechnology 24*, 426–432.
25. Ulker, B., Allen, G. C., Thompson, W. F., Spiker, S., and Weissinger, A. K. (1999) A tobacco matrix attachment region reduces the loss of transgene expression in the progeny of transgenic tobacco plants, *Plant Journal 18*, 253–263.
26. Hood, E. E., Bailey, M. R., Beifuss, K., Magallanes-Lundback, M., Horn, M. E., Callaway, E., Drees, C., Delaney, D. E., Clough, R., and Howard, J. A. (2003) Criteria for high-level expression of a fungal laccase gene in transgenic maize, *Plant Biotechnol J 1*, 129–140.
27. Young, T. E., Giesler-Lee, J., and Gallie, D. R. (2004) Senescence-induced expression of cytokinin reverses pistil abortion during maize flower development, *Plant Journal 38*, 910–922.
28. Nuttall, J., Vine, N., Hadlington, J. L., Drake, P., Frigerio, L., and Ma, J. K. C. (2002) ER-resident chaperone interactions with recombinant antibodies in transgenic plants, *European Journal of Biochemistry 269*, 6042–6051.
29. Stoger, E., Ma, J. K. C., Fischer, R., and Christou, P. (2005) Sowing the seeds of success: pharmaceutical proteins from plants, *Current Opinion in Biotechnology 16*, 167–173.
30. Streatfield, S. J., and Howard, J. A. (2003) Plant-based vaccines, *International Journal for Parasitology 33*, 479–493.
31. Sweeney, P. (2002) A Warehouse of Ideas; Developing and Using Intellectual Property, In *Plants as Factories for Protein Production* (Hood, E., and Hward, J., Eds.), pp 181–206, Kluwer Academic Pub.
32. OECD. (1993) Safety Evaluation of Foods Derived by Modern Biotechnology; Concepts and Principles, pp 1–77, Organization for Economic Co-operation and Development, Paris, France.
33. Huson, D., Richter, D., Mitra, S., Auch, A., and Schuster, S. (2009) Methods for comparative metagenomics, *BMC Bioinformatics 10*, S12.
34. Fierer, N., Breitbart, M., Nulton, J., Salamon, P., Lozupone, C., Jones, R., Robeson, M., Edwards, R. A., Felts, B., Rayhawk, S., Knight, R., Rohwer, F., and Jackson, R. B. (2007) Metagenomic and Small-Subunit rRNA Analyses Reveal the Genetic Diversity of Bacteria, Archaea, Fungi, and Viruses in Soil, *Appl. Environ. Microbiol. 73*, 7059–7066.
35. Daniel, R. (2005) The metagenomics of soil, *Nat Rev Micro 3*, 470–478.

Chapter 26

A Novel Plant Cell Bioproduction Platform for High-Yield Secretion of Recombinant Proteins

Jianfeng Xu and Marcia J. Kieliszewski

Abstract

Plant cell suspension culture integrates the merits of whole-plant systems with those of microbial fermentation and mammalian cell culture, and has been recognized as a promising alternative biosynthetic platform for valuable proteins. However, the low protein productivity dilemma has been the bottleneck toward commercializing this technology. Here, we describe a new technology, termed hydroxyproline (*Hyp*)-*Glyco* technology, that dramatically increases the yield of *secreted* recombinant proteins from cultured plant cells by expressing them as fusions with a novel glycomodule tag comprising an Hyp-rich repetitive peptide (HypRP) backbone that is subsequently glycosylated through the Hyp residues. The extensive glycosylation of the HypRP tags greatly extends the serum half-life of small therapeutic proteins, such as interferon α2b or human growth hormone, without significantly impairing their bioactivities and the tag greatly enhances solubility.

Key words: Plant cell culture, Recombinant protein, Secreted protein, Hydroxyproline-*O*-glycosylation, Enhanced green fluorescence protein, Interferon α2

1. Introduction

Plant-based "molecular farming" is emerging as a promising approach for the production of valuable recombinant proteins with significant advantages in cost and safety over other eukaryotic expression systems (1–4). Additional benefits gained using plant cell suspension cultures include a sterile, controlled, and contained environment, hence no issues with pathogen and herbicide contaminants and the environmental concerns linked to whole-plant systems (5, 6). Thus, good manufacturing practice (GMP) procedures can be readily implemented throughout the production pipeline, alleviating a number of regulatory concerns regarding

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_26, © Springer Science+Business Media, LLC 2012

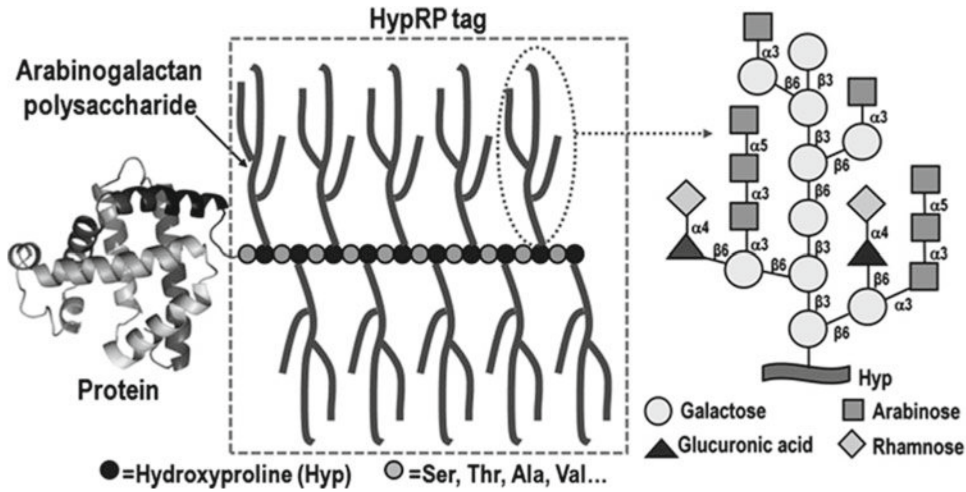


Fig. 1. Schematics of *Hyp-Glyco* technology. A recombinant protein is expressed as fusion with an HypRP tag comprising tandem Ser-Pro repeats. A typical Hyp-arabinogalactan polysaccharide (Hyp-glycan) consists of galactose, arabinose, rhamnose, and glucuronic acid.

plant-made pharmaceuticals. In addition, the procedures for protein purification are also simpler for cell cultures than for whole plants, especially when the product is secreted into the culture medium. However, low protein productivity remains a major obstacle that limits extensive commercialization of plant cell bioproduction platform (5).

We developed a novel technology, named hydroxyproline (*Hyp*)-*Glyco* technology, which significantly facilitated the secretion of the expressed proteins from plant cells, increasing the secreted protein yields as much as 1,500-fold (7, 8). The *Hyp-Glyco* technology involves the expression of recombinant proteins as fusions with a novel Hyp-rich repetitive peptide (HypRP) tag that directs extensive Hyp-*O*-glycosylation, a posttranslational modification consisting of proline hydroxylation, and subsequent Hyp-*O*-glycosylation, a modification that is unique to plants and green algae (9–15) (Fig. 1).

Hyp-*O*-glycosylation takes two forms in plants and is directed by the arrangement of Hyp residues. Thus, blocks of contiguous Hyp, such as the canonical Ser-Hyp_4 repeat that typifies network forming hydroxyproline-rich glycoproteins, are highly arabinosylated with small (1–5 residues) linear chains of L-arabinose. In contrast, clustered, noncontiguous Hyp residues, like the X-Hyp-X-Hyp repeats that typify the hydroxyproline-rich glycoproteins in many plant gums and exudates, are preferred sites of large (15–150 residues), acidic, and highly branched heteropolysaccharides designated arabinogalactan polysaccharides (Hyp-glycans) (11, 15). Most commonly, the HypRP tags we exploit are those that direct arabinogalactan polysaccharide addition to Hyp: tandem repeats of “Ser-Hyp” or “Ala-Hyp” (10, 15). Here, we describe the expression in tobacco cells of a reporter protein and a therapeutic

protein as fusions with an "Ser-Hyp"-based HypRP tag, specifically the enhanced green fluorescence protein (EGFP) and interferon α2b (IFNα2). In both cases, high-secreted protein yields were detected. Furthermore, the HypRP tags favorably impacted the stability and pharmacokinetic behavior of IFNα2.

2. Materials

2.1. Plasmid Construction

1. cDNAs: The *EGFP* gene isolated from plasmid *pEGFP* (Clontech); the *IFNα2* gene isolated from plasmid *pGL2BINF* (ATCC No. 53371); the tobacco extension signal sequence, *SStob*, described earlier (16).
2. Cloning vector: *pUC18*.
3. Binary vector: *pBI121* containing the *CaMV35S* promoter (Clontech) (see Note 1).
4. DNA Miniprep Kit.
5. DNA gel extraction kit.
6. QuickChange Mutagenesis kit (Stratagene).
7. MicroSpin® S-200 HR columns (GE Healthcare).
8. PCR thermal cycler.
9. PCR 2× Master Mix (New England BioLabs).
10. Thin-wall PCR tubes.
11. Luria-Bertani (LB) medium. Add 1.5% (w/v) agar for solid medium.
12. Custom-synthesized oligonucleotides (Integrated DNA Technologies).
13. Restriction enzymes: *Nco*I, *Xma*I, *Bsr*GI, *Bam*HI, and *Sac*I.
14. T4 DNA Ligase: 2–3 U/μL (Promega).
15. Antibiotics: Kanamycin and ampicillin.
16. 1% agarose gel containing ethidium bromide (0.5 μg/mL).
17. Electrophoresis apparatus.
18. Microcentrifuge.
19. TAE buffer 10× stock: 40 mM Tris–HCl, 1 mM EDTA, pH 8.0, 11.42 mL/L glacial acetic acid.

2.2. Stable Transformation of Tobacco Bright Yellow-2 Cells and Cell Suspension Culture

1. Tobacco Bright Yellow-2 (BY-2) cells.
2. *Agrobacterium tumefaciens* strain LBA4404.
3. LB-S medium: LB medium containing 100 mg/L streptomycin.

4. LB-SK medium: LB medium containing 100 mg/L streptomycin and 50 mg/L kanamycin; add 1.5% (w/v) agar for making LB-SK plates.
5. Antibiotics: Streptomycin, kanamycin, timentin (Phytotechnology Laboratories) (see Note 2). All of these antibiotics are filter sterilized and added to the plant cell culture medium as needed.
6. $CaCl_2$ solution: 20 mM $CaCl_2$.
7. NT medium: Murashige and Skoog (MS) basal salt (17), 30 mg/L sucrose, 6% (w/v) KH_2PO_4, 100 mg/L *myo*-inositol, 1.0 mg/L thiamine-HCl, 0.2 mg/L 2,4-dichlorophenoxyacetic acid (2,4-D). The medium pH is adjusted to 5.2 before autoclaving; for solid medium, add 2 g/L Phytagel (Sigma) before autoclaving.
8. Schenk-Hildebrandt (SH) medium: SH basal medium (18), 34 g/L sucrose, 2.1 mg/L *p*-chlorophenoxy acetic acid, 0.44 mg/L 2,4-D, 0.1 mg/L kinetin, pH 5.8; add 2 g/L Phytagel for solid medium.
9. NTK medium: NT medium, 200 mg/L kanamycin.
10. NTTK medium: NT medium, 400 mg/L timentin, 200 mg/L kanamycin.
11. SHK medium: SH medium, 200 mg/L kanamycin.
12. Erlenmeyer flasks (250 and 500 mL).
13. Petri dishes (100 × 15 mm).
14. Wide-mouth pipette (10 mL).
15. 37°C water bath.
16. Gyrotary shakers.
17. Spectrophotometer.
18. Centrifuge equipped with a swinging-bucket rotor that can hold 50-mL conical centrifuge tubes.
19. Laminar flow hood.

2.3. SDS-Polyacrylamide Gel Electrophoresis, Western Blot, and ELISA

1. Nitrocellulose membrane.
2. SDS-polyacrylamide gel electrophoresis (PAGE) running buffer (5×): 125 mM Tris–HCl, pH 6.8, 960 mM glycine, 0.5% (w/v) SDS. Store at room temperature.
3. Protein loading buffer (3×): 187.5 mM Tris–HCl, pH 6.8, 6% (w/v) SDS, 30% glycerol, 0.03% (w/v) bromophenol blue, 125 mM dithiothreitol (DTT).
4. Vertical mini gel electrophoresis apparatus (Bio-Rad).
5. Mini Trans-Blot electrophoretic transfer cell (Bio-Rad).
6. Polyacrylamide gel: Precast 4–15% and 12% Tris–HCl gel (Bio-Rad).

7. Transfer buffer: 20 mM Tris–HCl, pH 8.5, 150 mM glycine, 20% methanol.
8. Antibodies: Rabbit Anti-EGFP antibody (Open BioSystems); rabbit anti-IFNα2 antibody (Fitzgerald); goat anti-rabbit IgG alkaline phosphatase conjugate (Sigma).
9. 5-bromo-4-chloro-3′-indolyphosphate *p*-toluidine salt (BCIP) and nitro-blue tetrazolium chloride (NBT) substrate (Sigma) (see Note 3).
10. Tris-buffered saline with Tween (TBS-T) buffer (10×): 1.5 M NaCl, 20 mM KCl, 250 mM Tris–HCl, pH 7.4, 1% Tween-20.
11. Blocking buffer: 5% (w/v) BSA in TBS-T buffer.
12. Prestained molecular weight markers.
13. EGFP ELISA assay kit (Open Biosystems), IFNα2 ELISA assay kit (PBL InterferonSource).
14. Fluorometer: Modulus™ system (Turner Biosystems).
15. EGFP standard (BioVision), IFNα2 standard (Fitzgerald).

2.4. Recovery and Purification of Fusion Glycoproteins

1. Chromatographic column (16×700 mm) packed with Phenyl-Sepharose 6 Fast Flow resin (GE Healthcare).
2. Gilson 305 HPLC Piston Pump.
3. Spectra/Chrom model 280UV detector (Spectrum Laboratories).
4. Gilson FC 204 Fraction Collector.
5. Amicon stirred cell, 200 mL (Millipore).
6. Centricon YM10 membrane (Millipore).
7. Starting buffer: 20 mM PBS, pH 7.2, 2.0 M NaCl.
8. Washing buffer: 20 mM PBS, pH 7.2, 1.0 M NaCl.
9. Elution buffer: 20 mM PBS, pH 7.2, 0.2 M NaCl.
10. HPLC system: Agilent 1100 HPLC.
11. Hamilton polymeric reversed phase-1 (PRP-1) analytical column (4.1×150 mm, Hamilton).
12. HPLC solution A: 0.1% trifluoroacetic acid, solution B: 0.1% trifluoroacetic acid, 80% aqueous acetonitrile.

3. Methods

"Designer" HypRP tags comprising repeats of the dipeptide (Ser-Hyp) determines the number of glycans attached to the HypRP tag in the fusion protein, with each Hyp residue being glycosylated. Two different sizes of the arabinogalactosylated

Ser-Hyp-based HypRP tag are exploited: one with ten repeats of the "Ser-Hyp" motif, designated $(SO)_{10}$ (O denotes Hyp), and the other with 20 repeats, designated $(SO)_{20}$ to produce the fusion glycoproteins EGFP-$(SO)_{10}$, IFNα2-$(SO)_{10}$, and IFNα2-$(SO)_{20}$ (see Note 4). Each Hyp (O) residue in the $(SO)_{10}$ and $(SO)_{20}$ tags is glycosylated with one Hyp-glycan; thus, 10 and 20 Hyp-glycans are linked to $(SO)_{10}$ and $(SO)_{20}$ tag, respectively. The methods described here outline (1) construction of the plasmids for *Agrobacterium*-mediated plant cell transformation; (2) generation of transgenic tobacco BY-2 cells expressing three fusion glycoproteins; and (3) how to purify and characterize the fusion proteins. Please note that while $(SO)_{10}$ and $(SO)_{20}$ refer to the protein sequence of the tags expressed in plant cells, italicized $(SP)_{10}$ and $(SP)_{20}$ (P denotes Pro) refers to the name of *genes*/DNA sequence encoding these two peptide motifs. Hydroxylation of Pro to Hyp, presumably in the endoplasmic reticulum (ER), is a posttranslational modification; hence, there is no codon for Hyp.

3.1. Plasmid Construction

The *EGFP-*$(SP)_{10}$, *IFNα2-*$(SP)_{10}$, and *IFNα2-*$(SP)_{20}$ fusion genes preceded by DNA encoding a tobacco extensin signal peptide SS^{tob} (16) are constructed first in a small cloning vector *pUC18*. The $(SP)_{10}$ fragment is created directly from two custom-synthesized oligonucleotides (Fig. 2). The $(SP)_{20}$ fragment is generated by splicing two $(SP)_{10}$ sequences using typical molecular cloning methods. After the final constructs are made, the DNA sequences should be confirmed by DNA sequencing, and then subcloned into the binary vector *pBI121* containing the *NPT II* gene for selection (kanamycin-resistant cell colonies).

3.1.1. Construction of the $(SP)_{10}$ Synthetic Gene Fragment

1. Design two partially overlapping oligonucleotides (both sense and antisense sequences) encoding $(SP)_{10}$ peptide (Fig. 2) (see Note 5). Incorporate restriction sites, *Nco*I and *Brs*GI, at the 5′- and 3′-ends of the fragment, respectively, for subcloning purpose. Incorporate additional three to four nucleotide bases at 5′- and 3′-ends of the DNA sequence to facilitate restriction enzyme digestion.
2. After receiving the custom-synthesized oligonucleotides, dissolve them in nuclease-free water to a concentration of 0.1 nmol/μL.
3. Transfer 0.5 μL of each of the oligonucleotide solutions to a PCR tube. Add 12.5 μL of PCR Master Mix (2×) and 12 μL of nuclease-free water to bring the final volume to 25 μL.
4. Place the PCR tube in a PCR thermal cycler. Set up one thermal cycle program as follows: 95°C for 5 min; 55°C for 1 min; 72°C for 30 min. These conditions are suitable to make the double-stranded *(SP)10* fragment (see Note 6).

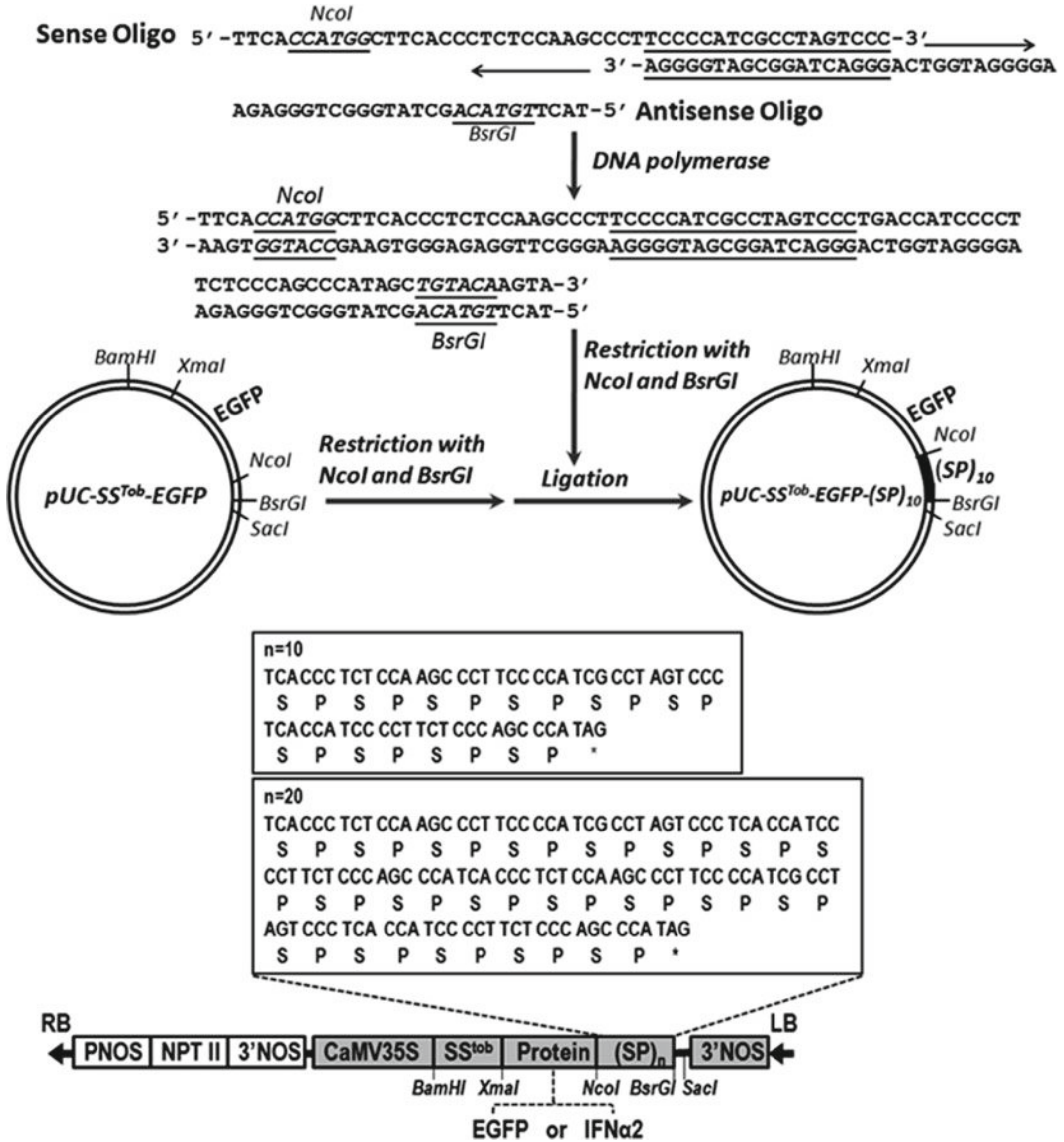


Fig. 2. Schematics of construction of synthetic *(SP)*$_{10}$ fragment (*top*) and plant expression vector *pBI121* containing *EGFP* or *IFN*α*2* fusion gene under the control of the *CaMV35S* promoter (*bottom*). *CaMV35S* 35S cauliflower mosaic virus promoter, *SS*^{tob} signal sequence of tobacco extensin for ER targeting, *(SP)*$_n$ Ser-Pro repeats, $n = 10, 20$, *PNOS* nopaline synthase promoter, *NPT II* neomycin phosphotransferase II gene conferring resistance to kanamycin, *3′NOS* nopaline synthase terminator.

5. Change the buffer with a MicroSpin® S-200 HR column or other PCR cleaning kits before subject to restriction digestion with *Nco*I and *Bsr*GI. The *(SP)10* fragment is now ready for subcloning.

3.1.2. Cloning

The construction of the *SS^tob^-EGFP-(SP)$_{10}$*, *SS^tob^-IFNα2-(SP)$_{10}$*, and *SS^tob^-IFNα2-(SP)$_{20}$* fusion genes in *pUC18* is based on the plasmid *pUC-SS^tob^-EGFP* generated earlier in the Kieliszewski Laboratory (15). Standard molecular cloning procedures are followed, which are not described in detail here due to space limitations. Briefly, the *(SP)$_{10}$* fragment created above is inserted as an *Nco*I and *Bsr*GI fragment immediately at 3′ of the *EGFP* gene in *pUC-SS^tob^-EGFP* generating the plasmid *pUC-SS^tob^-EGFP-(SP)$_{10}$*. The *IFNα2* gene is amplified by PCR from plasmid *pGL2BINF*, and then subcloned as an *Xma*I and *Nco*I fragment into *pUC-SS^tob^-EGFP-(SP)$_{10}$* replacing *EGFP* and generating *pUC-SS^tob^-IFNα2-(SP)$_{10}$*. The *IFNα2-(SP)$_{10}$* fusion is then amplified by PCR and subcloned into *pUC-SS^tob^-IFNα2-(SP)$_{10}$* at the *Xma*I and *Nco*I sites, replacing the *IFNα2* fragment only, to generate *pUC-SS^tob^-IFNα2-(SP)$_{10}$ -(SP)$_{10}$* or *pUC-SS^tob^-IFNα2-(SP)$_{20}$* (7). The extra nucleotides introduced into these plasmids for cloning purpose can be removed by site-directed mutagenesis using the QuickChange Mutagenesis kit. All the constructs are then subcloned into the plant expression binary vector *pBI121* as *Bam*HI-*Sac*I fragments (Fig. 2).

3.2. Generation of Transformed Tobacco BY-2 Cells and Cell Suspension Culture

Plant expression *pBI121* vectors are then transformed into *A. tumefaciens* strain LBA4404 by the freeze–thaw method. The stable nuclear transformation of tobacco BY-2 cells is then carried out by the *Agrobacterium*-mediated method (19). After the transformed cells are obtained, elite cell lines that grow fast and secret high yields of recombinant proteins into culture medium are screened by dot blot assays.

3.2.1. Transformation of A. tumefaciens with Binary Vector

In the following procedures, competent *A. tumefaciens* cells are made first and then transformed with the *pBI121* vector.

1. Grow *A. tumefaciens* in 5 mL of LB-S medium (see Note 7) at 28°C overnight.
2. Add 2 mL of the overnight culture to 50 mL of LB-S medium in a 250-mL flask and shake vigorously (250 rpm) at 28°C until the culture grows to an absorbance (at 600 nm) of 0.5–1.0.
3. Chill the culture on ice, and then centrifuge the cell suspension at 3,000 × *g* at 4°C for 5 min. Decant the supernatant.
4. Resuspend the cells in 1 mL of ice-cold $CaCl_2$ solution. Dispense 0.1-mL aliquots into prechilled Eppendorf tubes. Freeze immediately in liquid nitrogen and store the Eppendorf tubes in a −80°C freezer. This makes the competent *A. tumefaciens* cells.
5. Add ~1 μg of *pBI121* plasmid into one Eppendorf tube containing the frozen competent cells. Thaw the cells quickly by holding in fingers, and mix the DNA with the cells by gently

tapping the tube. Freeze the mixture immediately in liquid nitrogen.

6. Thaw the cells by incubating the tube in a 37°C water bath for 5 min.
7. Add 1 mL of LB medium to the tube and incubate at 28°C for 2–4 h with shaking.
8. Centrifuge the tube at full speed in a microcentrifuge for 30 s. Decant the supernatant and resuspend the cell pellets in 0.1 mL LB medium.
9. Spread the cells onto an LB-SK plate. Incubate the plate at 28°C. Transformed *A. tumefaciens* should grow within 2–3 days.

3.2.2. Stable Transformation of Tobacco BY-2 Cells

All the procedures involving the BY-2 cells should be carried out in a laminar flow hood.

1. At 3–4 days before transformation, subculture a 1-week-old, wild-type BY-2 cell in 500-mL Erlenmeyer flasks containing 200 mL NT medium. Culture the cells in the dark at room temperature (24–26°C) on a gyratory shaker at 90–100 rpm (see Note 8).
2. At 3–4 days before transformation, streak the *A. tumefaciens* line LBA4404 now transformed with the *pBI121*-based plasmid on an LB-SK plate. Incubate at 28°C for 24–48 h in an incubator.
3. At 2 days before BY-2 cell transformation, pick a single colony of the transformed *A. tumefaciens* from the LB-SK plate and culture in 5 mL of LB-SK medium at 28°C for 24 h on a shaker at 250 rpm.
4. Measure the absorbance of the transformed *A. tumefaciens* culture at 600 nm using a spectrophotometer. The optimum absorbance is between 0.5 and 0.6 (see Note 9).
5. Let the cells of the 3–4 days' BY-2 culture settle to the bottom of the tube for 5–10 min and then decant 1/2 to 2/3 vol. of the spent media. This concentrates the suspension cells (see Note 10). Transfer 4 mL of the suspension cells into each Petri dish (100 × 15 mm) with a wide-mouth 10-mL pipette.
6. Add 10–20 μL of *A. tumefaciens* to each Petri dish containing BY-2. Mix the bacteria and cells by swirling the Petri dish gently. Wrap each Petri dish with parafilm and incubate the cocultures at 28°C in the dark for 2–3 days (see Note 11).
7. Wash the cells in a sterile, disposable 50-mL conical centrifuge tube as follows: (a) Transfer the cells from each Petri dish transformation to a centrifuge tube (one tube per Petri dish of cells) and bring the total volume of the suspension to 45 mL

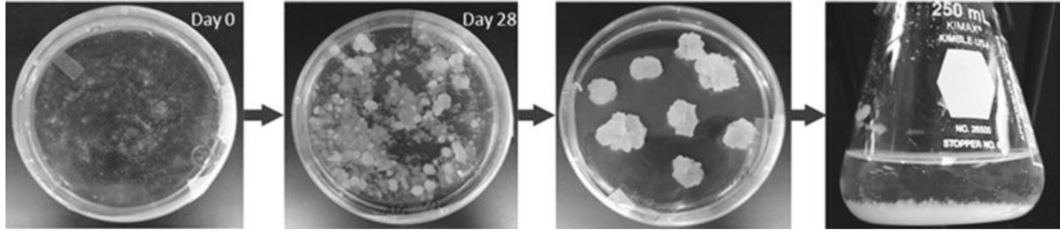


Fig. 3. Stable transformation of BY-2 cells and establishment of cell suspension culture. The cocultured cells are selected on NTTK plate (*first panel from left*). Transformed cell colonies (bumps) appear in 3–4 weeks (*second panel*). Individual cell colonies are picked up and transferred onto new NTTK plate; each individual colony grows to be a mass of calli (*third panel*). Individual callus is then transformed into liquid culture (*fourth panel*).

using NT medium. (b) Shake or invert the tubes to mix. (c) Centrifuge at $200 \times g$ (or 3,000 rpm) using a swinging-bucket rotor for 5 min. (d) Discard the supernatant and resuspend the cell pellet with 45 mL of NT medium. Repeat the steps (b)–(d) twice more for a total of three washes. The last wash is carried out with NTT medium. DO NOT resuspend the cell pellet after the last wash.

8. Spread the pelleted cells onto NTTK plates (100×15 mm) with a disinfected spatula. Wrap the plates in parafilm and culture at room temperature (23–25°C) in the dark. In 3–4 weeks, transformed calli appear as "bumps" (Fig. 3). These are masses of rapidly growing transformed cells in contrast to the background "lawn" of nongrowing or slow-growing cells. Each individual mass of cells (cell colony) is considered a line.
9. Pick up 30–40 individual colonies and transfer them onto fresh NTTK plates (Fig. 3). Subculture the colonies every 3–4 weeks for at least two generations before transferring the calli into liquid cultures. This prolonged treatment with timentin in NTTK plates removes all the *A. tumefaciens.* If no *A. tumefaciens* growth is present in the cells, NTK media (without timentin) can be used for subsequent subcultures (see Note 12).

3.2.3. Cell Suspension Culture for Secreting Recombinant Fusion Glycoproteins

After several transformed cell colonies are obtained for each construct, they are screened in liquid culture for elite cell lines that grow fast and secret high yields of recombinant protein into culture medium (see Note 13). The procedures for cell suspension cultures and cell line screening using dot blot assays are presented below.

1. Transfer 20–25 individual colonies (calli) into 60 mL of SHK medium contained in 250-mL Erlenmeyer flasks (see Note 14). Culture the cells under light at room temperature (24–26°C) on a gyratory shaker at 90–100 rpm.
2. After 2–3 weeks, take small volumes (0.5–1 mL) of culture medium from each flask in the laminar flow hood. These samples are screened for target protein with dot blot assay (see Note 15).

3. Heat the medium samples at 95°C for 10 min to deactivate the endogenous alkaline phosphatase.
4. Blot 1.0 μL each of the samples (including the medium from wild-type BY-2 cell culture) and 10 ng of protein standard onto a nitrocellulose membrane.
5. Proceed to antibody binding, membrane washing, and color development with the same procedure as the western blot shown below in Subheading 3.3.1.
6. After the dots are visible on the membrane, identify three to four cell cultures that result in the dots with the highest density. Maintain these cell lines and discard the rest of cell cultures.
7. Subculture the chosen cell lines every 8–10 days by mixing cell suspensions with fresh SHK medium at a 1:5 ratio (v/v). The cell suspensions can be maintained for several years (see Note 16).

3.3. Characterization of the Fusion Glycoproteins

One of the distinctive features of the *Hyp-Glyco* technology is the reliable high-yield secretion of recombinant proteins into culture medium. The cell culture media can be used directly for Western blots and ELISA assays to identify the secreted fusion glycoproteins and quantify their yields. Secreted proteins are particularly attractive for downstream protein purification because they circumvent several operations during the purification process that arise when cell must be disrupted and the product isolated from the intracellular debris. Additional advantage of purifying proteins from plant cell cultures resides in the simplicity of the starting medium which lacks proteins and other contaminants that complicate purification of the secreted product.

3.3.1. SDS-PAGE and Western Blotting

1. Collect the cell culture media by filtration on a coarse, sintered funnel and spin down to remove all particulate matter. The cell culture media can be used directly for a Western blot assay.
2. Combine 10 μL of medium with 5 μL of protein loading buffer (3×) and boil for 5 min.
3. Load 15 μL of each sample on a Tris–HCl gel. Include one well for prestained molecular weight markers.
4. Run the samples at 150 V on a vertical mini gel electrophoresis apparatus following manufacturer's procedures.
5. Transfer the protein to a nitrocellulose membrane at 100 V and 350 mA for 1 h using a Mini Trans-Blot electrophoretic transfer cell following manufacturer's procedures.
6. Block the membrane in blocking buffer for 1 h.
7. Incubate the membrane in TBS-T buffer containing primary anti-EGFP antibody or anti-IFNα2 antibody for 1 h (see Note 17). Wash three times with TBS-T buffer.

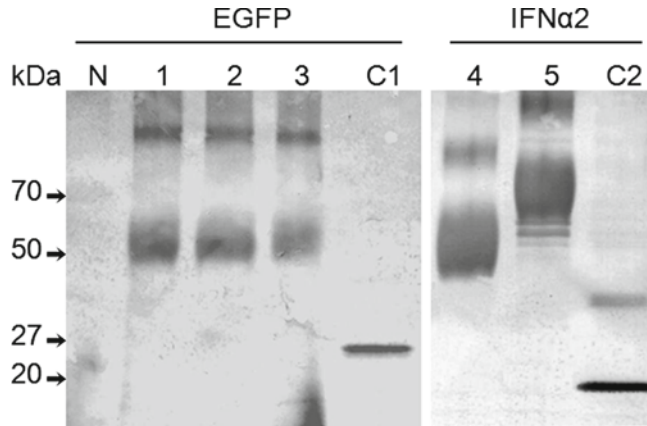


Fig. 4. Western blotting detection of recombinant EGFP-$(SO)_{10}$, IFNα2-$(SO)_{10}$, and IFNα2-$(SO)_{20}$ secreted into cell culture media using anti-EGFP and anti-IFNα2 antibodies. N: nontransformed BY-2 cell culture medium; 1, 2, and 3: EGFP-$(SO)_{10}$ cell culture media; 4: IFNα2-$(SO)_{10}$ cell culture medium; 5: IFNα2-$(SO)_{20}$ cell culture medium; C1: EGFP standard; C2: IFNα2 standard. EGFP and its fusion glycoproteins are separated on a 12% Tris–HCl gel; IFNα2 and its fusion glycoproteins are separated on a 4–15% Tris–HCl gel. The fusion glycoproteins are migrated as broad bands, probably resulting from microheterogeneity in the glycans and inefficient SDS binding by glycoproteins.

8. Incubate the membrane in TBS-T buffer containing secondary antibody goat anti-rabbit IgG conjugated with alkaline phosphatase (1:1,000 dilution). Wash three times with TBS-T buffer.
9. Detect the target proteins by adding BCIP/NBT substrate until the reactive bands are visible (see Note 3). Decant the substrate and wash the membrane with distilled water.
10. Dry the membrane in the air.

Western blot assays indicate that the molecular sizes of EGFP and IFNα2 are significantly increased when fused with the $(SO)_{10}$ or $(SO)_{20}$ tag (Fig. 4).

3.3.2. Quantification of Secreted EGFP and IFNα2 Yields

The secreted EGFP and IFNα2 yields are determined by sandwich ELISAs with an EGFP and an IFNα2 ELISA kit, respectively, following manufacturer's procedures. The culture media harvested from 10- to 12-day cultures can be used directly for ELISA assay. The highest secreted yields of EGFP-$(SO)_{10}$ that we measured were 55 mg/L EGFP equivalents, in contrast to 0.5 mg/L of EGFP equivalents determined when the protein was expressed without an $(SO)_{10}$ tag. The secreted yields of IFNα2-$(SO)_{10}$ and IFNα2-$(SO)_{20}$ were 30 and 28 mg/L of IFNα2 equivalents (7), respectively, in contrast to 0.02 mg/L of IFNα2 equivalents obtained when the protein was expressed without the $(SO)_{10 \text{ or } 20}$ fusion.

3.3.3. Recovery and Purification of Fusion Glycoproteins from the Culture Media

The secreted fusion glycoproteins can be recovered easily with a hydrophobic-interaction chromatography (HIC) when the therapeutic protein portion of the fusion has hydrophobic regions (see Note 18). This allows separation of the fusion glycoprotein from the hydrophilic endogenous Hyp-containing glycoproteins also made by BY-2 cells during growth. Final purification of protein can often be achieved using reverse-phased (RP) HPLC.

1. Harvest the cell culture medium (1 L) after 8–10 day of culture by filtration on a coarse, sintered funnel. Add NaCl to the harvested medium to a final concentration of 2 M and sit the medium at 4°C for 3 h or longer up to overnight.
2. Centrifuge the salt-adjusted medium at 25,000 × *g* for 20 min at 4°C to remove any insoluble materials, like pectins (see Note 19).
3. Equilibrate the HIC column with at least 10-column volumes of starting buffer, monitoring the effluent absorbance at 220 nm.
4. Load the clarified culture medium onto the HIC column at a flow rate of 1 mL/min. The effluent absorbance increases rapidly as nonbound material elutes from the column.
5. Wash the column with washing buffer until the baseline absorbance returns to equilibration levels.
6. Elute the fusion glycoproteins with elution buffer. Collect the eluant corresponding to the peak (10–15 mL) (see Note 20).
7. Wash the HIC column with distilled water and then store the HIC column in high salt under the recommended conditions provided by the manufacturer.
8. Concentrate the collected eluant to 1–2 mL via ultrafiltration at 4°C in the Amicon stirred cell with a Centricon YM10 membrane. The concentrated sample is ready for assay of the in vitro bioactivity and in vivo circulating half-life of the IFNα2 fusion glycoproteins. The sample can be further purified with RP-HPLC for characterization of the structure and glycosylation of the fusion glycoproteins.
9. Equilibrate the PRP-1 analytical column of the HPLC system with solution A for at least 15 min at a flow rate of 0.5 mL/min.
10. Spin down the sample to remove any particulate. Inject 0.1–0.2 mL of the sample into the PRP-1 column and keep running the buffer A at a flow rate of 0.5 mL/min for 3 min.
11. Start the gradient of 0–30% solution B for 10 min, followed by the gradient of 30–80% B for 90 min. Monitor the eluant absorbance at 220 nm.
12. Collect the major peak eluted from the PRP-1 column (Fig. 5) and freeze dry the sample. The resulting white, fluffy fusion

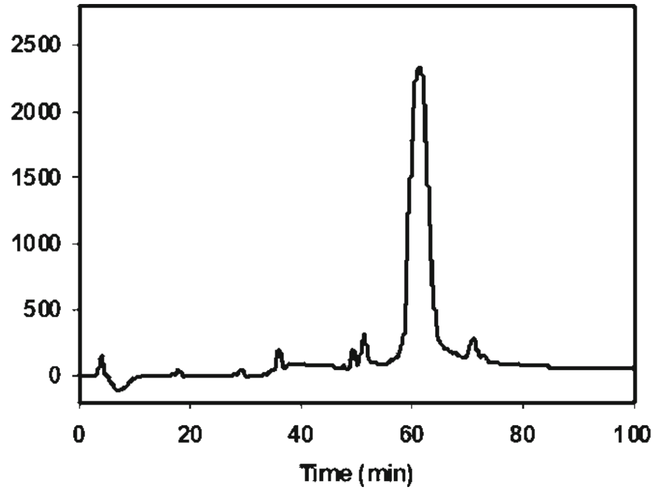


Fig. 5. Reverse-phased HPLC separation of IFNα2-$(SO)_{10}$ eluted from HIC column. The sample is concentrated tenfold with ultrafiltration before being injected into HPLC. Judging from the HPLC chromatogram, single step of HIC chromatography can yield protein preparation that is ~85% homogeneous.

glycoproteins can be used for sugar composition, sugar linkage analyses, glycosylation profiles, and amino acid composition analysis or peptide mapping (7, 9, 13–15).

3.3.4. Bioactivity and/or Clinical Effectiveness Assay of Fusion Glycoproteins

EGFP expressed as an $(SO)_{10}$ fusion is a fully functional fluorescent protein judged by fluorometry. The greatest fluorescent intensity of the culture media containing EGFP-$(SO)_{10}$ measures 95,000 RFU on the Modulus™ system, which corresponds to the fluorescent intensity of 40 mg/L of EGFP standard (see Note 21). The in vitro antiviral activity of IFNα2 fusion glycoproteins can be assayed using Madin–Darby Bovine Kidney (MDBK) cells challenged by *vesicular stomatitis virus* (VSV) (20). The specific antiviral activities of IFNα2-$(SO)_{10}$ and IFNα2-$(SO)_{20}$ are measured as 0.22 and 0.20 international unit (IU)/pg, respectively, representing 91 and 87% that of the IFNα2 standard (0.23 IU/pg). Here, the IU is determined against the NIH reference standard for human IFNα2b Gxa01-901-535. Following intravenous injection in mice, the antiviral half-life of IFNα2-$(SO)_{10}$ and IFNα2-$(SO)_{20}$ is prolonged by 7- and 13-fold, respectively (7).

4. Notes

1. The *pBI121* with a constitutive CaMV35S promoter is a relatively old binary vector for plant cell transformation. New binary vectors with a stronger constitutive promoter, e.g., double-enhanced CaMV35S promoter and/or a translational enhancer sequence, such as a 5′-translational elements from tobacco etch

virus (TEV), are expected to generate higher expression in plant cells. Alternatively, a binary vector with an inducible promoter can also be used.

2. Although other antibiotics, such as cefotaxime and carbenicillin, are generally used to kill *Agrobacterium* during plant cell transformation, timentin shows the most effective inhibitory effects for *Agrobacterium*. Timentin can be purchased from Phytotechnology Laboratories (http://www.phytotechlab.com/) without a prescription.
3. Alternatively, chemiluminescence reagents, such as Enhanced Chemiluminescent (ECL) reagents from Kirkegaard and Perry, can be used to detect protein bands in Western blot assay. In this case, X-ray film, a darkroom, and film developer are needed.
4. HypRP tags comprising many other clustered, noncontiguous Hyp, such as tandem repeats of Ala-Hyp and Thr-Hyp motifs, can be used. However, considering the relatively high prevalence of the Ala-Hyp motif in plant HRGPs and the possibility that non-Hyp-*O*-glycosylation on "Ser" and "Thr" residues might occur, an "Ala"-based HypRP tag design, $(AO)_n$, is recommended.
5. Codon bias should be considered in designing the oligonucleotides. In dicot plants such as tobacco, genetic codes ending with A or T (on third position) are preferred. In addition, varying genetic codes encoding the same amino acid across the DNA sequence are recommended to use. Purification with PAGE or HPLC is recommended for the custom-synthesized DNA oligonucleotides greater than 60 bases in length to improve oligonucleotide performance and reduce the possibility of sequence error.
6. An alternative approach to make $(SP)_{10}$ fragment is to use the Klenow fragment. Take 0.5 μL each of the oligonucleotides to an Eppendorf tube; add 2.5 μL of Klenow buffer (10×) and 19.5 μL of water; heat the mixture at 95°C for 5 min and cool down to room temperature; add 1.0 μL of 10 mM dNTP and 1 μL of Klenow (5 U/μL); incubate the tube at 37°C for 30 min.
7. YEP medium consisting of 10 g/L bacto-peptone, 10 g/L yeast extract, and 5 g/L NaCl other than LB medium is usually used by others to grow *A. tumefaciens*. It is important to observe the formation of a white precipitate during the culture of *A. tumefaciens* LBA4404.
8. The viability of the wild-type BY-2 cells is critical for transformation. If the cells are overgrown, they will form clumps leading to significantly impaired transformation. Cells can also be cultured under light as no significant reduction in transformation has been observed as a result.

9. If the culture has overgrown, dilute with additional LB-SK medium, allow growth for at least 1 h, and check the absorbance again. Periodically check the absorbance until the culture has reached optimum density.
10. This step is not critical. Unconcentrated BY-2 cell suspension can be used directly for transformation, but much fewer transformed cell colonies may be obtained later.
11. The optimal incubation time is between 2 and 3 days. Incubation for less than 2 days results in low transformation efficiency, though the cocultured cells remain robust and grow fast after being spread on selection plates. Sometimes, it is hard to identify the transformed cell colonies from the background "lawn" of untransformed cells. In contrast, incubation for more than 3 days leads to higher transformation efficiency, but this largely scarifies the cells' viability and actually causes most of the cells to die. It takes much longer time for the cells to grow on the selection plates and spread cells are easily contaminated by the *Agrobacterium*.
12. If *A. tumefaciens* growth reappears, return cultures to NTTK plates until bacterial growth is eliminated. After elite cell lines are screened and free of *A. tumefaciens*, they can be maintained on NTK solid medium; selection pressure should be maintained all the time.
13. It is important to screen the elite cells lines. In *Agrobacterium*-mediated cell transformation, T-DNA contained on binary vector is introduced into the cell genome randomly; therefore, protein production varies dramatically among individual cell lines. This variation in production is resulted from the chromosomal position of the inserted DNA, difference in gene copy number, and the occurrence of gene silencing.
14. Suspension cells should be grown in SHK medium in order to harvest high-yield secreted fusion glycoproteins. For unknown reason so far, the cells grown in NTK medium do not secrete high-yield fusion glycoproteins into culture medium.
15. For the cell suspension cultures expressing EGFP-$(SO)_{10}$, bright green fluorescence of the recombinant EGFP secreted into media can be viewed by eye. In this case, the dot blot assay is not necessary to screen the cell line.
16. When the cell suspension is established, the cells grow much faster (with a growth cycle of 8–10 days) than those just transferred from solid medium to liquid culture. In addition, the screened elite cell lines should also be maintained on solid SHK or NTK medium. Subculture can be performed every 3–4 weeks.
17. The primary antibody can be saved for subsequent experiments by addition of 0.02% final concentration sodium azide (caution: azide is highly toxic) and storage at 4°C.

18. Except for Phenyl-Sepharose 6 Fast Flow resin used here, Blue Sepharose 6 Fast Flow (GE Healthcare) is regarded as the ideal adsorbent for capturing interferon protein at both laboratory and bioprocess scales (21). However, when the Blue Sepharose 6 column is used for the recovery of the proteins from plant cell culture medium, a gel-like substance (probably pectin) accumulates at the top of the column even if the culture medium is clarified by centrifugation before being loaded onto the column. The gels make the chromatographic process inefficient or even fail.
19. Large amounts (40–50 g fresh weight/L) of precipitates are expected to form in the cell culture medium after 2 M sodium chloride is added. These precipitates consist of pectin and other polysaccharides secreted by the cultured cells which form gel-like residues on the Blue Sepharose column or other chromatographic columns if they are not removed.
20. Stepwise wash/elution procedure is used here to save time and buffer. Alternatively, fusion glycoproteins can be eluted with a gradient of 0–100% 20 mM PBS, pH 7.2, in 10–20-column volumes after column washing.
21. This secreted yield of EGFP-$(SO)_{10}$ (40 mg/L) measured by fluorescence intensity is smaller than that determined by ELISA (55 mg/L) due to many reasons, such as proteolytic degradation of EGFP in medium, fluorescence quenching, denature of protein, etc.

Acknowledgments

This work is supported by grants from the US Department of Agriculture (2002-06201-11919), National Science Foundation (IOS-0622428), Herman Frasch Foundation (526-HF02), as well as Arkansas NSF-EPSCoR funds and the Arkansas Biosciences Institute.

References

1. Ma, J. K. C., Drake, P. M. W., and Christou, P. (2003) The production of recombinant pharmaceutical proteins in plants, *Nature Reviews Genetics 4*, 794–805.
2. Fischer, R., Stoger, E., Schillberg, S., et al., (2004) Plant-based production of biopharmaceuticals, *Current Opinion in Plant Biology 7*, 152–158.
3. Streatfield, S. J. (2007) Approaches to achieve high-level heterologous protein production in plants, *Plant Biotechnology Journal 5*, 2–15.
4. Doran, P. M. (2000) Foreign protein production in plant tissue cultures, *Current Opinion in Biotechnology 11*, 199–204.
5. Hellwig, S., Drossard, J., Twyman, R. M., et al., (2004) Plant cell cultures for the production of recombinant proteins, *Nature Biotechnology 22*, 1415–1422.
6. Georgiev, M. I., Weber, J., and Maciuk, A. (2009) Bioprocessing of plant cell cultures for mass production of targeted compounds, *Applied Microbiology and Biotechnology 83*, 809–823.

7. Xu, J. F., Tan, L., Goodrum, K. J., et al., (2007) High-yields and extended serum half-life of human interferon alpha 2b expressed in tobacco cells as Arabinogalactan-protein fusions, *Biotechnology and Bioengineering 97*, 997–1008.
8. Xu, J. F., Okada, S., Tan, L., et al., (2010) Human growth hormone expressed in tobacco cells as an arabinogalactan-protein fusion glycoprotein has a prolonged serum life, *Transgenic Res. In press*, doi:10.1007/s11248-11010-19367-11248.
9. Tan, L., Leykam, J. F., and Kieliszewski, M. J. (2003) Glycosylation motifs that direct arabinogalactan addition to arabinogalactan-proteins, *Plant Physiology 132*, 1362–1369.
10. Kieliszewski, M. J. (2001) The latest hype on Hyp-O-glycosylation codes, *Phytochemistry 57*, 319–323.
11. Kieliszewski, M. J., and Shpak, E. (2001) Synthetic genes for the elucidation of glycosylation codes for arabinogalactan-proteins and other hydroxyproline-rich glycoproteins, *Cellular and Molecular Life Sciences 58*, 1386–1398.
12. Zhao, Z. D., Tan, L., Showalter, A. M., et al., (2002) Tomato LeAGP-1 arabinogalactan-protein purified from transgenic tobacco corroborates the Hyp contiguity hypothesis, *Plant Journal 31*, 431–444.
13. Xu, J. F., Tan, L., Lamport, D. T. A., et al., (2008) The O-Hyp glycosylation code in tobacco and Arabidopsis and a proposed role of Hyp-glycans in secretion, *Phytochemistry 69*, 1631–1640.
14. Shpak, E., Barbar, E., Leykam, J. F., et al., (2001) Contiguous hydroxyproline residues direct hydroxyproline arabinosylation in Nicotiana tabacum, *Journal of Biological Chemistry 276*, 11272–11278.
15. Shpak, E., Leykam, J. F., and Kieliszewski, M. J. (1999) Synthetic genes for glycoprotein design and the elucidation of hydroxyproline-O-glycosylation codes, *Proceedings of the National Academy of Sciences of the United States of America 96*, 14736–14741.
16. De Loose, M., Gheysen, G., Tire, C., et al.,(1991) The extensin signal peptide allows secretion of a heterologous protein from protoplasts, *Gene 99*, 95–100.
17. Murashige, T., and Skoog, F. (1962) A revised medium for rapid growth and bioassays with tobacco cultures, *Physiol. Plant 15*, 473–497.
18. Schenk, R. U., and Hildebrandt, A. C. (1972) Medium and techniques for induction and growth of monocotyledonous and dicotyledonous plant cell cultures, *Can. J. Bot. 50*, 199–204.
19. Mayo, K. J., Gonzales, B. J., and Mason, H. S. (2006) Genetic transformation of tobacco NT1 cells with Agrobacterium tumefaciens, *Nature Protocols 1*, 1105–1111.
20. Osborn, B. L., Olsen, H. S., Nardelli, B., et al., (2002) Pharmacokinetic and pharmacodynamic studies of a human serum albumin-interferon-alpha fusion protein in cynomolgus monkeys, *Journal of Pharmacology and Experimental Therapeutics 303*, 540–548.
21. Knight Jr, E., and Fahey, D. (1981) Human fibroblast interferon. An improved purification, *Journal of Biological Chemistry 256*, 3609–3611.

Chapter 27

Super-promoter:TEV, a Powerful Gene Expression System for Tobacco Hairy Roots

Luis Ñopo, Bonnie J. Woffenden, Deborah G. Reed, Scott Buswell, Chenming Zhang, and Fabricio Medina-Bolivar

Abstract

In order to identify a promoter system for high-level expression of transgenes in hairy roots, we characterized the chimeric super-promoter fused to the translational enhancer from tobacco etch virus (TEV). Transgenic tobacco plants and hairy roots were generated with the super-promoter:TEV sequence and a modified green fluorescence protein (*mGFP5*) as a reporter gene. To exploit the utility of hairy root cultures as a secretion-based expression system, the signal peptide of patatin was fused to *mGFP5* to direct its secretion into the culture medium. Levels of *mGFP5* RNA were on average sixfold higher in hairy roots than leaves. Likewise, GFP protein levels per gram of fresh weight were at least tenfold higher in hairy roots than leaves. Furthermore, more than 10% of the recombinant protein produced in the hairy root culture system was found in the medium. Immunoblotting with anti-GFP antibodies showed two products of 27.1 and 29.9 kDa in all leaf and hairy root tissue extracts, whereas a single 27.1-kDa product was detected in the medium. Inducibility of the promoter was studied with mature leaves and 14-day (midlog phase) hairy roots. A twofold increase in mRNA levels was found immediately after wounding in both mature leaves and hairy roots, with a corresponding increase in mGFP5 protein after 24 h. Our studies demonstrate the utility of the super-promoter:TEV system for high-level expression of recombinant proteins in hairy root bioreactors.

Key words: *ocs* element, Hairy roots, Inducible promoter, Transgenic tobacco, Super-promoter:TEV, GFP

1. Introduction

Hairy root cultures of many species are rapidly gaining recognition as facile biological systems for metabolic engineering of high-value, small molecules and therapeutic proteins not only at a laboratory scale, but also at the commercial level (1–3). To fully exploit the potential of transgenic hairy roots for bioproduction and facilitate

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_27, © Springer Science+Business Media, LLC 2012

recovery of engineered products, gene constructs may be strategically designed to include both a signal peptide to direct secretion of the recombinant protein into culture medium and an affinity tag for ease of purification (4).

The chimeric super-promoter was developed at Purdue University by the group of S.B. Gelvin (5) and is the subject of the current study. This promoter consists of three copies of the octopine synthase activator (*Aocs*) (6) and one copy of the mannopine synthase activator (*Aman*) located upstream of the mannopine synthase (*mas*) promoter (7): all sequences from the *Agrobacterium tumefaciens* pTiA6 plasmid. The effect of the super-promoter was measured using a *GUS* reporter gene in transgenic tobacco showing a 156- and 25-fold increase in GUS activity compared to the 35S and enhanced 35S promoters, respectively (5). The super-promoter was shown to be active in leaves and roots of tobacco. In a recent study, the super-promoter activity was higher in both tobacco and maize roots than in leaves. Furthermore, in tobacco, the activity was higher in mature leaves and apical stems than in young leaves (8).

This activity driven by the super-promoter is correlated with its structure. An analysis for regulatory DNA elements in the super-promoter nucleotide sequence was performed using the PLACE database curated at the National Institute of Agrobiological Sciences in Japan (9). A palindrome of 16 base pairs (5′-ACGTAAGCGCTTACGT-3′) from the *Aocs* sequence, present in the super-promoter, has been previously described as an essential structural component of the ocs enhancer in a transient expression system with maize protoplasts (10). Additionally, this 16-bp sequence functions as a constitutive enhancer in protoplasts of *Nicotiana plumbaginifolia* (11) and as an enhancer for the expression of proteins in tobacco roots (12). The super-promoter also contains the inducible mannopine synthase promoter (*Pmas*), which has been shown to be activated by tissue wounding, auxins, cytokinins, salicylic acid, methyl jasmonate, and 2,6-dichloroisonicotinic acid (13, 14).

Hairy roots could be an effective expression system for transgenes under control of the super-promoter. Hairy roots are produced by genetic transformation mediated by *Agrobacterium rhizogenes*, whereby the T-DNA, which contains genes known to modulate sensitivity to auxins, i.e., *rol* genes, is transferred and integrated into the plant genome. In addition, hairy roots developed with agropine strains of *A. rhizogenes*, such as ATCC 15834, can transfer T-DNA containing genes involved in auxin biosynthesis (15). The super-promoter may, therefore, be particularly well-suited to drive strong transgene expression in hairy roots derived from agropine *A. rhizogenes* strains, which have an inherent auxin-biosynthetic capacity and responsiveness.

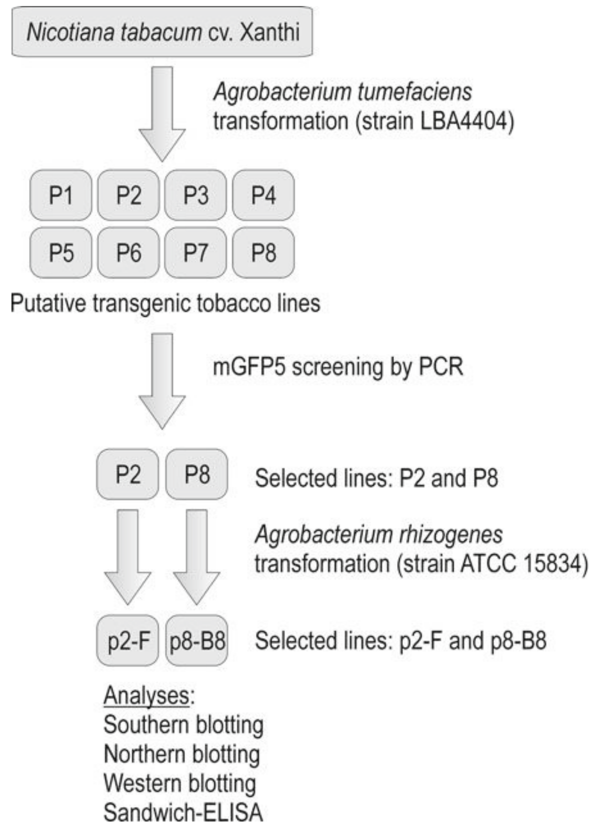


Fig. 1. Scheme followed to evaluate the use of the super-promoter:TEV in tobacco plants and hairy roots.

Based on these precedents, we evaluated the utility of the chimeric super-promoter in conjunction with the 5′ untranslated region (UTR) of the tobacco etch virus (TEV) for expression of recombinant proteins in tobacco hairy roots. The TEV 5′ UTR has been shown to function as a translational enhancer in tobacco (16) and other plants (17, 18). The modified green fluorescent protein (*mGFP5*) was used as the reporter gene. The signal peptide of the potato storage protein, patatin (*pat*), was fused to the *mGFP5* coding sequence to target the recombinant protein to the default secretory pathway (19). Herein, we evaluate the potential of the super-promoter:TEV system for directing gene expression in tobacco leaves and hairy roots by assessing transgene expression at both transcriptional and translational levels. A general scheme followed in this chapter is shown in Fig. 1. Our results show that the super-promoter:TEV drives high-level basal expression in tobacco hairy roots. Furthermore, expression can be increased upon wounding. Our work highlights the utility of the super-promoter in hairy root bioproduction systems.

2. Materials

2.1. Gene Cloning

1. Plasmid pR8-2 as source of the patatin signal peptide (*pat*) and *mGFP5* (20).
2. Binary vector pE1802 (Stanton Gelvin, Purdue University) as source of the super-promoter sequence ((Aocs)$_3$AmasPmas; 5) and TEV.
3. Restriction enzymes *Kpn*I, *Sst*I, and *Xba*I.
4. pBC SK+ (Stratagene) cloning vector.
5. *Escherichia coli* One Shot® TOP10 chemically competent cells (Invitrogen).
6. *A. tumefaciens* strain LBA4404.
7. Luria-Bertani (LB) medium: 5 g/L yeast extract, 10 g/L tryptone, and 10 g/L NaCl. For semisolid medium, add 15 g/L agar.
8. YEP medium: 10 g/L bacto-peptone, 10 g/L yeast extract, and 5 g/L NaCl pH 7. For semisolid medium, add 15 g/L agar.
9. Streptomycin (Strp) stock solution: 30 mg/mL. Dissolve 0.3 g in 10 mL H_2O; sterilize by filtration. Store at 4°C.
10. Kanamycin (Kan) stock solution: 100 mg/mL. Dissolve 1 g in 10 mL H_2O; sterilize by filtration. Store at 4°C.
11. Incubators at 28 and 37°C.
12. Vortex.
13. Microcentrifuge.

2.2. Plant Material Transformation

1. *Nicotiana tabacum* cv. Xanthi in vitro plants.
2. Phytatray™ II vessels (Sigma).
3. Medium mMS: MS salts and vitamins, 3% sucrose, 0.4 g/L $MgSO_4 \cdot 7H_2O$, 4 g/L phytagel; pH 5.7 (4).
4. Medium B5: Gamborg B5 salts and vitamins, 3% sucrose, 0.4 g/L phytagel; pH 5.7 (21).
5. *A. rhizogenes* strain ATCC 15834.
6. 250-mL glass Erlenmeyer flasks.
7. Plant growth incubator.
8. Orbital shaker for hairy root culture.

2.3. Nucleic Acids Analyses

1. DNA primers for mGFP5 detection by PCR: GFP3a (5′-TGCGAGCTCTCATTTGT ATAGTTCATCCAT-3′) and GFP5a (5′-GGCCTCTAGAAGTAAAGGAGAAGAA CTT-3′).
2. REDExtract-N-Amp™ Plant PCR kit (Sigma).
3. QIAquick® PCR Purification kit (Qiagen).

4. DNeasy® Plant Maxi kit (Qiagen).
5. Horizontal electrophoresis system with accessories and power supply.
6. 20× saline-sodium citrate (SSC) buffer: 0.3 M sodium citrate tribasic dihydrate and 3 M sodium chloride.
7. 2× SSC: 1:10 (v/v) dilution of 20× SSC with water.
8. Depurination Southern solution: 250 mM HCl.
9. Denaturation Southern solution: 0.5 M NaOH and 1.5 M NaCl.
10. Neutralization Southern solution: 0.5 M Tris–HCl, pH 7.5, and 1.5 M NaCl.
11. Positively charged nylon membrane (Roche).
12. Digoxigenin (DIG) system for detection of DNA by Southern blotting (Roche).
13. DIG-labeled DNA molecular weight markers (Roche).
14. DIG Easy Hybridization solution (Roche).
15. DIG Wash and Block Buffer Set (Roche).
16. CDP-Star substrate (Roche).
17. Washing Southern solutions I: 2× SSC with 0.1% SDS.
18. Washing Southern solution II: 0.5× SSC with 0.1% SDS.
19. RNA extraction buffer: 0.18 M Tris–HCl, pH 8.2, 0.09 M LiCl, 4.5 mM EDTA, and 1% SDS.
20. TE-equilibrated phenol, pH 4.3.
21. Chloroform.
22. 3 M sodium acetate, pH 5.2.
23. 100% ethanol.
24. 70% ethanol.
25. DEPC-treated water.
26. 2 M LiCl.
27. 37% formaldehyde (see Note 1).
28. 10× MOPS–EDTA–sodium acetate (MESA) buffer: 0.4 M MOPS, pH 7.0, 10 mM EDTA, and 0.1 M sodium acetate.
29. Hydrolysis Northern solution: 0.05 N NaOH.
30. Neutralization Northern solution: 0.5 M Tris–HCl, pH 6.8.
31. Transfer buffer for Northern: 25 mM sodium phosphate, pH 6.5.
32. DIG Easy hybridization solution (Roche).
33. DIG DNA labeling kit (Roche).
34. Washing solution for Northern blots: 0.5× SSC with 0.1% SDS.

35. Substrate and enhancer to develop reaction by chemiluminescence based on alkaline phosphatase (AP): Immun-Star™ AP substrate (Bio-Rad) and Nitro-Block Enhancer II™ (Tropix).
36. Image capturing system for ethidium bromide-stained gels.
37. UV cross-linker.

2.4. Protein Analyses

1. Sodium carbonate buffer: 0.1 M Na_2CO_3, pH 9.6.
2. Miracloth™ (Calbiochem-EMD Biosciences).
3. Advanced™ protein assay solution (Cytoskeleton).
4. Bovine serum albumin (BSA) for protein quantifications.
5. rGFPuv (GFP, Clontech) as control for Western blotting and sandwich ELISA.
6. NuPAGE™ 10% Bis-Tris acrylamide gels (Invitrogen).
7. MES running buffer (Invitrogen).
8. X-Cell SureLock™ Mini-Cell apparatus (Invitrogen).
9. 0.2 μm Trans-Blot® nitrocellulose membrane (BioRad).
10. NuPAGE™ transfer buffer: 25 mM Bicine, 25 mM Bis-Tris, 1 mM EDTA, 0.05 mM chlorobutanol (Invitrogen).
11. X-Cell II™ blot module (Invitrogen).
12. Primary antibodies to detect GFP by either Western blotting (Living Colors™ full-length *Aequorea victoria* GFP polyclonal antibody raised in rabbit (BD Biosciences)) or sandwich ELISA (rabbit anti-GFP polyclonal antibody (Molecular Probes)).
13. Secondary antibody to develop reactions by Western blotting and sandwich ELISA: Goat polyclonal antibody against rabbit IgG conjugated to AP enzyme (Sigma).
14. 10× PBS: 0.1 M Na_2HPO_4, 17.6 mM KH_2PO_4, pH 7.4, 1.37 M NaCl, and 26.83 mM KCl.
15. PBS-T: 1× PBS, 0.1% Tween-20.
16. Blocking buffer: 3% BSA in PBS-T.
17. Immun-Star™ AP substrate (BioRad).
18. Nitro-Block Enhancer II™ (Tropix).
19. AP substrate (Pierce).
20. PBS-T for ELISA: 100 mM sodium phosphate, pH 7.4, 150 mM NaCl, and 0.05% Tween-20.
21. 2 N NaOH.
22. 96-well Reacti-Bind™ Anti-GFP plate coated with polyclonal mGFP5 goat antibodies (Pierce).
23. Microplate Reader model EL808 with KC Junior v1.41 software (Bio-Tek).

3. Methods

3.1. Construction of Plasmid pWSB1

Basic DNA cloning techniques were used and they are not detailed here due to space limitations. When needed, plasmid pBC SK+ and *Escherichia coli* One Shot® TOP10 chemically competent cells were used in cloning steps. The DNA sequence encoding the patatin signal peptide (*pat*) was excised from plasmid pR8-2 with *Kpn*I and *Xba*I restriction endonucleases (Fig. 2a) (see Note 2). This fragment was subcloned into binary vector pE1802 (Fig. 2b) which contains the super-promoter ($(Aocs)_3$AmasPmas) upstream of the TEV translational enhancer (with flanking 3′ *Nco*I site), multicloning site, agropine synthase terminator (*ags-ter*), and neomycin phosphotransferase II (*nptII*) selectable marker under the control of the nopaline synthase promoter (Pnos).

The DNA fragment encoding *mGFP5* (GenBank accession number U87974, nucleotides 87–797) was excised from plasmid pR8-2 with *Xba*I and *Sst*I restriction endonucleases, and subcloned into the patatin signal peptide containing pE1802 plasmid, to create plasmid pWSB1 (Fig. 2c; details of fragment *Pmas:TEV:pat:mGFP5* are in Fig. 2d). Plasmid pWSB1 was introduced into *A. tumefaciens* LBA4404 by a freeze–thaw method described in ref. 22.

1. Sow *A. tumefaciens* LBA4404 on semisolid YEP medium containing 50 mg/L Strp in Petri dishes. Incubate plates overnight at 28°C.
2. Transfer one single colony to YEP liquid, containing 50 mg/L Strp. Incubate overnight at 28°C with shaking at 250 rpm.
3. Read OD at 600 nm. When absorbance is in the 0.5–1.0 range, collect the bacteria by centrifugation at 15,700 × *g* at 4°C. Discard supernatant.
4. Dissolve the pellet in 100 μL of 20 mM $CaCl_2$, filter sterilized (0.22 μm) and kept at 4°C.
5. Add 1 μg of plasmid pWSB1 contained in 20 μL H_2O and carefully mix by hitting externally the microtube by hand.
6. Immerse the tube in liquid nitrogen for 30 s and immediately transfer it to a water bath at 37°C, incubating it for 5 min.
7. Add 300 μL YEP, and incubate bacteria at 28°C for 2 h, with shaking at 250 rpm.
8. Plate bacteria in semisolid YEP containing 50 mg/L Strp and 50 mg/L Kan and incubate at 28°C until bacterial colonies are seen, usually for 24–48 h.

3.2. Plant Transformation

1. Sow *A. tumefaciens* harboring the pWSB1 plasmid in YEP semisolid medium containing 50 mg/L Strp and 50 mg/L Kan in Petri dishes. Incubate overnight at 28°C.

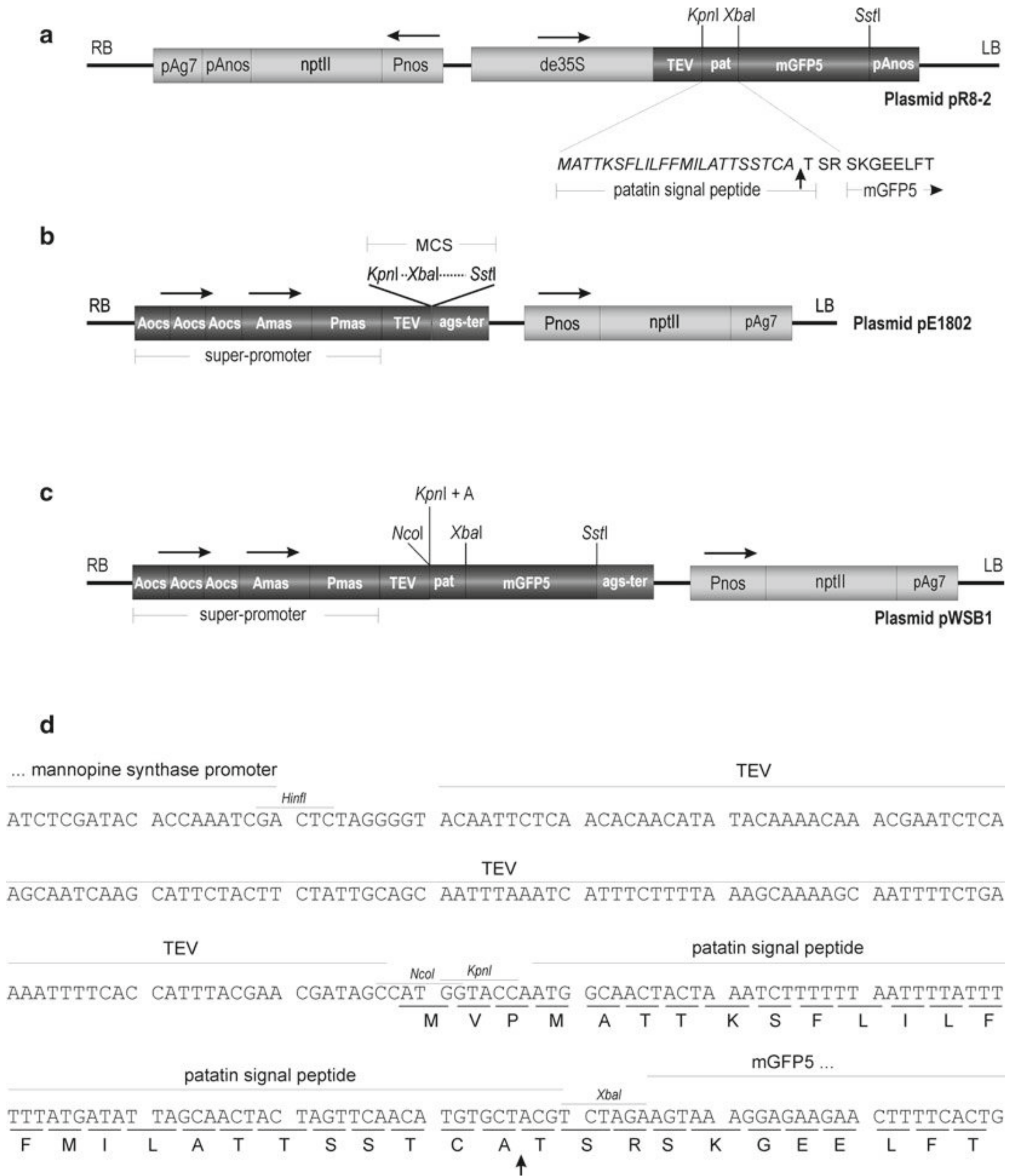


Fig. 2. Plasmid constructions used in tobacco transformations. (**a**) Plasmid pR8-2 (de35S:TEV:pat:mGFP5), (**b**) plasmid pE1802 (super-promoter:TEV), and (**c**, **d**) plasmid pWSB1 (super-promoter:TEV:pat:mGFP5). *de35S* double-enhanced 35S promoter, *TEV* tobacco etch virus translational enhancer, *pat* patatin signal peptide, *mGFP5* modified green fluorescent protein, *pAnos* polyadenylation sequence of the nopaline synthase gene, *Aocs* activating sequence of the octopine synthase gene, *Amas* activating sequence of the mannopine synthase gene, *Pmas* mannopine synthase promoter, *ags-ter* agropine synthase terminator, *Pnos* nopaline synthase promoter, *nptII* neomycin phosphotransferase II gene, *pAg7* T-DNA gene7 polyadenylation signal, *MCS* multicloning site, *RB* and *LB* right and left borders of the T-DNA, respectively. *Vertical arrow* between A and T amino acids in Fig. 1a and detail in Fig. 1d indicate the predicted signal peptide cleavage. Amino acid sequence MVP at the N-terminal end of patatin signal peptide created in pWSB1 is also shown (**c** and **d**).

2. Excise leaves containing the petiole from 1-month-old in vitro tobacco plants (*N. tabacum* cv. Xanthi) and place them onto semisolid mMS medium contained in Petri dishes.
3. With a dissecting knife, take one single-bacterial colony and proceed to make a longitudinal cut along the petiole. By this procedure, the plant cells are exposed to the infecting bacteria allowing cell transformation.
4. Seal plates with Parafilm® and incubate them at 24°C. Calli appear within 2 weeks following the inoculation, and putative transgenic plants develop before a month of inoculation.
5. Isolate putative transgenic plantlets and subculture them in Phytatray™ II vessels containing mMS medium. Incubate them under a 16-h photoperiod, 24°C, and average light intensity of 650 × 10 lx.

   Screen regenerated plants by PCR to verify the presence of the transgene (see Subheading 3.3) and proceed to the next step to induce hairy roots development.
6. Sow *A. rhizogenes* strain ATCC 15834 on YEP medium.
7. In a similar way as previous infection with *A. tumefaciens*, isolate a tobacco leaf from a previously verified transgenic plant. With a blade, take an *A. rhizogenes* colony and make a small (approximately 1 cm) cut along the midrib.
8. Place the leaf in B5 medium contained in a Petri dish. Small calli followed by putative hairy roots develop within 2 weeks.
9. Excise hairy roots as isolated lines and place them on semisolid B5 medium contained in Petri dishes. Incubate them at 24°C in the dark.
10. For growth characterization, inoculate ten root tips (approximately 1 cm length) of each line in 250-mL Erlenmeyer flasks containing 50 mL of liquid B5 medium (21), and incubate them at 24°C with continuous light (see Note 3) and shaking at 90 rpm. Do this by triplicate for each time point.
11. Measure fresh weight, pH, and conductivity every 3 days until day 33 (see Note 4).

3.3. Determination of Gene Insertion in Tobacco Plants by PCR

A fast screening for the presence of the transgene (*mGFP5*) in putative transgenic plants can be performed by PCR.

1. Cut a 4-cm² leaf section (equivalent to 200 mg) from a 4-week-old in vitro plantlet.
2. Mince it with a razor blade on a glass microscope slide, and transfer plant tissue into a microcentrifuge tube containing 100 μL extraction solution (REDExtract-N-Amp™ Plant PCR kit).
3. Vortex the sample for 30 s and incubate it at 95°C for 10 min followed by addition of 100 μL dilution buffer (REDExtract-N-Amp™ Plant PCR kit).

4. Vortex for 30 s and transfer 100 μL of this mixture to a DNA binding column (QIAquick® PCR Purification kit).
5. Place columns in a collection microcentrifuge tube and add 5 μL 3 M sodium acetate, pH 5.2, and 500 μL of binding buffer (QIAquick® PCR Purification kit).
6. Centrifuge for 1 min at 10,000 × *g*, and discard the flow-through solution.
7. Wash columns with 750 μL wash buffer (QIAquick® PCR Purification kit) and centrifuge as before.
8. Elute DNA by centrifugation with 30 μL sterile water.
9. Use 4 μL (see Note 5) of the eluted DNA extract for PCR using primers GFP3a and GFP5a. Amplification conditions were 30 cycles of 95°C for 1 min, 52°C for 1 min, and 72°C for 1 min.

We selected eight plantlets identified consecutively from P1 to P8. After PCR screening, lines P6 and P7 were discarded as they did not show the 733-bp band corresponding to *GFP* amplification. Several hairy root lines were further developed from leaves of other transgenic tobacco plants. Because of their fast growth, hairy root lines identified as p2-F and p8-B8 (according to their originating plant) were selected for subsequent experiments. Growth curves were developed for these hairy root lines.

For this and further experiments, transgenic tobacco plants and hairy roots containing the T-DNA from plasmid pR8-2 (de35S:TEV:pat:GFP, Fig. 2a) were used as controls (plant line R8 and hairy root line r8). They have been previously described (20). Wild-type Xanthi plant and hairy roots were also considered as negative controls for GFP detection (Fig. 3a).

3.4. Southern Blotting

Copy number of the transgene can be visualized in positive, PCR-analyzed, regenerated plants and hairy roots. For this, plant genomic DNA should be digested with a suitable restriction enzyme (single cutter within T-DNA). The enzyme *Xba*I was used in this case. DNA fragments are then detected by Southern blotting.

1. Extract DNA from PCR-positive plants and hairy roots following the procedure described by the DNeasy® Plant Maxi kit (Qiagen).
2. Digest 10 μg of genomic DNA with *Xba*I and resolve it on a 0.7% agarose gel. Include DIG-labeled DNA ladder as molecular weight marker.
3. After electrophoresis, stain the gel with ethidium bromide and photograph it under UV light.
4. Immerse the agarose gel in depurination solution for 20 min with shaking, followed by a brief rinse in water.

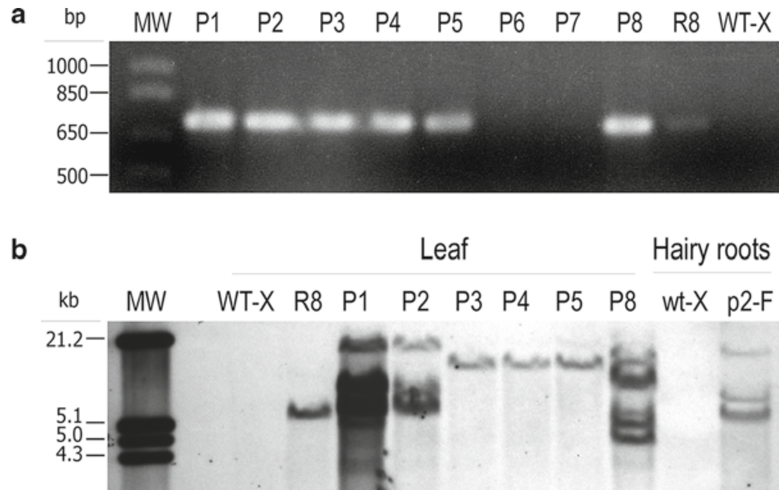


Fig. 3. Detection of mGFP5 gene in transgenic lines. (**a**) Screening by PCR of leaf DNA. (**b**) Southern blotting for PCR-positive leaf or hairy roots samples. 10 μg of genomic DNA were digested with *Xba*I and resolved on a 0.7% agarose gel, followed by transfer to positively charged nylon membrane. The blot was hybridized to a DIG-labeled GFP-specific probe, and detected via chemiluminescence. Analyzed DNA was extracted from regenerated transgenic tobacco plants (P1–P8), a positive control containing mGFP5 (R8), and nontransformed tobacco cv. Xanthi (WT-X); hairy root analysis included lines p2-F and wt-X derived from plants P2 and WT-X, respectively; *MW* DIG-labeled molecular weight markers.

5. Immerse the gel in denaturation solution, twice for 15 min each with shaking, followed by a brief rinse in water.
6. Immerse the gel in neutralization solution with shaking, twice for 15 min each.
7. Equilibrate the gel in 20× SSC for 10 min.
8. Proceed to transfer overnight the DNA to a positively charged nylon membrane (Roche) by capillarity in 20× SSC.
9. Following transfer, rinse briefly the membrane with 2× SSC and allow it to dry.
10. Immediately, fix DNA to the membrane by UV cross-linking. A DNA probe is necessary to detect the target gene in samples. In these experiments, the *mGFP5* probe was prepared by PCR from a plasmid containing this gene as follows.
    (a) In a microcentrifuge tube, pipette 1 ng of plasmid template.
    (b) Add GFP3a and GFP5a primers (described above) at a final concentration of 0.5 μM each.
    (c) Add 0.2 mM each dATP, dCTP, dGTP, 0.167 mM dTTP, 0.033 mM DIG-dUTP (Roche), 1.5 mM $MgCl_2$, and 2.5 units Taq polymerase in a final reaction volume of 100 μL. Optimal PCR parameters for these primers are 94°C for

3 min; 35 cycles of 94°C for 30 s, 52°C for 30 s, 72°C for 1 min; and 72°C for 10 min.

(d) Probe should be denatured by heating at 95°C for 10 min, just before hybridization.

(e) Prehybridize the membrane in DIG Easy Hyb solution (Roche) containing DIG Blocker (from DIG Wash and Block Buffer Set, Roche) for 1 h in a hybridization oven at 65°C.

(f) To the previous solution at 65°C, add 2 μL of labeled probe per mL, and hybridize overnight at 65°C in an oven.

(g) After hybridization, wash the blot twice for 20 min each in 2× SSC, 0.1% SDS, and then twice for 20 min each in 0.5× SSC, 0.1% SDS, preheated to 65°C.

(h) Proceed with the chemiluminescent detection. For this, follow manufacturer's recommendations using the DIG Wash and Block Buffer Set (Roche), and develop using CDP-Star substrate.

A Southern blot using a DIG-labeled mGFP5 probe confirmed the presence of the transgene in regenerated plants (Fig. 3b). The six PCR-positive plants likely represent four independent transformation events. Line P1 apparently contains three to four independent transgene insertion loci, whereas line P2 contains three independent integration loci; line P8 represents five independent insertions. Considering that the band pattern of plants P3, P4, and P5 are identical, they were probably derived from a single transformation event and are clonal plants. No bands were observed in the nontransformed Xanthi genomic DNA control. The positive-control R8-2 plant contains a single copy insert. In addition, hairy root line p2-F showed the same three bands as its parental P2 plant.

3.5. Verification of Transgene Transcription by Northern Blotting

Northern blotting analysis gives an outline of the expression of the foreign gene in modified plant tissues. This is a protocol to extract total RNA from plant tissues.

1. Grind 2.5-g plant tissue in liquid nitrogen.
2. Transfer tissue to a tube containing 25 mL RNA extraction buffer and 8.3 mL (0.3 volumes) of TE-equilibrated phenol, pH 4.3. Vortex vigorously for 2 min.
3. Add 0.3 aqueous volume of chloroform and vortex for 30 s.
4. Centrifuge at 17,700 × *g*, 10 min, at 4°C to separate the aqueous and organic phases.
5. Repeat twice the phenol/chloroform extraction (steps 2 and 3) of the aqueous layer, followed by two extractions with chloroform alone.
6. Transfer the aqueous layer to a fresh tube on ice.

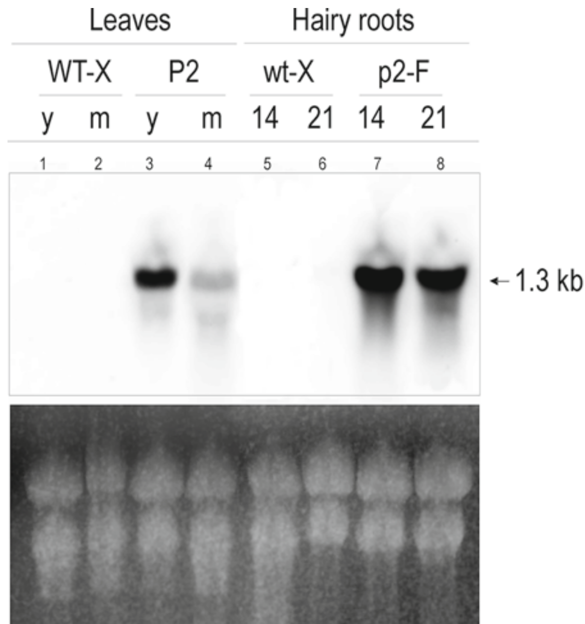


Fig. 4. Basal levels of mGFP5 mRNA in leaves and hairy roots. Leaves from 4-week-old transgenic P2 plants (**P2**) and 14- and 21-day-old hairy root cultures of p2 line F (p2-F) are compared to tobacco cv. Xanthi nontransformed leaves (WT-X) and hairy roots (wt-X). Upper block: northern blot showing the expected 1.3 kb mGFP5 mRNA band; lower block: equal loaded agarose gel of total RNA (20 μg per lane), visualized by ethidium bromide staining before Northern analysis.

7. Precipitate RNA by the addition of 0.1-volume 3 M sodium acetate, pH 5.2, and 2.5-volumes 100% ethanol. Incubate at –70°C for 30 min.
8. Centrifuge at 17,700 × *g* for 20 min at 4°C.
9. Resuspend pellet in 1 mL DEPC-treated water and selectively precipitate RNA with addition of 2 M LiCl. Incubate overnight at 4°C.
10. Centrifuge at 17,700 × *g* for 20 min at 4°C.
11. Wash the RNA pellet with 2 mL of ice-cold 2 M LiCl.
12. Resuspend RNA in 100 μL DEPC-treated water.

To estimate the steady-state levels of *mGFP5* mRNA, we analyzed young and mature leaves (Fig. 4) excised from 4-week-old P2 transgenic plantlets, as well as hairy roots (14- and 21-day-old cultures). Nontransformed plantlets and hairy roots of tobacco cv. Xanthi were used as controls. After harvesting, tissues were frozen in liquid nitrogen and stored at –70°C until use.

1. Prepare a 1.2% agarose gel containing 2.22 M formaldehyde (see Note 6).
2. Load 20 μg of RNA per sample.

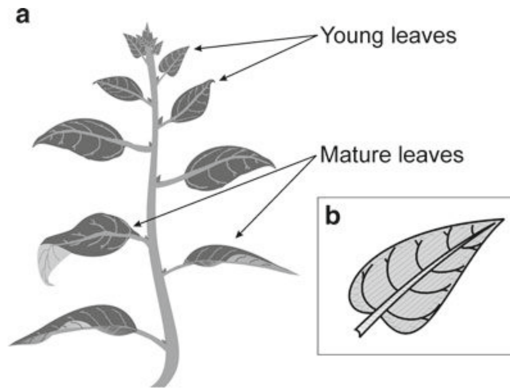


Fig. 5. Scheme for leaf sampling and dissection. Tobacco transgenic plantlets were grown in vitro for 4 weeks before sampling. (**a**) Leaf localization indicating young and mature leaves. (**b**) *Thin lines* show the wounding on leaves with a dissecting blade.

3. Run electrophoresis and photograph gel under UV light.
4. Wash the gel with 0.05 N NaOH to hydrolyze the RNA, for 20 min, with shaking at room temperature (RT).
5. Neutralize the gel in 0.5 M Tris–HCl, pH 6.8, for 1 h, with shaking at RT.
6. Transfer RNA to a positively charged nylon membrane by capillarity using 25 mM sodium phosphate buffer, pH 6.5, for 16 h.
7. Let the membrane to air dry, and cross-link RNA by UV irradiation.
8. Prehybridize with DIG Easy Hyb solution at 65°C for 1 h.
9. Add 6 μg DIG-labeled probe and hybridize overnight at 65°C (see Note 7).
10. After hybridization, proceed with washings of the membrane and detection according to steps 17 and 18 in Subheading 3.4.

We labeled *mGFP5* to be used as probe. The probe was prepared by PCR from plasmid R8-2 according to manufacturer's recommendations (Roche) using primers GFP5a and GFP3a described before.

Northern hybridization of total RNA confirmed the presence of an *mGFP5*-specific transcript of approximately 1.3 kb in all samples from transgenic tissues (Fig. 4). The size of the transcript is consistent with the predicted full-length mRNA, including the 5′ and 3′ untranslated regions. This fragment was absent in nontransformed Xanthi leaves and hairy roots.

To determine whether the promoter is differentially regulated at different stages of development, we compared the mRNA levels in young (apical) and mature leaves (Fig. 5a). The expression of *mGFP5* in P2 line was 3.5-fold higher in young leaves compared to that in mature leaves (Fig. 4, lanes 3 and 4), analyzed by

densitometry. The highest levels of expression were observed in hairy roots. Compared to mature leaves, p2 plant-derived hairy roots line F (p2-F) from 14-day-old cultures expressed sixfold higher *mGFP5* mRNA levels (Fig. 4, lane 7) and roots from 21-day-old cultures expressed 4.8-fold higher levels of p2 mGFP5 mRNA (Fig. 4, lane 8).

3.6. Total Protein Extraction from Leaves and Hairy Roots, and Detection of the Recombinant Protein mGFP5 by Western Blotting

1. Excise 1 g fresh weight of either young (average of 3 leaves) or mature leaves (average of 2 leaves) from 4-week-old in vitro plantlets (Fig. 5a). For hairy roots, use tissue from either 14- or 21-day-old cultures started in 50 mL B5 medium contained in 250-mL-capacity Erlenmeyer flasks.
2. Grind plant tissue under liquid nitrogen with mortar and pestle.
3. Transfer powder to a microcentrifuge tube containing 1 mL sodium carbonate buffer (0.1 M Na_2CO_3, pH 9.6), followed by vortexing for 15 min.
4. Centrifuge for 10 min at 4°C in a microcentrifuge at 9,279 × *g*.
5. Filter supernatant through two layers of Miracloth™.
6. Determine total protein content using Advanced™ protein assay using BSA as standard protein dissolved in sodium carbonate buffer.
7. Analyze protein extracts by NuPAGE™ 10% Bis-Tris acrylamide gels in an X-Cell SureLock™ Mini-Cell apparatus using MES running buffer and according to manufacturer's procedures (1 h at 200 V).
8. Proceed with the Western blot transfer. For this, cut a 0.2 µm Trans-Blot® nitrocellulose membrane of the same size of the gel (Bio-Rad) and assemble the transfer system (X-Cell II™ blot module; Invitrogen) according to the manufacturer's procedures using NuPAGE™ transfer buffer (25 mM bicine, 25 mM Bis-Tris, 1 mM EDTA, 0.05 mM chlorobutanol; Invitrogen).
9. Transfer for 1 h at 30 V.
10. Block the membrane with blocking solution for 1 h at RT.
11. Add the primary antibody (Living Colors™ full-length *A. victoria* GFP polyclonal antibody raised in rabbit (BD Biosciences)) at a 1:1,000 (v/v) dilution in blocking solution for 1 h at RT.
12. Wash membrane three times (15-min wash) in PBS, 0.1% Tween-20.
13. Add secondary antibody (goat anti-rabbit-IgG-whole molecule-AP-conjugated, Sigma) at a 1:3,000 (v/v) dilution for 45 min at RT.
14. Washed three times as above.
15. Continue with the chemiluminescence detection using Immun-Star™ AP substrate (Bio-Rad) and Nitro-Block Enhancer II™ (Tropix) following manufacturer's procedures.

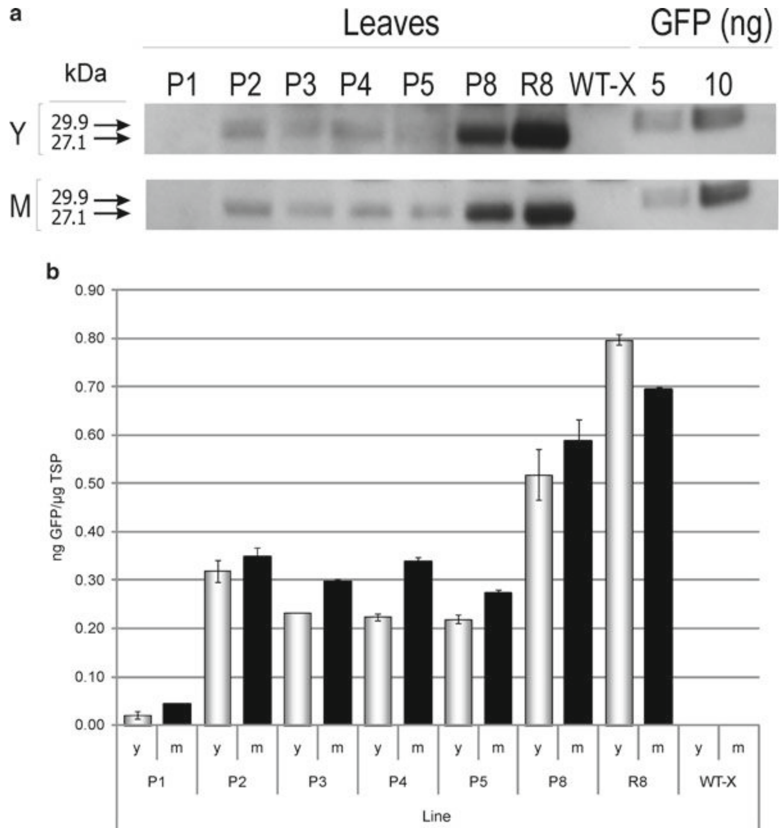


Fig. 6. Basal mGFP5 levels in a total protein extract from young and mature leaves in 4-week-old transgenic plants. (**a**) Western blot loading 50 µg TSP/well from young (Y) and mature (M) leaves; in both cases, the *arrows* show the 27.1- and 29.9-kDa bands for transgenic plant GFP. (**b**) GFP-ELISA from young (*gray bars*) and mature (*black bars*) leaves. P1, P2, P3, P4, P5, P8: Transgenic lines transformed with the plasmid pWSB1. R8: Positive control. WT-X: Nontransformed Xanthi. 5 and 10 ng/well commercial mGFP5 were used as detection control.

Total protein was extracted from all positive plants identified by Southern blot. Plantlet R8 and nontransformed Xanthi were used as mGFP5-positive and -negative controls, respectively.

Western blots show both the recombinant mGFP5 protein and the commercial rGFPuv as a double band of 29.9 and 27.1 kDa (see Note 8). This was observed without exception in all of the six transgenic lines engineered with the plasmid pWSB1 (Fig. 6a). Western blots performed with leaves from transgenic plants showed that mGFP5 was accumulated in both young and mature tissues (Fig. 6a, b), with the strongest mGFP5 expression observed in line P8. A positive control performed with a protein extract from the plant R8-2 (de35S:TEV:pat:mGFP5) showed a higher level of mGFP5 protein than super-promoter:mGFP5 plants.

3.7. mGFP5 Quantification by Sandwich ELISA

A method previously described by Richards et al. (23) was followed with modifications.

1. Ten microliters of total protein extracts from leaf and hairy roots samples (average of 2.6 and 2.3 μg total protein/μL, respectively) were assayed. In addition, secreted mGFP5 from hairy roots was quantified using 100 μL of the liquid culture medium. Recombinant mGFP5 (rGFPuv, BD Biosciences) ranging from 0.5 to 200 ng per well was used as standard.
2. Apply samples and standards to a 96-well Reacti-Bind™ Anti-GFP plate coated with polyclonal mGFP5 goat antibodies (Pierce) in a total volume of 100 μL per well with PBS-T, and incubate for 1 h at RT.
3. Wash wells three times with 200 μL PBS-T.
4. Add rabbit anti-GFP polyclonal antibody at a 1:10,000 (v/v) dilution in 200 μL blocking buffer, and incubate for 30 min.
5. Wash as before.
6. Add goat polyclonal antibody against rabbit IgG conjugated to AP enzyme at a 1:3,000 (v/v) dilution in 200 μL blocking buffer, and incubate for 30 min.
7. Wash as above.
8. Develop reaction by applying 100 μL AP substrate per well and let react for 15 min.
9. Stop reaction by adding 50 μL 2 N NaOH.
10. Read absorbance at 405 nm in a plate reader.

Agrobacterium transfers genes into random position throughout the plant genome; it is important to characterize the transgenic events to identify the best expresser genotype for scaling up. A sandwich ELISA assay confirmed that the best mGFP5 expresser was line P8 (588.29 ± 44.37 and 518.35 ± 52.42 ng mGFP5/mg total soluble protein (TSP) in mature and young tissues, respectively), followed by line P2 (348.10 ± 18.58 and 318.76 ± 23.30 ng mGFP5/mg TSP in mature and young tissues, respectively). For R8, mGFP5 amounts were 695.30 ± 3.91 and 797.42 ± 10.77 ng mGFP5/mg TSP in mature and young tissues, respectively. Nontransformed Xanthi showed zero levels of mGFP5 (Fig. 6b) (see Note 9).

Growth and culture characteristics of hairy root lines p2-F, p8-B8, wt-Xanthi, and R8 in liquid culture were characterized (Fig. 7a). It was observed that pH values increased and conductivity values decreased throughout the culture period (Fig. 7b). Accumulation of mGFP5 in hairy roots was assessed by a Western blot of total protein extracted from line p2-F. mGFP5 was visualized as a double band of 27.1 and 29.9 kDa in every sample harvested at all time points from days 6 to 21 of culture. Protein expression quantification was precluded prior to day 6 of culture since root tissue

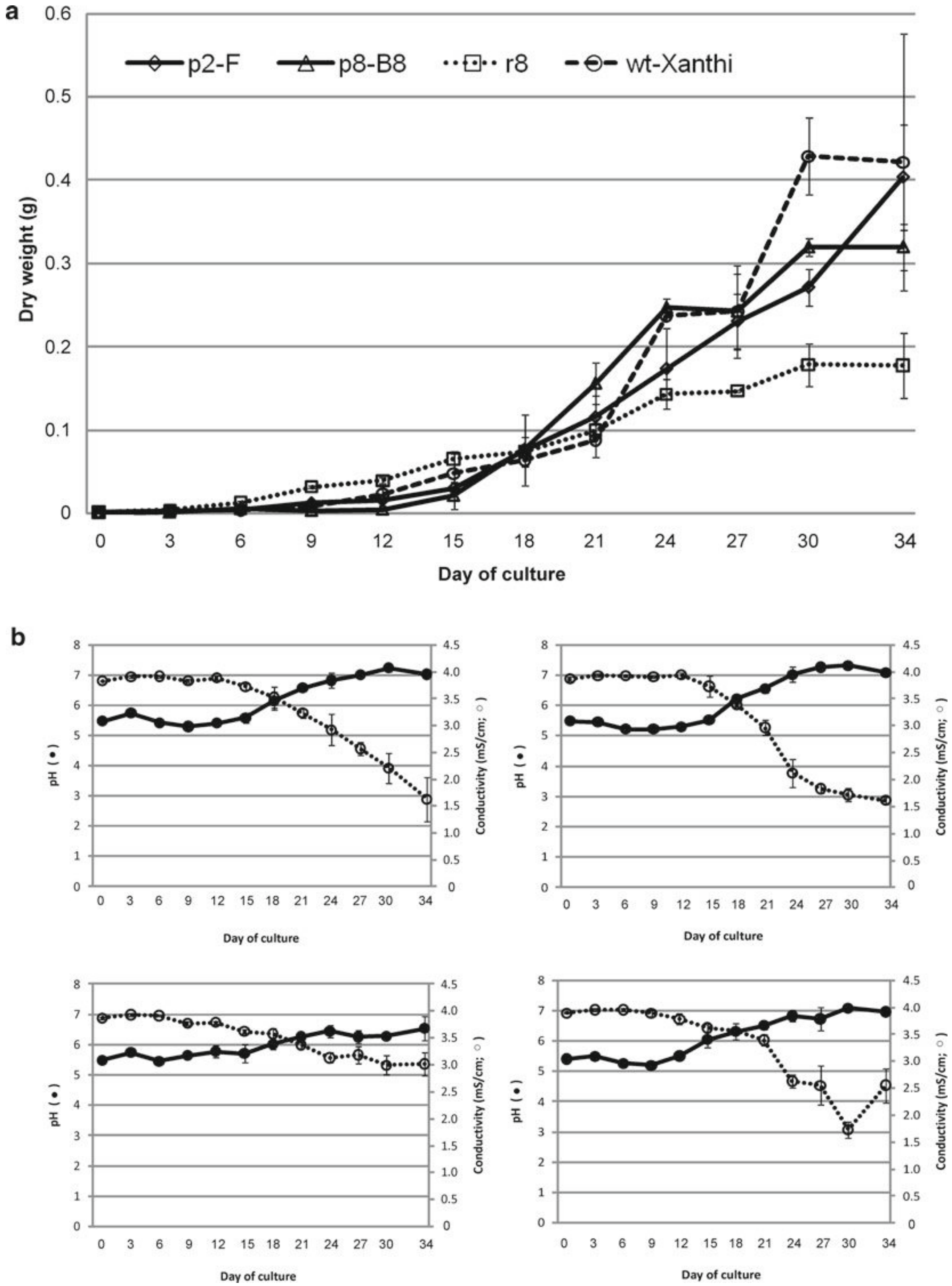


Fig. 7. Hairy root characterization of lines p2-F and p8-B8. (**a**) Growth curve in B5 liquid medium in a 34-day period. (**b**) Changes in pH (*filled circle*) and conductivity (*open circle*) of the B5 culture media. r8 and wt-Xanthi hairy roots were used as controls.

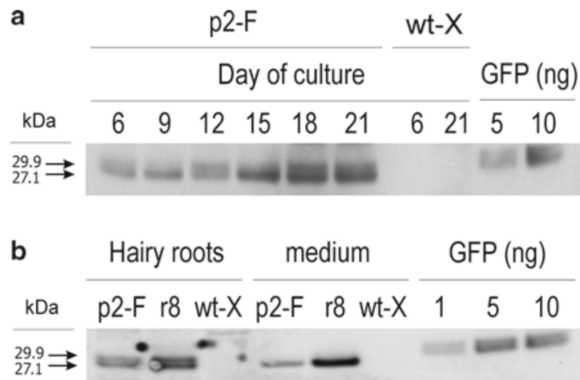


Fig. 8. mGFP5 in hairy root liquid cultures. (**a**) mGFP5 accumulation in tissue of hairy root line p2-F between 6 and 21 days in 30 mL cultures, compared to nontransformed Xanthi (wt-X), visualized by Western blot. (**b**) Western blot for mGFP5 in hairy roots on 14th day of growing in 50 mL cultures. For tissue samples, 50 μg TSP were loaded per lane, and 20 μL from B5 culture media were loaded per lane in **b**; in addition, 1, 5, and 10 ng commercial GFP were used as control. The *arrows* show the 27.1- and 29.9-kDa bands for transgenic plant GFP while commercial GFP is indicated as a double band in the 30-kDa region. Hairy roots r8 and wt-X were used as controls as indicated.

was not sufficient for analysis. A Western blot showed that mGFP5 protein levels increase steadily between days 6 and 15. No difference in the amount of mGFP5 in hairy root samples harvested on days 18 and 21 of culture was observed (Fig. 8a). Similarly, hairy roots from both p2-F and R8-2 lines showed two bands for GFP. Interestingly, a Western blot of proteins secreted from the hairy roots to the medium shows a single immunoreactive product of 27.1-kDa band in both p2-F and R8-2 cultures (Fig. 8b).

In order to assess super-promoter:TEV performance in another transgenic tobacco hairy root, line p8-B8 (derived from plant line p8) was examined by ELISA for its expression of recombinant mGFP5 protein. Ten root tips of this line were subcultured in flasks with 50 mL of B5 medium for 14 days at which time 1.7 g of hairy roots were harvested. At this stage, the root tissue contained 4,371 ng mGFP5/mg TSP representing a total of 10.4 μg mGFP5. In addition, 1.35 μg mGFP5 was found in the liquid medium. This amount of secreted mGFP5 represents 11% of the total mGFP5 produced in this hairy root culture system. In comparison, a 4-week-old p8 plant yielded 518 ng mGFP5/mg TSP (420 ng mGFP5/g FW) in young leaves and 588 ng mGFP5/mg TSP (1,026 ng mGFP5/g FW) in mature leaves, values which are approximately eightfold lower than those obtained from hairy roots cultured for half as long as leaf tissue. Similarly, the amount of recombinant mGFP5 protein was much higher in hairy roots of line P8 (containing super-promoter:TEV) than in leaves of line R8 (containing de35S:TEV promoter). Quantification of mGFP5 by ELISA was reproducible in both leaf and hairy root tissue, as well as hairy root culture media, over 3 years of in vitro subculture.

3.8. Induction of the Super-promoter by Wounding of Leaves or Hairy Roots

According to previous reports, sequences in the super-promoter respond to tissue wounding. To test this finding, GFP accumulation (*mGFP5* as reporter gene) was measured in mechanically wounded tissue (leaves and hairy roots) (see Note 10).

1. Harvest 1 g leaves from 4-week-old plantlets (1–2 leaves) and place them in empty Petri dishes.
2. Cut extensively plant tissue with a clean dissecting blade (Fig. 5b). Controls should not be wounded.
3. Seal Petri dishes with Parafilm™ to avoid dehydration and place them at RT for 6, 12, 24, and 48 h.
4. Immerse samples in liquid nitrogen and store them at –70°C until use.
5. Analyze samples by previously described techniques, including Northern blot, Sandwich ELISA, and Western blot.

To examine the potential for induction of the super-promoter by abiotic stress, we conducted a time-course experiment of wounding in mature leaf tissue. Wounded P2 leaves were processed for total RNA isolation after 0-, 2-, 6-, 12-, and 24-h incubation time post wounding. Leaves harvested for RNA isolation immediately after wounding (0 h) showed approximately twofold higher levels of expression (Fig. 9a, lane 3) compared to the nonwounded control sample (Fig. 9a, lane 2). Throughout the 2-12-h period post wounding, p2 *mGFP5* mRNA expression was induced between 3.8- and 4.0-fold over the levels observed in the nonwounded samples (Fig. 9a, lanes 4–9). Levels of p2 *mGFP5* mRNA declined 24 h after wounding (Fig. 9a, lanes 10–11) to approximately 0-h levels.

Induction of the super-promoter by wounding of hairy root line p2-F resulted in an immediate increase of *mGFP5* mRNA levels at 0 h, a similar response observed in leaves. A 1.3-fold increase was observed in samples harvested at 0 h (Fig. 9b, lanes 1–2), and 1.4-fold increase 2 h after wounding (Fig. 9b, lanes 3–4). Although *mGFP5* mRNA was present 6 and 24 h after wounding, no induction was observed compared to the corresponding nonwounded controls (Fig. 9b, lanes 5–8), with the levels of p2 *mGFP5* even lower than the levels observed in the nonwounded controls.

Analysis of the super-promoter in hairy roots was performed as follows. Liquid culture of hairy roots p2-F line was initiated with five root tips in 30 mL of liquid B5 media in 125-mL-capacity Erlenmeyer flasks and grown as before. After 14 days of culture, hairy roots were placed onto Petri dishes and cut transversely in approximately 2-mm sections with a dissecting blade. Plates were sealed with Parafilm™ and placed at RT for 0, 2, 6, and 24 h. Samples were stored and processed as described for Northern blot.

Six to forty-eight hours post induction by wounding, mGFP5 protein accumulated in P2 leaves. According to densitometry

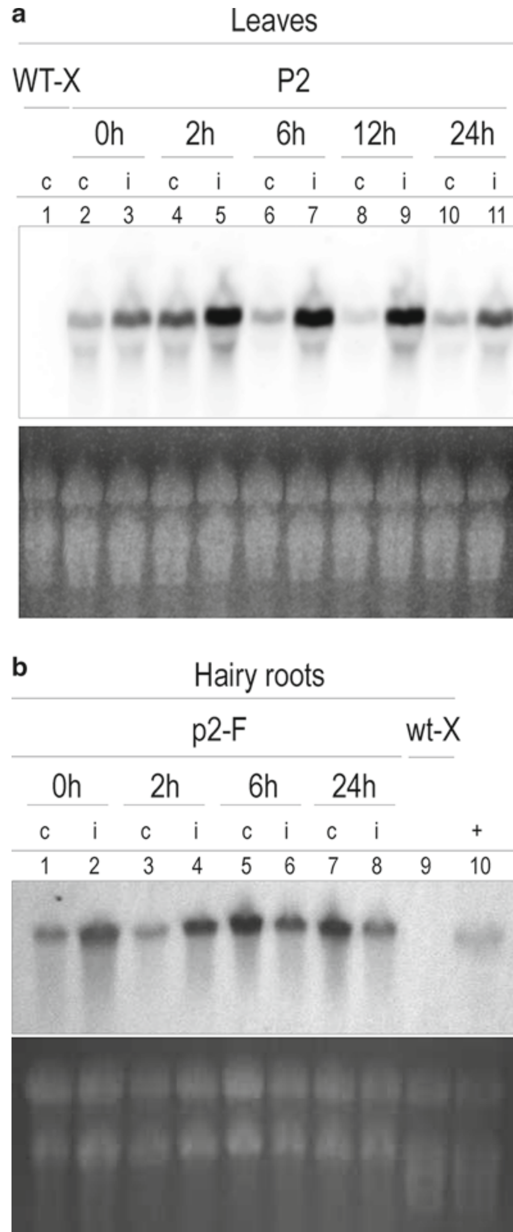


Fig. 9. Northern blotting of mGFP5 mRNA in both mature leaves (4-week-old plants) and hairy roots (14-day-old cultures) induced by wounding. (**a**) Levels of *mGFP5* mRNA in **P2** mature leaves 0, 2, 6, 12, and 24 h post induction of the super–promoter. (**b**) Amount of *mGFP5* mRNA in **p2-F** hairy roots 0, 2, 6, and 24 h after wounding. *WT-X* nontransformed Xanthi leaves, *wt-X* hairy roots derived from WT-X, *i* wound-induced tissues, *c* control, nonwounded tissues, + mGFP5 probe. In the lower blocks of **a** and **b**, the figure shows equal loading of 20 μg total mRNA per lane.

analyses, 24 h after wounding, this increase was 1.3-fold higher compared to nonwounded leaves. A Western blot showed two bands of 27 and 30 kDa, which matched the characteristic two bands of the standard mGFP5 in the 30-kDa region (Fig. 10).

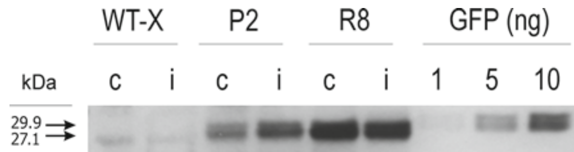


Fig. 10. Accumulation of mGFP5 24 h after mature leaf wounding. Wound induction (i) was done in leaves from 4-week-old in vitro grown P2 tobacco plants and compared to noninduced leaves (c). Western blot is shown for 50 μg TSP per well. Nontransformed Xanthi (WT-X) and R8 were used as negative and positive controls, respectively. Commercial GFP at 1, 5, and 10 ng/well was used as detection control.

An overall comparison of *mGFP5* mRNA levels indicates that the super-promoter drives a higher transcription of *mGFP5* gene in intact hairy roots of both 14- and 21-day-old cultures than in leaves, even upon mature leaf wounding. Wounding of 14-day-old hairy roots did not increase significantly the *mGFP5* mRNA levels.

In this study, we present evidence that super-promoter:TEV is a strong candidate for inclusion in gene constructs to drive overexpression of foreign proteins in tobacco hairy roots. A 14-day hairy root culture (mid-log growth phase) showed a higher amount of *mGFP5* mRNA than 21-day cultures (stationary growth phase). Likewise, *mGFP5* mRNA levels were higher in young than mature leaves (Fig. 4). However, higher levels of mGFP5 protein were observed in mature leaves as shown by Western blot (Fig. 6). The observation that the levels of *mGFP5* mRNA were higher in 14-day cultures of hairy roots than in young tobacco leaves suggests that the super-promoter may be particularly useful in a hairy root production system.

The analysis of the hairy root line p8-B8 grown in liquid culture for 2 weeks yielded 10.4 μg mGFP5 from hairy root tissue, plus a total of 1.35 μg mGFP5 of secreted protein. By comparison, the P8 transgenic tobacco plant from which p8-B8 was derived yielded 420 and 1,026 ng mGFP5/g FW in young and mature leaves, respectively, after 4 weeks of plant culture. Even though the mRNA levels of *mGFP5* were higher in young leaves than in mature leaves (Fig. 4), there was a higher amount of mGFP5 protein accumulated in mature leaves. However, the levels of *mGFP5* mRNA and protein were always higher in hairy roots, with 11% of the recombinant protein produced in the hairy root system present in the growth medium. An additional advantage of hairy roots as a recombinant protein production system is that protein purification from hairy root medium may be expected to be simple compared to extraction and purification from whole plant tissues. Currently, different strategies, such as the addition of salts into the medium, are being tested to increase the secretion of the recombinant mGFP5 protein.

Our results support the use of the super-promoter:TEV in a hairy root production system for overexpression of important recombinant proteins, such as those used as biopharmaceuticals (24, 25),

or to express enzymes involved in the biosynthetic pathways of valuable secondary metabolites (26). Moreover, hairy roots are genetically very stable and its culture in contained conditions (i.e., bioreactors) is a benefit for a long-term production of bioactive molecules. A further advantage is the wound inducibility of this promoter system in hairy root tissue. Additional exploration of inducible mechanisms of this promoter leads to increased yield of recombinant proteins. These approaches combined with novel technologies for scale-up of hairy roots in bioreactors along with associated purification strategies of recombinant proteins further the use of the super-promoter:TEV expression system from the laboratory scale to commercial production of valuable therapeutic proteins.

4. Notes

1. Formaldehyde (methanal) is a denaturing agent for RNA. It prevents RNA to form secondary structures. Formaldehyde can be toxic, allergenic, or carcinogenic. Handle it in a fume hood and wear gloves.
2. An adenine was added immediately downstream to the *Kpn*I site and upstream of the DNA sequence encoding the initial methionine of the patatin signal peptide. Therefore, the encoded protein had an MVP amino acid sequence at the N-terminus, before the initial methionine of *pat*.
3. Growth conditions need to be optimized for each species. The effect of light versus dark culture conditions on root growth was not evaluated in this study.
4. As root tissue develops, it is expected that pH alkalinizes and conductivity value reduces as salts in medium are incorporated into root tissue.
5. DNA concentration using this method is in the range of 10–30 ng/μL (tobacco leaf tissue). Then, yield is between 1.5 and 4.5 ng DNA/mg tissue (considering 200 mg fresh tobacco leaves as starting material).
6. To prepare 100 mL of gel solution for RNA analysis, mix 72 mL DEPC-treated water, 10 mL 10× MESA buffer, and 1.2 g agarose and boil. Let cool down to 65°C (hand tolerant) and add 18 mL of a 37% formaldehyde solution (commercially available) in a safety hood. The final formaldehyde concentration is 2.22 M.
7. Roche recommends using 20 mL DIG Easy Hyb solution per 100-cm^2 membrane. According to manufacturer's procedures, 6 μg DIG-labeled probe is contained in 40 μL labeling reaction, and this is the recommended amount for 20 mL hybridization solution.

8. *Protein size*. The recombinant protein (MVPpat:mGFP5) obtained from tobacco tissues may correspond to an incompletely processed protein. Vector NTI software (Invitrogen) predicted the premature protein to be composed of 266 amino acids, with an estimated molecular weight of 29.9 kDa and pI 6.07. Using both Neural Networks and Hidden Markov Models (27, 28), SignalP predicts a cleavage between adenine and threonine at the C-terminal side of the *pat:mGFP5* sequence (vertical arrow in Fig. 2d), resulting in a protein of 240 amino acids with an estimated molecular weight of 27.1 kDa and pI 5.93 for mGFP5 according to Vector NTI. This cleavage is consistent with reports on the size of the mature peptide of patatin (29). Both MVPpat:mGFP5 and mGFP5 may be present in protein extracts of whole plant cells (leaves and hairy roots), and this could explain the reason why we observed the 29.9- and 27.1-kDa bands in Western blots (Fig. 8b). In the case of rGFPuv used as standard, a similar doublet was also observed, but this may be explained by the presence of a 6-His tag which has been added to the coding sequence (Clontech, personal communication). When the culture medium of the hairy root cultures was analyzed, we observed only the 27.1-kDa band, which may correspond to the fully processed protein expected to be present in the apoplast and medium following the secretory pathway (19).
9. *Autofluorescence in tobacco*. We first attempted to quantify mGFP5 by detection of fluorescence using different dilutions of protein extracts from transgenic tobacco leaves, but found this technique extremely unreliable. To do this, we used a commercially available rGFP-uv as standard, which has the same excitation wavelength as mGFP5. Fluorescence readings were very consistent for rGFP-uv protein alone, but no correlation was found when a nontransformed tobacco plant protein extract was added as background to the standard curve. Phenolic compounds and other secondary metabolites naturally present in tobacco tissues could be responsible for this background that may interfere with quantification of mGFP5 protein in transgenic tobacco using this fluorescence-based detection method (data not shown). Our assessments of mGFP5 protein expression results were significantly more consistent when a sandwich ELISA method was used.
10. *Tissue wounding*. Nishiuchi et al. (30) reported that elements of the super-promoter sequence are involved in transcriptional repression of the ethylene-responsive factor 3 (ERF3) gene following by wounding of tobacco tissue. Our experiments with tissue wounding may also indicate that these elements in the super-promoter are regulated by wounding in both leaves and hairy roots (Fig. 9). Interestingly, the super-promoter responds

immediately after leaf or root wounding by accumulating *mGFP5* mRNA, possibly as a response of the repair machinery to injury. In leaves, we found that the amount of *mGFP5* mRNA in nonwounded tissue reached a peak 2 h after excising the leaf from the plant, and then decreased. In wounded p2 leaves, the mRNA levels increased immediately after wounding, and reached a plateau between 2 and 12 h after wounding (Fig. 9a). In the case of hairy roots, the induction by wounding was higher than nonwounded controls at time zero and even 2 h after wounding. The inducibility by wounding of the super-promoter in leaves and hairy roots is a potentially convenient and advantageous condition when using this system as a production tool for recombinant proteins. As induction by wounding is being analyzed, special care must be taken into account when harvesting plant tissue to avoid unnecessary damage.

Acknowledgments

The authors would like to thank Dr. Gastón Zolla for his assistance with some of the assays, and also to the members of our laboratories at the Virginia Polytechnic Institute and State University (Blacksburg, VA) and at the Arkansas Biosciences Institute at the Arkansas State University (Jonesboro, AR) for their contributions at different stages of this work. This research was supported by the State of Virginia Commonwealth Technological Research Fund and the Arkansas Biosciences Institute.

References

1. Wildi, E., Wildi, R., and Ripplinger, P. (2004) Device for cultivating plant or animal tissue cultures. US Patent **6**, 794–183.
2. Guillon, S., Trémouillaux-Guiller, J., Pati, P. K., Rideau, M., and Gantet, P. (2006) Harnessing the potential of hairy roots: dawn of a new era. *TRENDS Biotechnol.* **24**, 403–409.
3. Liu, J., Dolan, M., Reidy, M., Cramer, C. (2008) Expression of bioactive single-chain murine IL-12 in transgenic plants. *J. Interferon Cytokine Res.* **28**, 381–92.
4. Medina-Bolivar, F., Wright, R., Funk, V., Sentz, D., Barroso, L., Wilkins, T. D., Petri, W. Jr., and Cramer, C. L. (2003) A non-toxic lectin for antigen delivery of plant-based mucosal vaccines. *Vaccine* **21,** 997–1005.
5. Ni, M., Cui, D., Einstein, J., Narasimhulu, S., Vergara, C. E., and Gelvin, S. B. (1995) Strength and tissue specificity of chimeric promoters derived from the octopine and mannopine synthase genes. *Plant J.* **7,** 661–676.
6. Leisner, S. M., and Gelvin, S. B. (1988) Structure of the octopine synthase upstream activator sequence. *Proc. Natl. Acad. Sci. USA* **85,** 2553–2557.
7. Ellis, J. G., Ryder, M. H., and Tate, M. E. (1984) *Agrobacterium tumefaciens* TR-DNA encodes a pathway for agropine biosynthesis. *Mol. Gen. Genet.* **195**, 466–473.
8. Lee, L. Y., Kononov, M., Bassuner, B., Frame, B. R., Wang, K., and Gelvin, S. B. (2007) Novel plant transformation vectors containing the super-promoter. *Plant Physiol.* **145,** 1294–1300.
9. Higo, K., Ugawa, Y., Iwamoto, M., and Korenaga, T. (1999) Plant cis-acting regulatory DNA elements (PLACE) database:1999. *Nucleic Acids Res* **27**, 297–300. http://www.dna.affrc.go.jp/PLACE/. Accessed 05 Feb 2007.

10. Ellis, J. G., Llewellyn, D. J., Walker, J. C., Dennis, E. S., and Peacock W. J. (1987) The ocs element: a 16 base pair palindrome essential for activity of the octopine synthase enhancer. *EMBO J.* **6,** 3203–3208.
11. Singh, K., Tokuhisa, J. G., Dennis, E. S., and Peacock, W. J. (1989) Saturation mutagenesis of the octopine synthase enhancer: correlation of mutant phenotypes with binding of a nuclear protein factor. *Proc. Natl. Acad. Sci. USA* **86,** 3733–3737.
12. Fromm, H., Katagiri, F., and Chua, N. H. (1989) An octopine synthase enhancer element directs tissue-specific expression and binds ASF-1, a factor from tobacco nuclear extracts. *Plant Cell* **1,** 977–984.
13. Guevara-García, A., López-Ochoa, L., López-Bucio, J., Simpson, J., and Herrera-Estrella, L. (1998) A 42 bp fragment of the *pmas1*' promoter containing an *ocs*-like element confers a developmental, wound- and chemically inducible expression pattern. *Plant Mol. Biol.* **38,** 743–753.
14. Guevara-García, A., Mosqueda-Cano, G., Argüello-Astorga, G., Simpson, J., and Herrera-Estrella, L. (1993) Tissue-specific and wound-inducible pattern of expression of the mannopine synthase promoter is determined by the interaction between positive and negative cis-regulatory elements. *Plant J.* **4,** 495–505.
15. Shen, W. H., Davioud, E., David, C., Barbier-Brygoo, H., Tempé. J., and Guern, J. (1990) High sensitivity to auxin is a common feature of hairy root. *Plant. Physiol.* **94,** 554–560.
16. Carrington, J. C., and Freed, D. D. (1990) Cap-independent enhancement of translation by a plant potyvirus 5 nontranslated region. *J. Virology* **64,** 1590–1597.
17. Carrington, J. C., Freed, D. D., and Leinicke, A. J. (1991) Bipartite signal sequence mediates nuclear translocation of the plant potyviral NIa protein. *The Plant Cell* **3,** 953–962.
18. Mason, H. S., Ball ,J. M., Shi, J. J., Jiang, X., Estes, M. K., Arntzen, C. J. (1996) Expression of Norwalk virus capsid protein in transgenic tobacco and potato and its oral immunogenicity in mice. *Proc. Natl. Acad. Sci. USA* **93,** 5335–5340.
19. Denecke, J., Botterman, J., and Deblaere, R. (1990) Protein secretion in plant cells can occur via a default pathway. *The Plant Cell* **2,** 51–59.
20. Medina-Bolivar, F., and Cramer, C. (2004) Production of recombinant proteins in hairy roots cultured in plastic sleeve bioreactors In: P. Balbas and A. Lorence (eds) Recombinant Gene Expression: Reviews and Protocols, Humana Press, Totowa, pp. 351–363.
21. Gamborg, O., Miller, R. A., and Ojima, K. (1968) Nutrient requirements of suspension cultures of soybean root cells. *Exp. Cell. Rep.* **50,** 151–158.
22. Holsters, M., de Waele, D., Depicker, A., Messens, E., van Montagu, M., and Schell, J. (1978) Transfection and transformation of *A. tumefaciens. Mol. Gen. Genet.* **163,** 181–187.
23. Richards, H. A., Halfhill, M. D., Millwood, R. J., and Stewart, C. N. Jr. (2003) Quantitative mGFP5 fluorescence as an indicator of recombinant protein synthesis in transgenic plants. *Plant Cell Rep.* **22,** 117–121.
24. Shadwick, F. S., and Doran, P. M. (2007) Propagation of plant viruses in hairy root cultures: A potential method for *in vitro* production of epitope vaccines and foreign proteins. *Biotechnol. Bioeng.* **96,** 570–583.
25. Sharp, J. M., and Doran, P. M. (2001) Characterization of monoclonal antibody fragments produced by plant cells. *Biotechnolog. Bioeng.* **73,** 338–346.
26. Köhle, A., Sommer, S., Yazaki, K., Ferrer, A., Boronat, A., Li, S. M., and Heide, L. (2002) High level expression of chorismate pyruvate-lyase (UbiC) and HMG-CoA reductase in hairy root cultures of *Lithospermum erythrorhizon. Plant Cell. Physiol.* **43,** 894–902.
27. Bendtsen, J. D., Nielsen, H., von Heijne, G., and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* **340,** 783–795 http://www.cbs.dtu.dk/services/SignalP. Accessed 05 Feb 2007.
28. Nielsen, H., and Krogh, A. (1998) Prediction of signal peptides and signal anchors by a hidden Markov model. Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology (ISMB 6), AAAI Press, Menlo Park, California, pp. 122–130. http://www.cbs.dtu.dk/services/SignalP. Accessed 05 Feb 2007.
29. Mignery, G. A., Pikaard, C. S., Hannapel, D. J., Park, W. D. (1984) Isolation and sequence analysis of cDNAs for the major potato tuber protein, patatin. *Nucleic Acids Res.* **12,** 7987–8000.
30. Nishiuchi, T., Shinshi, H., and Suzuki, K. (2004) Rapid and transient activation of transcription of the ERF3 gene by wounding in tobacco leaves: Possible involvement of NtWRKYs and autorepression. *J. Biol. Chem.* **279,** 55355–55361.

Chapter 28

Bioseparation of Recombinant Proteins from Plant Extract with Hydrophobin Fusion Technology

Jussi J. Joensuu, Andrew J. Conley, Markus B. Linder, and Rima Menassa

Abstract

Two main hurdles hinder the widespread acceptance of plants as a preferred protein expression platform: low accumulation levels and expensive chromatographic purification methods. Fusion of proteins of interest to fungal hydrophobins has provided a tool to address both accumulation and purification issues. In this method, we describe the one-step purification of a GFP–HFBI fusion from crude plant extract using an aqueous two-phase system (ATPS). ATPS can be carried out in a very short time frame, yields relatively pure protein with very few contaminants, and does not require any chromatographic column steps. This purification system takes advantage of the affinity of hydrophobins to the micellar phase of widely available nonionic surfactants, such as Triton X-114, and can be easily scaled up for industrial-scale protein purification.

Key words: Aqueous two-phase system, ATPS, Hydrophobin, HFBI, Protein purification, Protein fusion, Molecular farming

1. Introduction

1.1. Rationale

For the past two decades, plants have shown promise as bioreactors for the large-scale production of various therapeutic and industrially important recombinant proteins (1). To obtain high-recombinant protein yields, *Agrobacterium* infiltration in *Nicotiana benthamiana* leaves (2), when combined with the coexpression of a suppressor of gene silencing (3), has widely established itself as the most utilized transient expression system in plants. Furthermore, this flexible technology allows for rapid gram-quantity production of recombinant protein within a couple of weeks (4, 5), which is a significant advantage over conventional production technologies.

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_28, © Springer Science+Business Media, LLC 2012

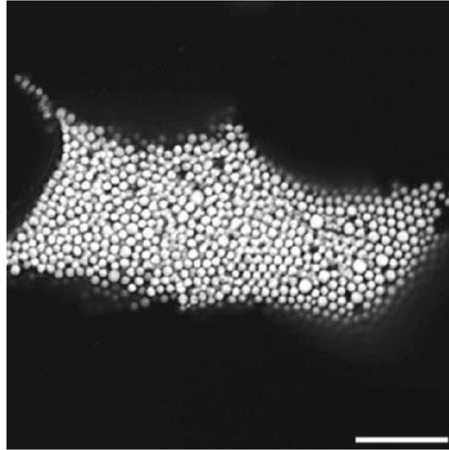


Fig. 1. Expression of an ER-targeted GFP–HFBI fusion promotes the formation of protein bodies (PBs). A leaf epidermal cell accumulating GFP–HFBI in PBs was imaged with a Leica TCS SP2 confocal laser scanning microscope equipped with a 63× water-immersion objective. For the imaging of GFP expression, excitation with a 488-nm argon laser was used and fluorescence was detected at 500–525 nm. Bar = 10 μm.

However, the recovery of recombinant proteins from the complex plant proteome has remained challenging, which can contribute to greater than 80% of the product cost (6). Traditionally, affinity chromatography is used for the purification of recombinant proteins, but this method is typically only suitable for purifying low quantities of high-value products. To address this limitation, we investigated the use of hydrophobin fusion technology, which has shown promise as a simple, rapid, scalable, and inexpensive nonchromatographic means of purifying plant-made recombinant proteins (7).

Hydrophobins are small and surface-active fungal proteins capable of altering the hydrophobicity of their respective fusion partner, thus enabling efficient purification using a surfactant-based aqueous two-phase system (ATPS) (8, 9). In an ATPS, a surfactant is added to crude protein extracts which concentrates the hydrophobin fusions inside micellar structures and partitions them toward the surfactant phase while the majority of the proteins remain in the aqueous phase (10). Most importantly, the one-step ATPS purification is particularly attractive because it can be easily and effectively scaled up for industrial-scale protein purification (9, 11).

In addition to the role as a purification aid, hydrophobins have also been shown to enhance the accumulation of recombinant proteins in plants through the formation of novel protein bodies (PBs) (7) (Fig. 1). Presumably, the packing of these fusion proteins into PBs protects them from proteolytic degradation associated with normal physiological turnover (12). PBs are stable, intracellular storage organelles which enable high local concentrations of

heterologous proteins to exist within the confined space of the cell, without subjecting the host cell to an intolerable level of stress (13). In summary, hydrophobin fusion technology provides a general method for enhancing the yield of recombinant proteins in plants while also providing a simple, efficient, and inexpensive approach for their subsequent purification.

1.2. Nonionic Surfactants in ATPS

There are several different forms of ATPS and they can be based on various polymers or surfactants (14). The type of ATPS that has been the most successful with hydrophobins involves the use of nonionic surfactants that phase separate in a temperature-dependant manner (15). Nonionic surfactants are molecules that have two different components (as surfactants in general): one is hydrophilic and the other is hydrophobic. In nonionic surfactants, the hydrophilic moiety does not have a net electric charge (in contrast to most common detergents, such as SDS); instead, it is a chain of a hydrophilic polymer. Polyethoxy chains are often used as the hydrophilic moiety, whereas long-chain alcohols often serve as the hydrophobic moiety (16). Surfactants, such as Tween and Triton, that are commonly used as laboratory chemicals are typical nonionic surfactants. However, not all nonionic surfactants are useful in ATPS. For a functioning ATPS, conditions must exist, where the aqueous surfactant mixture forms a two-phase system. The hydrophilic polyethoxy chain is useful in this context. The solubility of the polyethoxy chain decreases with increased temperature. This "reversed" solubility property most likely stems from conformational changes in the polymer that decreases its interaction with water molecules. For the nonionic surfactant to be useful in ATPS, the temperature at which this phase separation occurs must be in the practical range of 5–30°C (17). There are several such surfactants, one example being Triton X-114.

It has not been thoroughly studied why hydrophobins are so efficiently separated in two-phase systems. However, it is likely that the hydrophobins are somehow compatible with the micellar structure of the surfactant phase. It is also well-known that membrane proteins are efficiently purified by ATPS using nonionic surfactants (18). However, a significant difference between membrane proteins and hydrophobins is that hydrophobins are highly soluble in water while membrane proteins are not. There is also an interesting connection between the size of the fusion partner linked to the hydrophobin and which type of surfactant it most efficiently partitions into. Large fusion partners tend to partition more efficiently into surfactants that have longer hydrophilic polyethoxy chains (8). Most likely, such surfactants are more hydrated and can accommodate the hydrophilic proteins better.

2. Materials

All solutions and equipment should be clean, sterile, and detergent free. Water should be deionized, distilled, and autoclaved (conductivity <10 μS). Eye protection, lab coat, and insulated gloves should be used when handling liquid nitrogen.

2.1. Protein Extraction from Tobacco Leaves

1. Tobacco leaves expressing (transient or stable expression) a protein of interest fused to a C-terminal hydrophobin HFBI tag. see Fig. 2 for construct details. For transient expression in *N. benthamiana*, see ref. 7.
2. Laboratory scale.
3. Ceramic mortar and pestle.
4. Liquid nitrogen and suitable insulated container for handling.
5. Phosphate-buffered saline (PBS) extraction buffer: 137 mM NaCl, 2.7 mM KCl, 8.1 mM Na_2HPO_4, and 1.8 mM KH_2PO_4; pH 7.4.
6. Refrigerated centrifuge.

2.2. Aqueous Two-Phase Separation

1. Eppendorf or other plastic test tubes/separation funnel (for volumes over 25 mL).
2. Triton X-114 surfactant (Sigma–Aldrich).
3. Isobutanol.
4. Temperature-adjustable water bath.
5. Laboratory pipettes and suitable tips for liquid handling.
6. Vortex mixer.

2.3. SDS-PAGE Gel Analysis

1. Tris–HCl Polyacrylamide Ready gel 4–15% (Bio-Rad).
2. SDS-PAGE running buffer: 25 mM Tris, 200 mM glycine, and 0.1% (w/v) SDS.
3. MiniProtean III SDS-PAGE electrophoresis apparatus (Bio-Rad).

Fig. 2. Schematic representation of the expression cassette that was used for transient expression of GFP–HFBI fusion protein in *Nicotiana benthamiana* leaves. p35S, enhanced cauliflower mosaic virus 35S promoter; tCUP, translational enhancer; Ss, Pr1b tobacco secretory signal peptide; GFP, green fluorescent protein gene; HFBI, *Trichoderma reesei* hydrophobin HFBI gene; $(G_3S)_3$, flexible 12 amino acid linker that was inserted between GFP and HFBI; KDEL, endoplasmic reticulum retention signal; vsp3′, soybean vegetative storage protein terminator.

4. 4× SDS-PAGE loading buffer: 240 mM Tris–HCl, pH 6.8, 40% glycerol, 8% (w/v) SDS, 0.1% (w/v) bromophenol blue, and 5% (v/v) 2-mercaptoethanol.
5. GelCode Blue Stain Reagent (Thermo Scientific).

3. Methods

3.1. Protein Extraction from Tobacco Leaves

1. Weigh 1 g of freshly harvested tobacco leaves expressing hydrophobin fusion protein (see Note 1). Put leaves into a mortar and cover them with a generous amount of liquid nitrogen. Precool the pestle with liquid nitrogen and crunch leaves to small pieces. Once the excess of liquid nitrogen has evaporated, grind leaves thoroughly to a fine powder.
2. Set the mortar on ice and let the powder thaw before adding 10 mL of PBS. Keep grinding for at least 5 min to release total soluble proteins (TSPs) from the leaves.
3. Transfer the leaf extract to precooled test tubes and centrifuge for 5 min at 20,000 × *g* at 4°C.
4. Collect the supernatant to new tubes and repeat centrifugation.
5. Collect the supernatant. Keep the leaf extract on ice before proceeding to the ATPS.
6. For SDS-PAGE samples, take 75 μL of the leaf extract and mix it with 25 μL of 4× SDS-PAGE loading buffer (Sample 1, Fig. 3).

3.2. Aqueous Two-Phase Separation

1. Weigh 40 mg of Triton X-114 to a 1.5-mL Eppendorf tube (see Note 2).
2. Preheat 1 mL of leaf extract in a 24°C water bath for 5 min (see Note 3).
3. Add the leaf extract to an Eppendorf tube containing Triton X-114.
4. Vortex the extract/Triton X-114 mixture thoroughly for 1 min (see Note 4).
5. Set the tube back to 24°C water bath for 5 min to separate the phases (see Note 5, Fig. 3).
6. Pick up the tube carefully, take 75 μL of upper phase to be analyzed on SDS-PAGE, and mix it with 25 μL of 4× SDS-PAGE loading buffer (Sample 2, Fig. 3).
7. Carefully collect the lower phase with a thin gel-loading tip through the upper phase. Start pipetting from the bottom of the tube. Transfer this lower phase containing the hydrophobin fusion protein to a new tube (see Note 6).
8. Add 400 μL of isobutanol to the collected lower phase and mix well by vortexing (see Note 7).

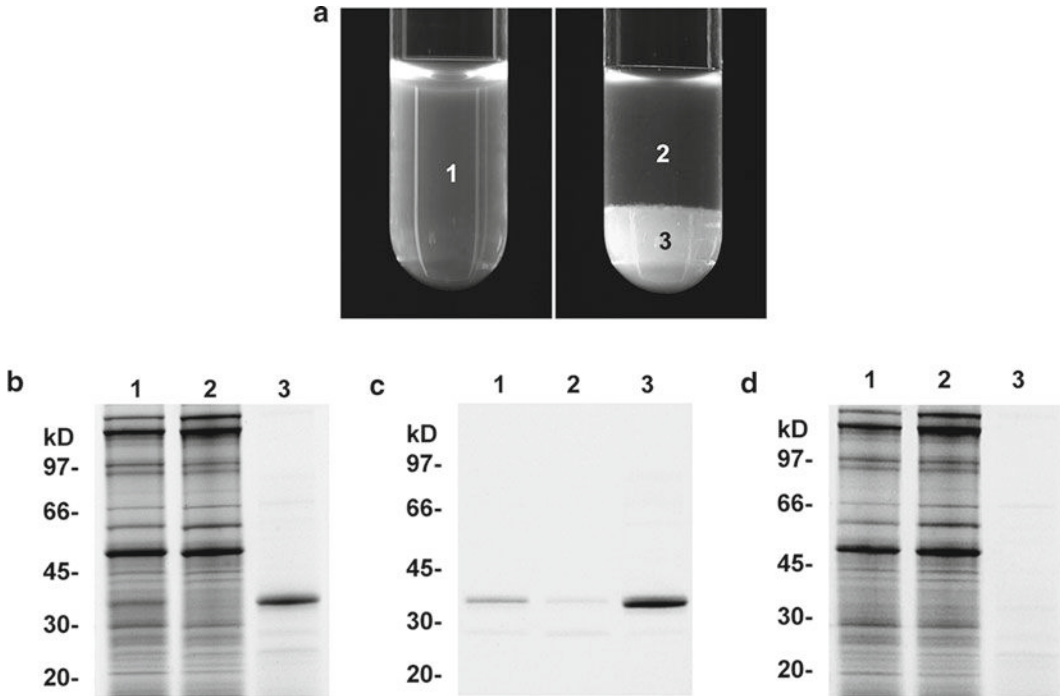


Fig. 3. Aqueous two-phase separation (ATPS) of the GFP–HFBI fusion protein from plant extract with surfactant Triton X-114. (**a**) Photographs of the GFP–HFBI leaf extract before (*left*) and after (*right*) the phase separation indicating the accumulation of GFP–HFBI in the lower surfactant phase. The tubes were imaged under ultraviolet light. The numbering corresponds to lane numbers in (**b**), (**c**), and (**d**). (**b**) SDS-PAGE demonstrates a selective recovery of the GFP–HFBI fusion protein (37 kDa) in the lower phase with minimal background from the leaf extract. Equal volumes of the following samples were loaded onto the gel: 1, leaf extract containing GFP–HFBI before addition of Triton X-114; 2, upper phase after ATPS; 3, lower phase after ATPS. Similar experiments were performed with purified GFP–HFBI protein (**c**) and wild-type leaf extracts (**d**). They confirm the concentration of GFP–HFBI protein by ATPS (**c**) with Triton X-114 while extracting out very few contaminating plant proteins (**d**).

9. Set the tube to a 24°C water bath for 5 min to separate the phases (see Note 8).
10. Collect the lower phase containing the hydrophobin fusion protein to a new tube as mentioned above (see Note 9).
11. Take 75 μL sample to be analyzed on SDS-PAGE and mix it with 25 μL of 4× SDS-PAGE loading buffer (Sample 3, Fig. 3).

3.3. SDS-PAGE Gel Analysis

1. Boil the collected samples 1, 2, and 3 (75 μL+25 μL of 4× SDS PAGE loading buffer) for 5 min.
2. Prepare a polyacrylamide gel of appropriate percentage and load 20 μL of the samples per well.
3. Run gel at 200 V until the dye front reaches the end of the gel.
4. Disassemble the running apparatus and desalt the gel twice for 15 min with H_2O.
5. Stain the gel overnight with GelCode Blue Stain Reagent.
6. Destain the gel with H_2O for 1 h.

4. Notes

1. In this example, *N. benthamiana* leaf extract expressing GFP–HFBI fusion protein was used (Fig. 3).
2. The amount of surfactant may vary between 2 and 10% (w/v) depending on which volumes are practical to handle. If the volume of supernatant is large (over 100 mL), a smaller amount of surfactant may be used. The concentrating effect of the ATPS is better when a smaller volume of surfactant is used, although higher amounts of surfactant tend to give higher total recovery of the hydrophobin fusion. The scale of the separation can be flexibly adjusted. We have used the same protocol for volumes from 100 μL to 1,500 L.
3. This example is for extracts from plant leaves, but needs only slight modification for applying to any homogenized cell extract or cell culture supernatant.
4. The sample should be carefully mixed until all surfactant is well-dispersed. For larger volumes, allow the sample to stand for 20–60 min or until the surfactant has clearly separated into its own phase. In the case of Triton X-114, the surfactant phase is the lower one. For other nonionic surfactants, it can be the top phase. If volumes are 20–500 mL, the separation is most easily performed in a separation funnel.
5. The sample can be centrifuged to speed up clarification. Separation times might be longer if the samples contain a lot of dispersed cell debris.
6. When working with a separation funnel, the lower phase containing the protein of interest can be easily drained from the bottom valve.
7. The volume of isobutanol should be about ten times the volume of surfactant originally added. For example, if 1 mL of surfactant was added, then 10 mL of isobutanol should be used. For scaling-up, a smaller amount of isobutanol can be used, but the volume must be optimized separately.
8. Isobutanol and surfactant are mixed carefully and allowed to separate. The solution separates into two phases. One phase (the upper one) contains a mixture of isobutanol and surfactant. The other phase is an aqueous phase containing the hydrophobin fusion protein. The water in this phase is separated from the surfactant by the action of isobutanol. This means that when the surfactant was separated from the supernatant in the previous step some water was entrapped within the surfactant. The separation is usually complete within minutes. If the aqueous phase is very small (much less than one-tenth of the original supernatant), the yield may be improved by adding some extraction buffer (or water) to the isobutanol extraction step.

9. The target fusion protein is in the aqueous phase that is drawn off from the isobutanol–surfactant mixture. It is typically quite pure from other proteins (Fig. 3), but this aqueous solution is saturated with isobutanol. The further use of the target protein may require removal of residual isobutanol. This can be achieved easily by filtration on a column, such as the BioGel P-6 (Bio-Rad) or equivalent.

References

1. Ma, J. K., Barros, E., Bock, R., et al., (2005) Molecular farming for new drugs and vaccines. Current perspectives on the production of pharmaceuticals in transgenic plants, *EMBO Rep.* **6**, 593–599.
2. Kapila, J., De Rycke, R., Van Montagu, M., et al., (1997) An Agrobacterium-mediated transient gene expression system for intact leaves, *Plant Sci.* **122**, 101–108.
3. Silhavy, D., Molnar, A., Lucioli, A., et al.,(2002) A viral protein suppresses RNA silencing and binds silencing-generated, 21- to 25-nucleotide double-stranded RNAs, *EMBO J* **21**, 3070–3080.
4. Giritch, A., Marillonnet, S., Engler, C., et al., (2006) Rapid high-yield expression of full-size IgG antibodies in plants coinfected with noncompeting viral vectors, *Proc. Natl. Acad. Sci. USA* **103**, 14701–14706.
5. Vezina, L. P., Faye, L., Lerouge, P., et al., (2009) Transient co-expression for fast and high-yield production of antibodies with human-like N-glycans in plants, *Plant Biotechnol. J.* **7**, 442–455.
6. Kusnadi, A. R., Nikolov, Z. L., and Howard, J. A. (1997) Production of recombinant proteins in transgenic plants: Practical considerations, *Biotechnol. Bioeng.* **56**, 473–484.
7. Joensuu, J. J., Conley, A. J., Lienemann, M., et al., (2010) Hydrophobin fusions for high-level transient protein expression and purification in *Nicotiana benthamiana*, *Plant Physiol.* **152**, 622–633.
8. Linder, M. B., (2009) Hydrophobins: Proteins that self assemble at interfaces**,** *Current Opinion in Colloid & Interface Science*, **14**, 356–363
9. Linder, M. B., Qiao, M., Laumen, F., Set al., (2004) Efficient purification of recombinant proteins using hydrophobins as tags in surfactant-based two-phase systems, *Biochemistry* **43**, 11873–11882.
10. Linder, M., Selber, K., Nakari-Setälä, T., et al., (2001) The hydrophobins HFBI and HFBII from *Trichoderma reesei* showing efficient interactions with nonionic surfactants in aqueous two-phase systems, *Biomacromolecules* **2**, 511–517.
11. Selber, K., Tjerneld, F., Collen, A., et al., (2004) Large-scale separation and production of engineered proteins, designed for facilitated recovery in detergent-based aqueous two-phase extraction systems, *Proc. Biochem.* **39**, 889–896.
12. Vitale, A., and Ceriotti, A. (2004) Protein quality control mechanisms and protein storage in the endoplasmic reticulum. A conflict of interests?, *Plant Physiol.* **136**, 3420–3426.
13. Galili, G. (2004) ER-derived compartments are formed by highly regulated processes and have special functions in plants, *Plant Physiol.* **136**, 3411–3413.
14. Quina, F. H., and Hinze, W. L. (1999) Surfactant-mediated cloud point extractions: An environmentally benign alternative separation approach, *Ind. Engineer. Chem. Res.* **38**, 4150–4168.
15. Walter, H., and Johansson, G., (eds.) (1994) *Aqueous Two-Phase Systems*, In *Methods in Enzymology* **228**, Academic Press, San Diego; Toronto.
16. Holmberg, K. (2003) *Surfactants and polymers in aqueous solution*, 2nd ed., John Wiley & Sons, Chichester, West Sussex, England; Hoboken, NJ.
17. Terstappen, G. C., Ramelmeier, R. A., and Kula, M.-R. (1993) Protein partitioning in detergent-based aqueous two-phase systems, *J. Biotechnol.* **28**, 263–275.
18. Sivars, U., and Tjerneld, F. (2000) Mechanisms of phase behaviour and protein partitioning in detergent/polymer aqueous two-phase systems for purification of integral membrane proteins, *Biochim. Biophys. Acta* **1474**, 133–146.

Chapter 29

Quality Assessment of Recombinant Proteins Produced in Plants

Giuliana Medrano, Maureen C. Dolan, Jose Condori, David N. Radin, and Carole L. Cramer

Abstract

Plant-based expression technologies for recombinant proteins have begun to receive acceptance for pharmaceuticals and other commercial markets. Protein products derived from plants offer safer, more cost-effective, and less capital-intensive alternatives to traditional manufacturing systems using microbial fermentation or animal cell culture bioreactors. Moreover, plants are now known to be capable of expressing bioactive proteins from a diverse array of species including animals and humans. Methods development to assess the quality and performance of proteins manufactured in plants are essential to support the QA/QC demands as plant-produced protein products transition to the commercial marketplace. Within the pharmaceutical arena, process validation and acceptance criteria for biological products must comply with the Food and Drug Administration (FDA) and ICH Q6B guidelines in order to initiate the regulatory approval process. Detailed product specifications will also need to be developed and validated for plant-made proteins for the bioenergy, food, chemical synthesis, or research reagent markets.

We have, therefore, developed assessment methods for important qualitative and quantitative parameters of the products and the manufacturing methods utilized in plant-based production systems. In this chapter, we describe a number of procedures to validate product identity and characteristics including mass analyses, antibody cross-reactivity, N-terminal sequencing, and bioactivity. We also address methods for routine assessment of yield, recovery, and purity. The methods presented are those developed for the synthesis and recovery of the avian cytokine, chicken interleukin-12 (ChIL-12), produced in the leaves of *Nicotiana benthamiana*. The ChIL-12 protein used as a model for this chapter includes a C-terminal histidine epitope (HIS-tag) and, thus, these methods may be directly applicable to other HIS-tagged proteins produced in plants. However, the overall strategy presented using the ChIL-12^{HIS} example should provide the basis of standard procedures for assessing the quality of other plant-based protein products and manufacturing systems.

Key words: Plant-based bioproduction, Recombinant protein, Product validation, HIS-tag, Western analyses, ELISA, Endotoxin, Bioactivity, MS-MALDI, Interleukin-12

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_29, © Springer Science+Business Media, LLC 2012

1. Introduction

Plants have emerged as a relatively simple, safe, and economical expression platform for bioproduction of recombinant proteins including those requiring eukaryotic post-translational modifications common in many complex animal proteins of interest for pharmaceutical and veterinary applications. Numerous reports have demonstrated that a wide range of valuable proteins and compounds can be expressed efficiently in plants (1–15). However plant-based manufacturing has yet to be widely adopted in comparison to the more traditional bioproduction platforms employing mammalian cells or microbial expression systems. In part this may be due to a lack of well-developed and clearly defined manufacturing quality assessment protocols for recombinant protein production in plants (16). In order to develop a well-characterized plant-derived protein product, quality control procedures for product identity, recovery, quantification, structure, activity, and in some cases "bioequivalency" are needed. A harmonized FDA and EU guidance document, ICH Q6B (the ICH Guideline Q6B: Test Procedures and Acceptance Criteria for Biotechnological/Biological Products, http://www.ich.org/LOB/media/MEDIA432.pdf), has been developed for protein "biologics" targeted for clinical applications that required current Good Manufacturing Practices (cGMP) compliant manufacturing. While the procedures described herein do not fulfill all the parameters required for cGMP compliance, they do provide an entrée into the tools and analyses required to initiate preclinical studies or to move a recombinant protein to the research reagent market.

This chapter describes the overall workflow and steps required for the assessment of quality parameters for recovery of proteins produced in plants. The workflow described uses a model protein of interest, chicken interleukin-12 with a histidine tag at the C-terminus (ChIL-12^{HIS}) (3), expressed transiently in plants (2). IL-12 is a complex heterodimeric glycoprotein that is responsible for stimulating key immune-responsive cells to product interferon-gamma, a key regulator of cell-mediated immunity and inflammatory responses. Thus, IL-12 has significant potential as an immuno-stimulant and vaccine adjuvant. Recombinant IL-12s have been successfully produced in a single chain form (p40 subunit:linker:p35 subunit:HIS-tag), yielding a product of approximately 70 kDa that elicits strong immune responses both in vitro and in vivo. Our interest in chicken IL-12 is focused on its use as a vaccine adjuvant for feed-based delivery to poultry targeting avian influenza and other viral and intracellular pathogens. The initial cloning and expression of ChIL-12^{HIS} in plants has been previously described (3). In contrast to most human proteins, protein standards and many of the assessment tools simply did not exist for

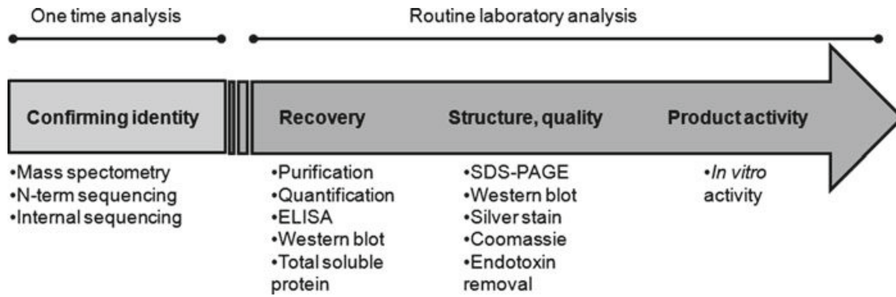


Fig. 1. Workflow for recombinant protein product quality assessments.

this protein when the project was initiated. Therefore we believe ChIL-12^{HIS} provides an excellent model to describe the development of key tools and protocols necessary for implementing effective plant-made protein product assessment strategies.

A very useful tool in facilitating development of detection and purification processes for a new recombinant protein product is incorporation of an epitope tag (for example, the 6×-histidine epitope; HIS-tag). This tag is typically placed on either the N- or C-terminus of the protein; however, it is important to establish that positioning of this tag does not interfere with the function of the protein of interest. Although the long-term goal may be to produce a protein without an epitope tag, we often include a HIS-tag for exploratory work in quantification and purification process development (3, 17). The HIS-tag is widely used in recombinant protein expression in diverse production platforms. We have therefore focused on protocols developed for a HIS-tagged version of the ChIL-12, reasoning that these methods may have broader utility to researchers targeting recombinant protein production.

Quality assessment procedures described here consist of standardized procedures and calculations that we have developed for routinely confirming identity, measuring the quality/quantity of the product as it moves through the downstream processing steps, and establishing function and unit activity (Fig. 1). These assessment procedures include mass spectrometry, N-terminal sequencing, protein purification, ELISA, protein deglycosylation, western immunoblot, silver stained polyacrylamide gel electrophoresis (PAGE), and product-specific activity assays. These methods are designed to provide routine and consistent laboratory analysis check points with the goal of ensuring quality (e.g., product integrity and purity testing), quantity (yield and recovery analyses), and activity (e.g., bioassays protocols) of the plant-made product. Identity confirmation is also included as part of quality assessment procedure. Because of the growing market demand for high-valued proteins, and the potential of plants as production systems for recombinant proteins, the development of quality standards of production compatible with standards for human use is essential.

2. Materials

2.1. Total Protein Extraction ("SDS-grind")

1. 0.5 g of plant tissue.
2. Mortar and pestle.
3. 1.5 mL microcentrifuge tubes.
4. Liquid nitrogen.
5. SDS (sodium dodecyl sulfate) grinding buffer: 150 mM Tris–HCl pH 6.8, 30% glycerol, 6% SDS, 5 mM EDTA pH 8.0 in H_2O. SDS should be of high quality (e.g., electrophoresis/molecular grade SDS from Bio-Rad or Fisher Scientific).

2.2. SDS Polyacrylamide Gel Electrophoresis

1. 4× NuPAGE LDS Sample Buffer (Invitrogen).
2. 4× Bromophenol blue solution: 4% (w/v) Bromophenol blue in dH_2O. Not required if sample buffer, such as LDS Sample buffer, contains running dye.
3. 1 M DTT (dithiothreitol, Research Products).
4. 1× Reducing agent: Dilute 10× NuPAGE sample reducing agent (Invitrogen) 1:10 (v/v) with dH_2O.
5. NuPAGE Novex 10% Bis-Tris precast gels (Invitrogen).
6. 1× SDS-PAGE running buffer: Dilute 20× NuPAGE MOPS/MES SDS running buffer (Invitrogen), 1:20 (v/v) with dH_2O.
7. Molecular weight standards (e.g., Precision Plus Protein, Bio-Rad; stained version for Western blots; unstained for gels visualized with silver- or Coomassie-stains).
8. Mini-format protein electrophoresis unit (e.g., XCell Surelock Electrophoresis Apparatus, Invitrogen).
9. Power supply (e.g., Power Pac Basic, Bio-Rad).

2.3. Western Immunoblot

1. Electrophoretic gel transfer unit (e.g., XCell II™ Blot Module, Invitrogen) and power supply (e.g., Power Pac Basic, Bio-Rad).
2. Electrotransfer buffer: 20 mM Tris–HCl, 150 mM glycine, 20% methanol.
3. Pre-cut nitrocellulose or PVDF membrane (Bio-Rad).
4. Pre-cut filter paper (e.g., Whatman paper).
5. Blotting pads, pre-cut sheet protector.
6. 1× PBS: 137 mM NaCl, 2.7 mM KCl, 10.1 mM Na_2HPO_4, 1.76 mM KH_2PO_4, pH 7.4.
7. Washing buffer: 1× PBS with 0.1% Tween-20.

8. Blocking/dilution buffer: 1× PBS with 3% bovine serum albumin (BSA).
9. Antibody (western analysis): Mouse anti-6×HIS(C-term)-alkaline phosphatase conjugated antibody (e.g., Invitrogen).
10. Platform shaker (e.g., Gyrotory Shaker Model G2; New Brunswick Scientific).
11. Detection reagents: CDP-Star (Boehringer-Mannheim); Nitroblock Enhancer II (Tropix); CDP-Star Detection Buffer (0.1 M Tris–HCl pH 9.5, and 0.1 M NaCl).
12. X-ray film (e.g., Kodak BioMax XAR).
13. X-ray cassette.
14. Darkroom.
15. Fixer and developer.

2.4. Protein Purification for Model His-Tagged Protein

1. Cation exchange resin (e.g., UNOSphere-S cation resin, Bio-Rad).
2. Disposable plastic gravity columns (e.g., Poly-prep Chromatography Column, Bio-Rad).
3. S Cation Equilibration/Wash Buffer: 50 mM sodium phosphate buffer pH 7.4.
4. S Cation Elution Buffer: 50 mM sodium phosphate buffer pH 7.4, and 500 mM NaCl.
5. Ni affinity resin columns (e.g., HIS60 Ni Superflow Columns, Clontech).
6. Ni Equilibration/Wash Buffer: 50 mM sodium phosphate pH 7.4 and 300 mM sodium chloride (see Note 1).
7. Ni Elution Buffer: 50 mM sodium phosphate pH 7.4, 300 mM sodium chloride, and 300 mM imidazole (see Note 2).

2.5. Endotoxin Removal and Endotoxin Test

1. Toxin Eraser™ Endotoxin Removal Kit (GenScript).
2. Centrifugal filtration device (e.g., Amicon Ultracel-30K, Millipore).
3. Centrifugal filtration unit (e.g., 0.22 μm GV Durapore, Millipore).
4. Endotoxin Assay (e.g., ToxinSensor™ Chromogenic LAL Endotoxin Assay, GenScript).
5. 96-well flat bottom plate (e.g., 96-well flat bottom, Nunc Cat # 269620, Thermo Scientific).
6. Incubation oven; 37 ± 1.0°C.
7. Spectrometer with a 545 nm filter.

2.6. Matrix-Assisted Laser Desorption Ionization: Time of Flight Mass Spectrometry

1. Matrix-assisted laser desorption ionization: time of flight (MALDI-ToF) instrument such as Waters Micromass® MALDI microMX™ (MALDI-ToF MS) with MassLynx 4.1 software (Waters).
2. SDS-PAGE gel (Invitrogen).
3. SimplyBlue™ SafeStain (Invitrogen).
4. Trypsin Gold, mass spectrometry grade (Promega).
5. Trifluoroacetic acid (TFA), HPLC grade (Fisher Scientific).
6. Acetonitrile (ACN), HPLC grade (Fisher Scientific).
7. 200 mM NH_4HCO_3 buffer pH 7.8.
8. Ultrapure water.
9. $ZipTip_{C18}$® pipette tips (Millipore).
10. Savant SDP121P Speedvac Concentrator with refrigerated vapor trap RVT4104 (Thermo Scientific).
11. MALDI stainless steel target plate (Waters).
12. Matrix (see Note 3): 10 mg/mL α-Cyano-4-hydroxy-cinnamic acid (CHCA), 10 mg of CHCA is dissolved in 1 mL of solution containing 49.5% ethanol, 49.5% ACN, and 1% aqueous 0.1% TFA. It was prepared according to MALDI microMS standard operating procedures. This solution should be prepared fresh every time.
13. Adrenocorticotropic hormone 18–39 clip (2,465.1989 Da, Sigma-Aldridge) (ACTH): Mix 1 μL of 10 pmol/μL solution with 9 μL of 0.1% TFA to give 1,000 fmol/μL. Mix 1:1 with CHCA solution to give 500 fmol/μL.
14. Sodiated polyethylene glycol (PEG): Prepare 10 mg/mL solution in 50% aqueous ACN. Mix 1 μL PEG 1000, 2 μL PEG 2000, and 6 μL PEG 3000. Add 9 μL of CHCA matrix and 3 μL of 2 mg/mL of sodium iodide (prepared in 50% aqueous 2-propanol).
15. Glu-Fibrinopeptide B (1570.6774 Da, Sigma-Aldridge) (GFPB): Mix 1 μL of 10 pmol/μL GFPB stock solution with 9 μL of 0.1% TFA to give 2,000 fmol/μL. Mix 1:1 with CHCA solution to give 1,000 fmol/μL.

2.7. Protein Preparation for N-Terminal Sequencing

1. PVDF membrane (e.g., Sequi-Blot PVDF 0.2 μm for protein sequencing, Bio-Rad).
2. Methanol (HPLC grade).
3. Coomassie Blue R250 or G250 (Invitrogen).

2.8. Silver Staining of Proteins on PAGE Gels

1. The Pierce SilverSNAP Stain II (e.g., Thermo Scientific).
2. Fixing solution: 10% ethanol and 30% acetic acid.
3. 10% Ethanol.

4. Stop solution: 5% Acetic acid.
5. Gel imaging system (e.g., Versa Doc 4000 Imaging System, Bio-Rad).

2.9. ELISA

1. HIS-tagged protein standard (e.g., HSP90α^{HIS}, SignalChem Cat# H36-50 H).
2. 96-well flat bottom plates (e.g., Nunc 96-well flat bottom, Cat # 3855, Thermo Scientific).
3. 1× PBS: 137 mM NaCl, 2.7 mM KCl, 10.1 mM Na_2HPO_4, 1.76 mM KH_2PO_4, pH 7.4.
4. Blocking buffer: 1× PBS with 3% BSA.
5. Dilution buffer: 1× PBS with 1% BSA.
6. Washing buffer: 1× PBS with 0.05% Tween-20.
7. Detection antibody: Anti 6×HIS-HRP (Invitrogen).
8. Substrate Solution: Tetramethylbenzidine (TMB) Substrate Solution (KPL).
9. Stop solution: 1 M H_2SO_4 stock.
10. Disposable troughs or basins for dispensing solutions with a multichannel pipette.
11. Multichannel pippettes and/or robotic multichannel liquid handler system (e.g., epMotion 5075 Eppendorf).
12. Microplate washer (e.g., Model ELx405, BioTek).
13. Microplate reader spectrophotometer capable of reading at 450 nm (e.g., Model 680 Bio-Rad).

2.10. Deglycosylation

1. Peptide-N-glycosidase F (PNGase F) (Sigma-Aldrich).
2. 10× G7 Reaction buffer (supplied with enzyme).
3. 10× Glycoprotein Denaturing Buffer (supplied with enzyme).
4. NP-40 (supplied with enzyme).

3. Methods

3.1. Factoring Protein Recovery and Assessment into Expression Construct Design

Regardless of the production platform, in order to isolate and purify a protein, a polyhistidine-tag is often designed into the protein expression construct to enable affinity purification and to facilitate protein detection. We typically use six histidine residues (6×HIS-tag), which, in the case of our model IL-12 protein, is added at the C-terminus. A HIS-tag can be placed at either the C- or N-terminus of the protein of interest with selection of site often based on knowledge of active sites or other important structural domains. Because most of our “proteins of interest”

are glycoproteins requiring routing through the endomembrane system, we commonly engineer the HIS-tag at the C-terminus as opposed to the N-terminus, which comprises the signal peptide and is subsequently cleaved in the endoplasmic reticulum. If the HIS-tag is to be part of the target product, it will be important to determine that the tag does not negatively impact bioactivity.

In higher eukaryotic expression systems, another design consideration for expression of complex recombinant proteins is the signal peptide to use. This sequence can derive from the native gene (as is the case for the model IL-12 protein) or can be a classic signal peptide that is well characterized in the heterologous bioproduction system. We will often exchange the native signal peptide for a signal peptide derived from a plant protein [typically using the sequence corresponding to the potato patatin protein signal peptide (*Pat)*]. For some recombinant proteins, use of a plant signal peptide increases product yields in a plant-based production platform. In either case the recombinant-expressed protein is expected to emerge fully processed from the endomembrane system (as confirmed by N-terminal sequencing; see below). Although outside the scope of this chapter, it should be noted that both the trafficking route and the subcellular site of accumulation will potentially impact product yield, post-translational modifications, extractability, and quality/stability of a particular recombinant protein (18).

The starting material for the procedures described below is typically leaves of *Nicotiana benthamiana* containing the heterologous HIS-tagged protein generated via an *Agrobacterium tumefaciens*-mediated transient expression system as described by (2). Typically, leaves are harvested 48–72 h (optimized for each protein/construct) after vacuum infiltration of *A. tumefaciens* bearing the appropriate gene construct and processed immediately after harvest or stored at −80°C. For quality assessment, it is necessary to provide a parallel negative control using the same procedures employed with the recombinant protein preparation. Our negative control typically comprises *N. benthamiana* leaf tissue infiltrated with *A. tumefaciens* bearing an “empty” vector, i.e., the same vector used for expression of transgene but lacking the coding region for the protein of interest. This material is processed in parallel for all steps.

3.2. Validation of Recombinant Protein Expression Levels: Production Benchmark

It is important to have a reliable method for determining the productivity of your bioproduction system that is independent of extraction and purification variables. This is critical in process development and as a routine check point that the production run is meeting minimal productivity levels. We have instituted an initial analysis that involves extracting proteins from a small amount of tissue (typically 0.5 g of infiltrated *N. benthamiana* leaf tissue) by homogenization directly in SDS sample buffer and detecting the protein of interest by Western immunoblotting. This procedure

effectively extracts all proteins from the tissue providing an estimation of the total amount of target protein produced and establishing the baseline for product recovery during extraction and purification. It should be noted that this analysis does not measure "active" protein (SDS denatures the proteins) but does provide size estimates and indications of degradation. This procedure is very useful during initial screening or assessment of new proteins whose extraction behavior is unknown. During process development, the method provides a simple assessment tool that enables (1) optimization of the vector, gene construct, and expression system parameters, (2) determination of optimal leaf tissue harvest time(s), and (3) development of extraction protocols that effectively recover the recombinant protein. When this procedure is used to make relative comparisons between different expression constructs or harvest times, it is important to sample leaves of similar size and in comparable positions on the plant in an effort to mitigate any difference due to developmental stage of the leaf sample. Once production and downstream protocols are established, this test can also be integrated into process control for validation of your "raw materials" to ensure a given protein production run meets minimal yield criteria and to assess lot-to-lot variation.

3.2.1. Total Protein Extraction ("SDS-grind")

1. In the presence of liquid nitrogen, grind 0.5 g leaf tissue sample to a fine powder with a pre-chilled mortar/pestle. Do not let tissue thaw during grinding process. Use clean mortar and pestle for each sample in an effort to avoid cross-contamination.
2. Add two volumes (for 0.5 g; add 1 mL) of SDS grinding buffer to frozen tissue powder and thoroughly mix with a pestle to saturate sample with buffer. Allow frozen sample mixture to thaw before further grinding sample to a homogeneous liquid extract.
3. Transfer each extract using a clean transfer pipet to a 1.5 mL microcentrifuge tube.
4. Centrifuge all samples at 18,000 × *g* at room temperature (RT) for 10 min.
5. Carefully transfer supernatant to a clean 1.5 mL microcentrifuge tube. Samples may be immediately processed (see Subheading 3.2.2) or an aliquot(s) volume anticipated to run on SDS PAGE can be dispensed and stored at −20°C.

3.2.2. SDS Polyacrylamide Gel Electrophoresis

SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) is a technique used to separate proteins according to their size. Following gel electrophoresis, the total soluble protein in a sample can be visualized in gels by Coomassie or silver staining (see Subheading 3.5.1) or your protein of interest can be detected by Western immunoblot analysis based on protein-specific antibodies (see Subheading 3.2.3).

1. Set up NuPAGE precast gels (see Note 4).
2. Prepare 1× SDS-PAGE running buffer (MOPS or MES) by diluting 40 mL of the 20× running buffer with 760 mL of water in a measuring cylinder. Cover with Para-film and invert to mix (see Note 5).
3. Carefully remove comb and peel off tape from back of pre-cast gel, wash briefly with dH_2O and set up in an electrophoresis apparatus as per instructions.
4. Add 200 mL running buffer to the upper and lower chamber. Add 600 mL remaining buffer to the lower chamber (see Note 6).
5. Prepare samples derived from the "SDS-grind" procedure (Subheading 3.2.1) for electrophoresis under reducing conditions by adding DTT and heating. Depending on the expression level of your protein of interest, the sample volume load may vary. We have found that if your protein is moderately to well-expressed in plants, 5–10 μL of crude extract supernatant can easily be visualized by western analysis. For our model protein, ChIL-12^{HIS} a final gel loading volume of 10 μL is prepared as follows:
   (a) Add 6.5 μL of the extract supernatant in a 1.5 mL microcentrifuge tube (see Note 7).
   (b) Add 1 μL 1 M DTT stock to a final concentration of 100 mM DTT. Alternatively, one can utilize a commercial reducing agent for the NuPAGE gel system (10× NuPAGE® Sample Reducing Agent) (see Note 8).
   (c) Add 2.5 μL of 4× Bromophenol Blue.
   (d) Boil samples at 100°C for 5 min or heat at 70°C for 10 min.
   (e) Load samples onto an SDS-PAGE gel.
6. For samples comprising crude leaf extracts (i.e., not extracted in presence of SDS) or purification fractions (see Subheading 3.3), samples are processed as follows:
   (a) Add sample plus dH_2O to total of 6.5 μL in a 1.5 mL microcentrifuge tube.
   (b) Add 1 μL 1 M DTT stock to a final concentration of 100 mM DTT. Alternatively you can use 1 μL of 10× NuPAGE® Sample Reducing Agent (see Note 8).
   (c) Add 2.5 μL 4× NuPAGE LDS Sample Buffer to a final concentration of 1× loading buffer.
   (d) Boil samples at 100°C for 5 min or heat at 70°C for 10 min.
   (e) Load samples onto an SDS-PAGE gel.
7. Protein molecular weight markers, negative control (pBK), and positive control (if you have a commercial standard for the

protein of interest +/− HIS-tag) or a HIS-tagged protein standard to confirm subsequent antibody detection on western analysis should be included on this gel.

8. Load samples onto NuPAGE gel (see Note 9).
9. Run at 200 V constant voltage for about 60 min (see Note 10).

3.2.3. Western Immunoblot Analysis

Western immunoblotting involves the transfer of proteins that have been size-separated by SDS-PAGE to membranes and the detection of specific proteins based on their cross-reactivity to either protein-specific or epitope tag-specific antibodies. This analysis allows us to determine the approximate molecular size of our product and to measure relative amounts of the protein of interest. The protocol we use for our model protein ChIL-12^{HIS} permits detection of proteins in the range of 5–100 ng. It should be noted that many antibodies show non-specific binding to endogenous plant proteins in crude extracts at the antibody concentrations recommended in the manufacturer's protocols. Thus, both monoclonal and polyclonal antibodies will need to be optimized using the negative control plant extracts (processed in parallel) as the comparator to your transgenic samples. The protocol presented here involves detection based on the HIS-epitope engineered into our ChIL-12^{HIS} product and thus should be applicable to many HIS-tagged proteins. In general, the detection of proteins for which antibodies are not commercially available can be challenging, and therefore characterization and purification is greatly facilitated by the HIS-tag. The following outlines the assembly of a SDS-PAGE for electrotransfer to a membrane for performing western immunoblotting using the Invitrogen XCell II™ Blot Module. However, much of the process and technical tips are applicable to other gel transblotting units:

1. Resolve the protein samples by SDS-PAGE as indicated above (Subheading 3.2.2).
2. Pre-soak the blotting sponges, filter papers, and nitrocellulose membrane in a Pyrex glass container filled with 1× transfer buffer.
3. Carefully separate the pre-cast gel plates maintaining the polyacrylamide gel on one of the plates. To expose the gel, insert a metal spatula into the gap between the two gel plates, push up and down on the spatula handle to separate the plates (see Note 11). Mark the upper left corner of the gel to facilitate orientation throughout the process. For NuPAGE gels, the manufacturer recommends covering the gel with a pre-soaked filter to avoid gel drying out. However, we typically pre-equilibrate the gel briefly (2–4 min) in transfer buffer.
4. To assemble the membrane/gel sandwich, place two pre-soaked blotting sponges into the deeper of the two blot module cores

(designated the negative electrode). In a stepwise fashion, layer one pre-soaked filter paper followed by the pre-equilibrated gel, the nitrocellulose or PVDF membrane, and another pre-soaked filter paper to construct the gel/membrane sandwich. Remember to use clean gloves when handling gels and membranes. It is important to remove any air bubbles during this layering process that would void the transfer of protein to the membrane in these areas; a small glass rod can be gently rolled across the sandwich a couple of times in the same direction to remove any bubbles. Complete the assembly by layering two to three pre-soaked blotting pads to create a tight fit upon closing the blot module. Place the more shallow blot module core (designated the positive electrode) on top of the pads (see Note 12).

5. The sandwich should be held securely when transferred to the rails on the electrophoretic chamber. The sandwich assembly will only fit into the blot module chamber in one direction. Place the gel tension wedge in vertically against the blot module and push forward to lock into place.
6. Fill the inner chamber, which contains the sandwich assembly with transfer buffer (see Note 13).
7. Add 650 mL of dH_2O to the outer buffer chamber to reduce heating of the gel during the run (see Note 14).
8. Place the lid on top of the chamber and connect the color coded leads (red +; black –) into the corresponding positions on the power supply.
9. Run the gel electrotransfer at constant voltage; 35 V for 1–2 h (see Note 15).
10. Following transfer, dissemble and remove the membrane from the blot assembly. Place membrane on a Kimwipe with the side of the membrane that was in contact with the gel face-up and allow the membrane to air-dry for about 5 min.
11. Block the membrane with 20 mL blocking buffer and shake for 1 h.
12. Add anti-6×HIS(C-term)-alkaline phosphatase conjugated antibody in blocking/dilution buffer and incubate membrane for 1 h with rotation on a platform shaker at RT (see Note 16).
13. The membrane is then subject to three 10-min washes in washing buffer on platform shaker at RT.
14. Mix detection reagents: 3 mL of CDP-Star Detection Buffer, 150 μL of Nitroblock Enhancer II, and 30 μL CDP-Star. Place the membrane face-up inside of a sheet protector, cover the membrane with detection mix and incubate for 5 min at RT.

15. Lay the other face of the sheet protector over the membrane avoiding formation of air bubbles and using a Kimwipe to squeeze out the excess detection reagent before placing in a film cassette. In a darkroom, expose membrane to film. Exposure times can range from 30 s to 2 h and we typically capture multiple exposures. For a new protein, 5 min exposure is a good starting point and then you can increase or decrease exposure time based on these results. Fix and develop film using an automated film processor.

3.3. Protein Purification

In order to purify plant-derived proteins harboring a HIS-tag from complex plant material (e.g., leaf tissue), we recommend using a two-step column chromatography procedure. We generally use ion-exchange chromatography as a first column purification step, which accommodates the large volumes associated with initial plant extracts. This is followed by affinity chromatography using a nickel-based resin, which selectively binds the HIS-tag. For ChIL-12^{HIS} isolation and purification, we use the following steps: (1) plant tissue homogenization in liquid N_2 and buffer, (2) centrifugation to clarify, (3) cation exchange chromatography, (4) Ni affinity chromatography, (5) endotoxin removal, and, (6) final polishing steps (buffer exchange, concentration, and filter-sterilization) prior to structural and activity characterization (Fig. 2). "Empty vector" negative control samples are processed equivalently through all steps. Chromatography steps are typically carried out by fast pressure liquid chromatography (FPLC) when processing plant extracts at large scale. However, when first analyzing a new protein, small-scale purification (5 g per batch) is typically performed by gravity flow chromatography. The small scale provides a rapid source of transgene product for assessing post-translational modifications and purification strategies, as well as an effective system for optimizing construct elements to test if there is product expression.

The small-scale purification is described as follows.

3.3.1. Homogenization and Clarification

1. Grind 5 g of transient infiltrated leaf tissue in the presence of liquid nitrogen using a mortar and pestle. Add 10 mL of S cation Equilibration/Wash buffer and grind again.
2. After thawing plant extract, transfer to a 15 mL conical tube and centrifuge at 12,000 × *g* for 10 min. Transfer the supernatant (crude extract) to a new 15 mL tube (see Note 17).

3.3.2. Ion Exchange Chromatography

1. To set up cation exchange column, add 1 mL of slurry (e.g., UNOSphere S resin) to a disposable chromatography column. Equilibrate with 5–10 column volumes of S cation Equilibration/Wash Buffer. S cation resin has up to 60 mg/mL of binding capacity, which provides sufficient binding for small-scale plant extracts that typically have 30–50 mg TSP per 10 mL extract.

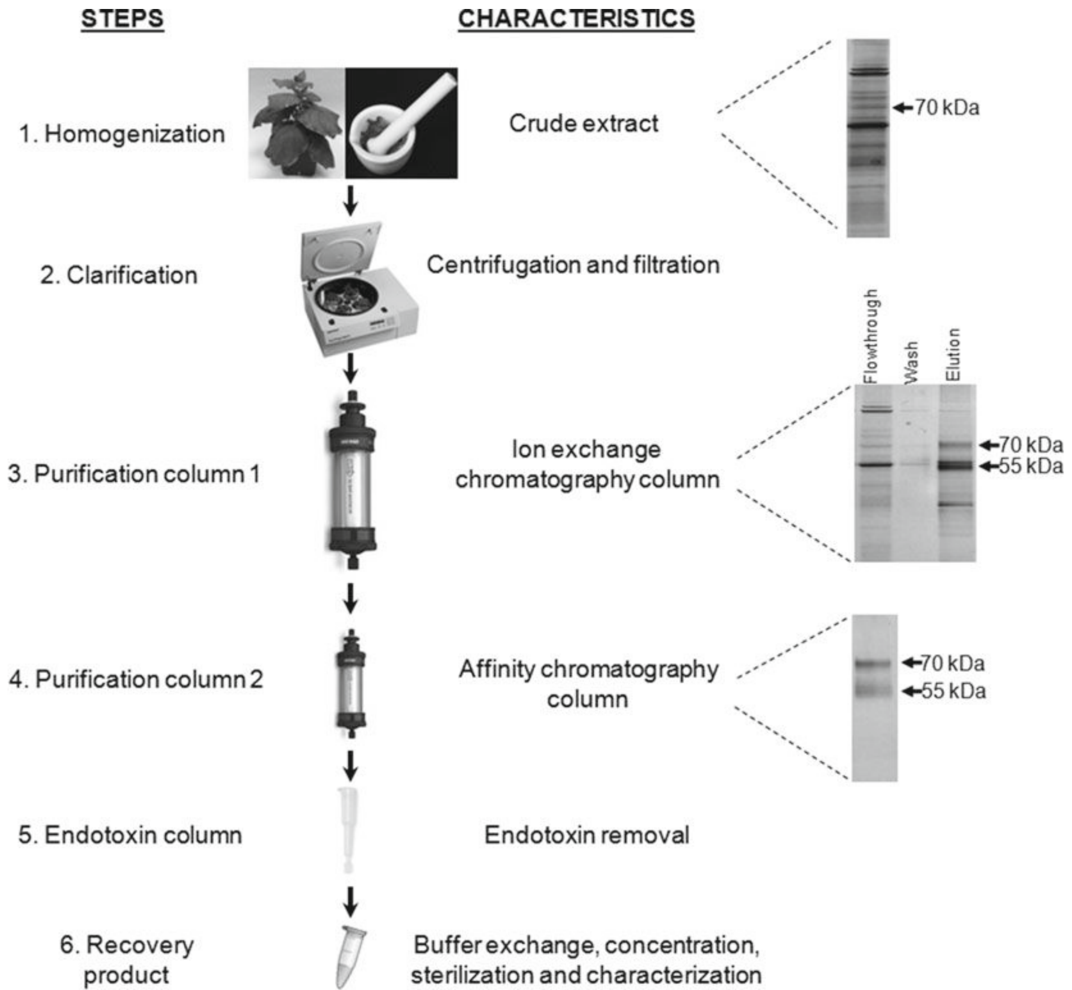


Fig. 2. Purification process for plant-derived ChIL-12HIS. Crude extracts were clarified by centrifugation and filtration. Samples were collected following elution from ion exchange chromatography column (S cation) and affinity chromatography column (Ni-NTA). The molecular size of purified ChIL-12HIS (70 and 55 kDa) is shown by gel analysis of silver-stained proteins on the right. Endotoxin removal column was included as a last column step followed by buffer exchange, concentration using Centricon 30 K, and filter sterilization prior to molecular characterization. Proteins from leaves infiltrated with pBIB-Kan empty vector were processed in parallel with ChIL-12HIS samples through all steps of purification and loaded onto SDS-PAGE based on equivalent volumes.

2. Add 10 mL of supernatant (crude extract) and collect flow-through fraction for downstream analysis.
3. Apply 5–10 column volumes of Equilibration/Wash Buffer to column.
4. Elute column with 1–5 column volumes of Elution Buffer. If proceeding to the next purification step, store eluted protein fraction containing your protein of interest at 4°C; otherwise store this elution fraction at –20°C until resuming purification the following day.

3.3.3. Nickel Affinity Chromatography for HIS-Tagged Proteins

1. Add 1 mL of His60 Ni Superflow slurry to a 15 mL tube and 5 mL (5–10 column volume) of HIS60 Ni Equilibration/ Wash Buffer. Mix by inversion.
2. Spin down or let the resin settle to separate the resin from buffer. Decant to discard the buffer.
3. Add the sample previously purified by S cation column to the 15 mL tube and allow target protein to bind by slowly inverting (rocking or rotating) the tube for 1 h preferably at 4°C. The His60 Ni Superflow resin has 60 mg/mL of binding capacity.
4. With a clean disposable column positioned in a stand, transfer the resin allowing it to slowly settle.
5. Wash the column with 5–10 column volumes of HIS60 Ni Equilibration/Wash Buffer.
6. Elute the target protein with 10 column volumes of Elution Buffer; collect as 1 mL fractions. Store samples at −20°C for further analysis.

3.3.4. Larger Scale Purification

Purification protocols at a scale of 35 g fresh weight/batch have been described previously (3). Product recovery at this scale provides sufficient quantities of purified ChIL-12^{HIS} for confirming protein identity by N-terminal sequencing and MALDI-ToF analysis. For larger scale purification runs (e.g., starting with 200 g leaf fresh weight/batch), additional concentration or ultrafiltration steps are often integrated to deal with increased volumes. At this scale, we have integrated a polyethylene glycol (PEG) precipitation of our crude protein extract in order to reduce the volume (and provide a potential enrichment step) prior to the initial ion exchange chromatography step. Ammonium sulfate precipitation can also be used as a scalable strategy for volume reduction as the amount of tissue is increased. Optimal amounts of PEG or NH_4SO_2 for effective recovery of the protein of interest need to be determined for each protein. As one considers scaling to very high volumes, microfiltration and other strategies that do not require centrifugation may be warranted.

3.3.5. Endotoxin Removal and Endotoxin Detection Test

Endotoxin test is one of the most critical quality control tests required by the FDA for all drugs. Endotoxins are lipopolysaccharides (LPS) associated with presence of gram-negative bacteria and can cause severe reactions in humans and animals. Plant biomass resulting from greenhouse or field-grown stable transgenic lines or, in our case, plants infiltrated with *A. tumefaciens* has a significant bacterial "bioload". Endotoxins must be removed from proteins used for in vivo applications to avoid the potential of unwanted inflammatory responses. This is particularly important when your protein of interest is an immune activating protein, such as is the case with ChIL-12^{HIS}. Quantification of endotoxin levels and

documentation to ensure the endotoxin load in the protein product is below a certain level is often required as part of quality assurances for the product to be sold as a research reagent or used in animal trials or clinical applications. The Limulus Amoebocyte Lysate (LAL) Endotoxin Assay described below meets industry standards for endotoxin assessment.

1. Load purified samples on an endotoxin removal column (GenScript) in accordance to manufacturer's instructions (see Note 18).
2. Concentrate the eluted fraction(s) using a centrifugal filter device (e.g., Amicon Ultracel-30K) at 5,000 × *g* for about 10–30 min to facilitate further processing.
3. For buffer exchange, add 3 mL PBS and centrifuge again as outlined in the previous step.
4. Repeat concentration of protein product by centrifugation for 10–30 min at 5,000 × *g* using centrifugal filter 30 K (see Note 19).
5. Filter-sterilize by centrifugation for 1 min at 12,000 × *g* using a 0.22 μm centrifugal filter unit.
6. Quantify endotoxin levels using ToxinSensor™ Chromogenic LAL Endotoxin Assay according to manufacturer's instructions (see Note 20).

3.4. Identity Confirmation

While cross-reactivity with antibodies, or while a confirmed specific binding or enzyme activity strongly supports that the protein you produced and purified is in fact the recombinant protein of interest, direct proof of protein identity is required prior to moving forward with commercial applications. Mass spectrometry (typically of tryptic peptides derived from your purified protein) or N-terminal sequencing will provide direct peptide sequence information that can be used to establish protein identity. The amounts of purified protein required for these analyses range from 1 to 100 pmol for MALDI-ToF and 10 to 100 pmol for N-terminal sequencing. According to the ICH Q6B Guideline (http://www.ich.org/LOB/media/MEDIA432.pdf), N-terminal sequencing may be required for characterization of recombinant proteins for clinical testing. Some proteins are "blocked" for N-terminal sequencing due to chemical modification of the protein (e.g., acetylation) that occur in vivo or during purification. Although "unblocking" by enzymatic treatment is an option in some cases, often mass spectrometry approaches need to be substituted in order to establish identity.

For purified plant-derived ChIL-12^{HIS}, Western immunoblot analyses (see below) typically revealed two major products at approximately 70 kDa (p70; expected size of single-chain IL-12) and 55 kDa (p55). MALDI-ToF analyses confirmed that both of these bands were products of the *ChIL-12^{HIS}* gene. N-terminal

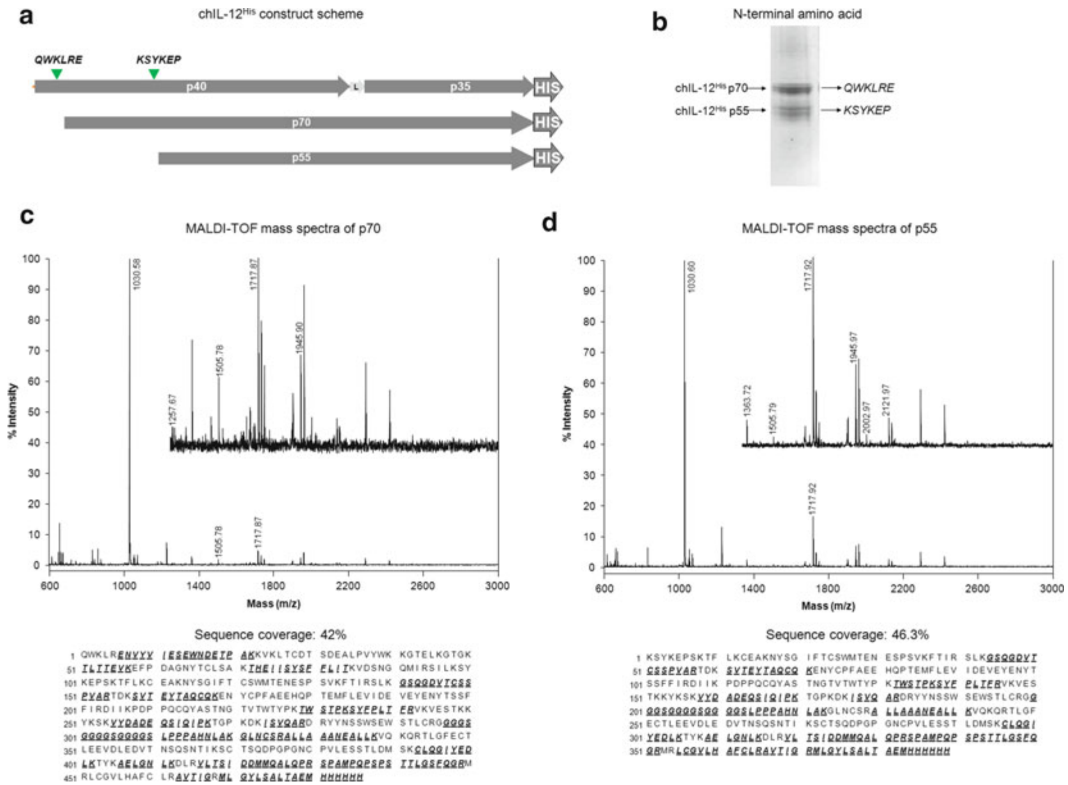


Fig. 3. Analytical techniques for establishing protein identity. (**a**) Schematic representation of the ChIL-12HIS construct including the chicken interleukin-12 p4*0 subunit (p*40) including its own signal peptide, a $(G_4S)_3$ gly-ser linker (L), the mature p35 subunit (p35), and a 6× histidine epitope tag (HIS) at the C-terminus. (**b**) Coomassie stained PVDF membrane of purified ChIL-12HIS submitted for N-terminal sequencing and the resulting amino acid sequence obtained for the p70 and p55 products. The N-terminal sequence of the 70-kDa protein, *QWKLRE*, is consistent with precise cleavage of the p40 signal peptide at a site between amino acids 20 and 21. The N-terminal sequence of the 55-kDa product, *KSYKEP*, indicates a loss of 97 amino acids from the N-terminus of the mature peptide. (**c**, **d**) MALDI-ToF spectra of the ChIl-12HIS peptides from the p70 (**c**) and p55 (**d**) products after trypsin digestion. Sequence coverage for p70 was 42% whereas for p55 was 46.3%. Inset spectra are amplifications of higher mass regions. Peptide masses are given if they comply with the in silico prediction. *Bolded underlined sequences* represent peptide identified by mass analyses.

sequence analysis provided additional identity confirmation as well as establishing the precise processing site for signal peptide cleavage of the p70 product. N-terminal analysis of the p55 product confirmed that this protein was a degradation product of the p70 ChIL-12HIS and established the precise site of cleavage.

3.4.1. Matrix-Assisted Laser Desorption Ionization: Time of Flight Mass Spectrometry

MALDI-ToF can be used to confirm identity of a recombinant protein (see Chapter 4 for greater details on MALDI-MS for protein identity). For our ChIL-12HIS example, MALDI-ToF proved effective in analyzing both the p55 and p70 products recovered from the ChIL-12HIS purification procedure. A peptide mass fingerprint (Fig. 3) was obtained after the two ChIL-12HIS products were treated with trypsin, a protease that cleaves polypeptide sequences

at arginine and lysine residues. MALDI-ToF (in positive reflectron mode) is most effective at visualizing fragments in the range of 0.8–3.5 kDa. However in practice, we set the range between 0.7 and 3.0, thus eliminating the over-saturation of the detector by the smaller fragments. It should be noted that glycosylated peptides may be underrepresented (i.e., may not "fly") in MALDI-ToF analyses and that analyses may provide less than 60% coverage of the complete sequence due to size, modifications, or low ionization potential of specific fragments. However, sequence confirmation of two to four peptides provides high confidence of product identity.

1. Size-separate purified protein, e.g., ChIL-12^{HIS}, by SDS-PAGE under reducing and denaturing conditions. Following electrophoresis, stain the gel with SimplyBlue™ SafeStain according to manufacturer's procedure.
2. Cut out protein bands, visualized with SimplyBlue™ SafeStain, from the gel (in our case, the p70 and p55 products) and slice into approximately 1 mm pieces (see Note 21). Perform in-gel trypsin digestion using Trypsin Gold followed by sample cleanup and peptide enrichment (Zip-Tip treatment; see Note 22) according to manufacturer's procedures. Elute tryptic peptides from the Zip-Tip with 10 μL of CHCA matrix.
3. Spot a portion (1 μL) of the resulting tryptic peptide mixture on the MALDI target plate and allow to dry at RT (can also dry under vacuum). Spot 1 μL each of ACTH, PEG, and GFPB, prepared as described in Subheading 2.5, onto the plate as well.
4. Perform MALDI-ToF MS measurements using an instrument such as the Waters Micromass® MALDI microMX™ equipped with a 337 nm Nitrogen UV laser. Use the positive reflectron mode (5,200 V) to acquire the spectra and collect data within the mass range of 700–3,000 m/z. The acquisition voltage we used was 12,000 V (see Note 23).
5. Optimize peak resolution using ACTH (see Note 24) and calibrate using PEG (provides prescribed sizes equivalent to a "ladder" in electrophoresis). Use GFPB as the "lock mass".
6. Use the peptide mass analysis software at http://www.expasy.org/tools/peptide-mass.html to facilitate mass prediction of tryptic peptides, selecting "trypsin" as the proteolytic enzyme and "no missed cleavage allowed". Use the default option for all the remaining selections. Set the software to display peptide masses between 600 and 3,000 Da.
7. Confirm the identity of the protein by comparing the masses of peaks obtained with MALDI-ToF MS and those computationally predicted for tryptic digests. This comparison is performed automatically by using MASCOT at the EXPASY Web site (see above). Alternatively, this software can be accessed directly at http://www.matrixscience.com/search_form_select.html (See also Chapter 4 of this book).

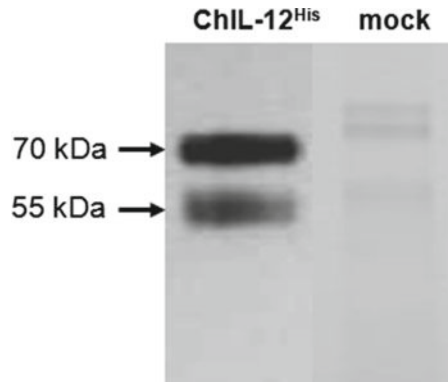


Fig. 4. Western immunoblot analysis. Purified ChIL-12^{HIS} (lane 1) or pBIB-Kan "mock" control (lane 2) were compared by Western analyses. Proteins were separated by SDS-PAGE run under reducing conditions. Immunoblotting was performed using a mouse monoclonal anti-6×HIS alkaline phosphatase-conjugated antibody as described in Subheading 3.2.3. The positions of the p70 and p55 bands are indicated *on the left.*

3.4.2. N-Terminal Sequencing

N-terminal sequencing via Edman degradation is valuable as an analysis tool for the sequencing and determination of the first 5–15 amino acids (from the N-terminus) of a protein or peptide. Success is dependent on amount, purity, and amino acid composition of the sample. The N-terminus of some proteins is "blocked" by specific chemical modifications such as acetylation. Typically purified protein is submitted to a protein sequencing facility and technical experts associated with these service centers will recognize a blocked protein and recommend changes in your procedures to minimize N-terminal modifications or enzymatic strategies to "unblock" the N-terminus. N-terminal sequence analysis of purified ChIL-12^{HIS} was reported previously (3) and is delineated as an example here (see Fig. 4):

1. Run a SDS-PAGE gel with the standard protocol. However, leave an empty lane between each of the samples including the stained molecular size ladder. This will facilitate cutting the bands out and minimize opportunities for cross-contamination. It is best to run more than one lane to ensure sufficient protein sample recovery for this analysis. For ChIL-12^{HIS}, at least 1 μg was loaded per lane. If you can visualize the bands on the PVDF membrane with Coomassie stain (see below), there should be sufficient material for sequencing. Transfer to PVDF by electrotransference (see Note 25).
2. Use clean gloves when handling the membrane as well clean containers that have not been used for other protein-related work for procedure outlined below to avoid cross-contamination.
3. PVDF membrane must be pre-wet in 100% methanol until the entire membrane is translucent before equilibrating in the transfer buffer (see Note 26).

4. Set up membrane transfer as described above (Subheading 3.2.3). Transfer for 2 h at 200 mA.
5. Following transfer, remove the membrane from the cassette and rinse 3 × 5 min in dH_2O.
6. Stain for 10–30 min in 40% methanol with 0.025% Coomassie Blue R250 or G250.
7. Destain for 15 min with 50% methanol.
8. Repeat step 5 as many times as needed, until the background is light blue.
9. Air-dry the membrane for a minimum of 2 h or overnight.
10. Capture an image of the membrane using a gel imaging system or scanner. A stained membrane picture must be submitted to the sequencing facility with your sample with the band selected for sequencing clearly marked.
11. Cut the band of interest on PVDF membrane with a clean new razor blade and place it into a micro-centrifuge tube. If you run more than one band for the same sample, place them in the same tube. Alternatively, the entire PVDF membrane can be submitted to the sequencing lab with instructions as to the bands to be used for sequencing.
12. Fill out the appropriate sequencing forms and provide samples to an appropriate sequencing facility. For ChIL-12^{HIS}, we have used the Protein Facility in the Office of Biotechnology at Iowa State University.

3.5. Protocols for Assessment of Protein Yield, Quantity, and Quality

The tools and procedures used to determine the quality and quantity of the recombinant protein of interest will vary depending on the stage of R&D and the particular product to be produced. If industry standards are available, much of your assessment can target "equivalency" i.e., demonstrating that your product meets the specifications of the current industry standard. However, for many proteins, these standards do not exist and must be developed and validated as part of the process of product development. In early stages of R&D, monitoring a protein based on immunological tools (e.g., Western immunoblots, ELISAs) and/or activity (e.g., enzyme assays, binding assays, etc.) provides initial assessment of (1) product yields and recovery through purification steps, (2) product heterogeneity and the presence (or generation) of degradation products or aggregation, and (3) activity and function. As your protein product moves toward commercialization (e.g., research reagent) or preclinical assessment in support of regulatory approval, the goals shifts to developing and validating the key assessment tools that will constitute or support production parameters and product specifications. These analyses will include product characterization (physicochemical properties, bioactivity, purity, impurities, quantity, etc.), development of reference standards

and reference materials, and establishment of in-process acceptance criteria and limits. In contrast to the early R&D process, which focuses on information gathering on product characterization and process optimization, the later stages target the establishment of standard operating procedures for production, purification and product specifications, quality controls, and quality assurance parameters. The protocols below provide additional methods that aid in the process of characterizing and quantifying a recombinant protein as you move through these stages.

3.5.1. Silver Staining Proteins in PAGE Gels

Silver stain is a highly sensitive method for detecting proteins in polyacrylamide gels. We typically use the silver stain technique to assess purity of our protein and to monitor enrichment during purification.

1. Separate proteins by SDS-PAGE as described above (Subheading 3.2.2).
2. After the electrophoresis run, separate gel (Subheading 3.2.3).
3. Carefully transfer gel to a container compatible with the staining process.
4. Follow the SilverSNAP Stain II kit instructions (see Note 27).

3.5.2. Western Blot Analysis for Product Characterization (Fig. 5)

In addition to providing an estimate of product amounts and apparent molecular size (Subheading 3.2.3), Western immunoblots provide a key assessment tool to "follow" the recombinant protein during purification and to detect the presence of degradation products or protein oligomers. In conjunction with glycan degrading enzymes (Subheading 3.7), Westerns blots can be used to assess presence and number of N-linked glycans [see ref. 3 for application of this technique to ChIL-12^{HIS}]. The protocol for Western analyses of ChIL-12^{HIS} utilizing antibodies directed against the HIS tag is described previously (Subheading 3.2.3).

3.5.3. Direct ELISA

An ELISA permits high throughput results in determining if a particular protein is present in a sample and to quantify product levels. Of the four standard ELISA formats (see Chapter 4 for detailed discussion), we have developed a direct ELISA as well as indirect and direct sandwich ELISA in establishing our quality assessment program for the recombinant ChIL-12^{HIS} product. In the example highlighted below, we describe methods for the "direct" anti-HIS ELISA used in determining how much of our particular HIS-tagged ChIL-12 protein is present in plant-derived purified product. We have selected this format as this will likely be useful for detection/quantification of any HIS-tagged protein and is applicable to situations where specific antibodies to one's protein of interest are not available. For a direct ELISA for HIS-tagged proteins, the bottom of each well is coated with the HIS-tagged protein you want to measure – typically using multiple wells as

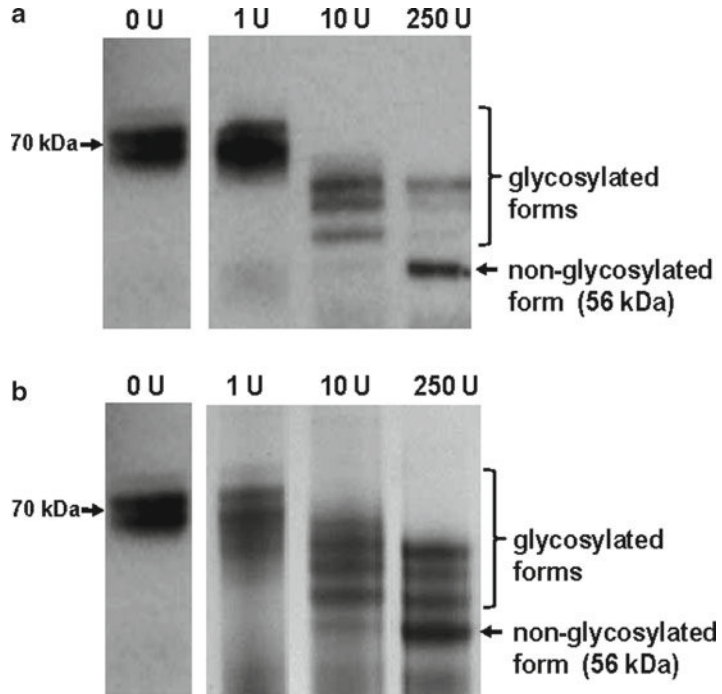


Fig. 5. Deglycosylation analysis of plant-derived ChIL-12^{His}. ChIL-12^{His} purified from *N. benthamiana* leaves was treated with different concentrations (0, 1, 10, and 250 units) of peptide-N-glycosidase F (PNGase F) and subjected to reducing SDS-PAGE. Glycosylated and non-glycosylated forms were visualized by (**a**) western blotting using monoclonal anti-6×His-alkaline phosphatase-conjugated antibody and (**b**) by silver stain. *Arrow* on the *left* indicates the molecular weight for ChIL-12^{His} not treated with PNGase F.

replicates and analyzing serial dilutions of your target protein. In addition, a positive control commercial HIS-tagged protein (e.g., HSP90α^{HIS}) and a negative control (e.g., an analogous extracts from plants expressing an "empty vector") would be included among the 96 samples being tested. Samples are then allowed to cross-react with anti-HIS antibody that has been conjugated to horse radish peroxidase (anti-HIS-HRP). In the absence of a fully characterized ChIL-12^{HIS} standard, this anti-HIS direct ELISA provides the ability to establish a concentration value for our ChIL-12^{HIS} relative to a known protein (HSP90α^{HIS}) whose concentration has been standardized for commercial sale. It should be noted that this direct anti-HIS ELISA is only quantitative for purified protein and not in detecting ChIL-12^{HIS} in crude plant extracts. Once optimized and validated, ELISA serves as an effective quantification tool and can be largely automated through use of liquid handling systems.

1. Coat an ELISA plate with 100 μL/well serial dilutions of the purified protein (ChIL-12^{HIS}) or HIS-tag protein standard (e.g., HSP90α^{HIS} starting with 100 ng/mL) diluted in PBS (see Notes 28 and 29).

2. Seal the plate and incubate for 2 h at RT or overnight at 4°C.
3. Wash the plate three times with washing buffer and block the wells with 300 μL/well of blocking buffer for 1 h at RT. Seal the plate.
4. Wash the plate and add the anti-6×HIS-HRP conjugated detection antibody (dilution 1:2,000) to a final volume of 100 μL/well. Seal the plate and incubate at RT for 1 h.
5. Wash the plate and add 100 μL/well of TMB substrate solution to each well. Incubate plate at RT for 30–45 min (see Note 30).
6. Add 50 μL of stop solution to each well and determine the absorbance of each sample well using a plate spectrophotometer at 450 nm (see Note 31).

3.6. Biological Activity of a Recombinant Protein

The biological function of a protein can be enzymatic (e.g., β-glucuronidase), structural (e.g., collagen), or serve a carrier (hemoglobin), regulatory (receptor proteins and insulin), binding, or mechanical (myosin) function. The biological activity of a recombinant protein is routinely measured using a bioassay, such as cell proliferation assay, enzyme assay, or a functional ELISA. For enzymes, activity characterization and bioequivalency assessments typically included biochemical parameters such as V_{max}, K_m for key substrates and inhibitors, pH optima, temperature optima, and co-factor requirements. Cytokines such as IL-12 do not catalyze simple reactions and must be assessed based on their ability to activate complex reactions in specific responsive cells. To demonstrate bioactivity of our purified plant-derived ChIl-12^{HIS} described in this chapter, the concentration-dependent production and secretion of chicken interferon-gamma (ChIFN-γ) by chicken spleen cells 48 h after treatment with ChIL-12^{HIS}, positive control (ConA) or negative control (mock) was measured by using a direct and specific ChIFN-γ capture ELISA assay (3, 19).

3.7. Deglycosylation

Many proteins of interest expressed in higher eukaryotes contain N-linked or O-linked glycans that are added co- or post-translationally as proteins traffic through the endomembrane system. Depending on the application of one's particular recombinant protein, it may be important to determine whether the recombinant protein is glycosylated and to characterize the glycan moieties. Because glycans can vary in the number of sugars present at a particular site and whether a consensus site (e.g., the N-X-S/T motif for N-linked glycosylation) is occupied, glycoproteins often exhibit size heterogeneity (commonly seen as "fuzzy" bands on SDS-PAGE). We typically use enzymatic deglycosylation as an initial tool to confirm the presence of N-linked glycans and to assess the number of sites that are glycosylated. In this procedure (described below), the purified protein is incubated with specific endoglycosidases at multiple concentrations (or for various periods of time) and the protein sizes with and without glycosidase are compared. An

N-linked glycan (6–11 sugars) can contribute about 2 kDa of mass to the protein and a deglycosylation time-course can reveal a "ladder" of products indicting the number of sites that are occupied (e.g., Fig. 5). Some protein structures and some sugars provide more recalcitrant substrates for specific glycosidases. For example, complex plant glycans containing xylose and/or fucose are often not efficiently cleaved by PNGase-F and may be more effectively cleaved by PNGase-A (endo A) (20).

An endoglycosidase-mediated change in molecular size, as described below and in Fig. 5, establishes that the recombinant protein is glycosylated. For some applications (e.g., as research reagent), an estimate of the average proportion of the product that is protein and carbohydrate may be included in the product specification. This may be important for enzymes whose activity may be affected by the presence or size/composition of glycans that are added in a particular expression system (21). The composition of N-linked glycans can differ among proteins expressed in plant, mammalian, fungal, and insect cells. For certain human proteins, the presence or composition of the glycans can have significant impacts on bioactivity and/or pharmacokinetics. Thus, for most plant-made pharmaceutical glycoproteins, preclinical studies establishing the specific composition of the glycans will be required. This can involve determination of the total sugar *composition* (glycan are acid hydrolyzed and the amount and proportion of each sugar is determined) or determination of the glycan *structure* (typically involving sequential exoglucosidase digestions). Unless these analyses are routine within one's laboratory or institution, we recommend using a well-established center (e.g., the Complex Carbohydrate Center at the University of Georgia, Athens, GA, USA) or commercial entity. If the presence or absence of specific glycans impacts in vivo efficacy, batch-to-batch variation will need to be established and documentation of the glycan content will likely be instituted as part of the QA/QC parameters for lot release. The following protocol describes the digestion of ChIL-12^{HIS} with peptide-N-glycosidase F (PNGase F) and was used to confirm that the plant-derived product was glycosylated and determine how many sites were modified.

1. Denaturation step: Prepare the following denaturation reaction, mix 200 ng of purified ChIL-12^{HIS} with 1× Glycoprotein Denaturing Buffer (included in the kit) in a final volume of 5 μL.
2. Denature the sample by boiling for 10 min.
3. PNGase-F digestion step: The PNGase-F stock we used had a concentration of 500 U/μL. Prepare enzymatic deglycosylation reactions with 0, 0.1, 1, 10, 100, and/or 250 U of PNGase-F as shown in Table 1 (see Note 32). Keep samples in

Table 1
Reaction components for enzymatic deglycosylation

	0 Units	0.1 U	1 U	10 U	100 U	250 U
Denaturation reaction	5 µL	5 µL	5 µL	5 µL	5 µL	5 µL
10× G7 buffer	1 µL	1 µL	1 µL	1 µL	1 µL	1 µL
10% NP40	1 µL	1 µL	1 µL	1 µL	1 µL	1 µL
PNGase	0 µL	0.2 µL (from dil 1/1,000)	2 µL (from dil 1/1,000)	2 µL (from dil 1/100)	2 µL (from dil1/10)	0.5 µL (from stock)
dH_2O	3 µL	2.8 µL	1 µL	1 µL	1 µL	2.5 µL
Final volume	10 µL	10 µL	10 µL	10 µL	10 µL	10 µL

ice while preparing and add all components except enzyme to all samples.

4. To initiate reaction, add enzyme and incubate at 37°C for 1 min. Incubate at 100°C for 5 min to stop the reaction.
5. Digests are resolved by reducing SDS-PAGE and analyzed by either Western analysis using the anti-6× His antibody or silver staining described in Subheadings 3.2.2, 3.2.3, and 3.5.1, respectively.

4. Notes

1. Although the manufacturer recommends 20 mM imidazole in the "Ni Equilibration buffer" and 40 mM imidazole concentration in the "Ni Wash buffer", we have used a single "Ni Equilibration/Wash buffer" with no imidazole to avoid potential loss of precious protein.
2. Use a highly pure low-absorbance imidazole that is designated for liquid chromatography applications (e.g., Fisher, Product No. BP 305-50).
3. Matrix selection is based on strong light absorption at the wavelength of the laser flux. CHCA matrix is used for peptides < 10 kDa. This matrix is light sensitive, so a brown tube or aluminum foil should be used to protect the matrix.
4. NuPAGE Pre-Cast Gel System is a polyacrylamide gel system for high performance gel electrophoresis using specially optimized buffers and a low operating pH during electrophoresis, which results in sharper band resolution and longer shelf-life. An aliquot of 4× NuPAGE sample buffer can be stored at

RT, as this is a very viscous reagent and needs to be brought to RT before use.

5. Use MOPS SDS running buffer for better separation of proteins ranging from >200 kDa to 14 kDa. MES-SDS running buffer is better suited for resolving proteins in the range of 60–2.5 kDa. Be sure to check for leaks after adding running buffer to the chamber.
6. Before running your samples, add 0.5 mL antioxidant (NuPAGE® Antioxidant; Invitrogen) to the upper chamber if running SDS-PAGE under reducing conditions. This helps keep the samples reduced as the gel runs (the DTT in the sample does not migrate in neutral pH gels as it does in Tris-glycine gels).
7. In some cases the higher amounts of SDS can impact the band resolution or protein separation behavior on the gel (narrowing/band smearing). If your protein of interest is of sufficient concentration for detection, we recommend using 3.3 μL sample plus 2.5 μL of water to create the 6.5 μL of sample. This reduces the SDS to "2×" before gel loading and should eliminate these gel artifacts.
8. In crude plant extracts, attaining a fully reduced total soluble protein fraction is not always accomplished due to the complexity of the mix. While we routinely use the 10× NuPAGE® Sample Reducing Agent for reduction, be aware of the shelf life as the reducing agents can break down over time. If you begin to notice that samples do not appear fully reduced, use freshly made DTT (recipe in Subheading 2).
9. The 15-well comb gel can hold a total of 15 μL; the 12-well comb can hold a total of 20 μL; and the 10-well comb can hold a total of 25 μL.
10. We run gels at a constant voltage of 200 V, for approximately 60 min to obtain best separation of the bands for detection of ChIL-12^{HIS}. Running time may need to be adjusted depending on the size of your protein of interest.
11. When opening the cassette, the gel will likely adhere to one of the cassette plates. Carefully remove and discard the plate without the gel allowing the gel to remain on the other plate for further manipulations. Remove gel wells with a knife to establish the upper boundary of the gel that will define the top of the membrane.
12. For sandwich setup, it is important to remove bubbles using a roller or glass pipette rolling over the surface. Although the manual indicates two gels can be processed at the same time, we have noticed repeatedly that the second gel does not transfer as well as the first. We recommend running one gel at the time.
13. To avoid overheating of the unit during transfer, do not overfill the sandwich blot chamber with transfer buffer.

14. Use dH_2O instead of transfer buffer in the outer buffer chamber to avoid salts precipitation in the chamber.
15. In the blotting step, the SNAP i.d. protein detection system (Millipore) can be used as an alternative to conventional diffusion-based blotting. In general, following short incubations, a vacuum is used to actively drive reagents (antibodies, wash solutions, blocking agents) through the membrane. This novel method allows you to optimize your blotting conditions in 25–30 min for maximum results. However, we have observed that while SNAP i.d. is very useful for purified proteins, increased background is observed with plant crude extract and when re-using the cassette several times. In addition, the utility of this vacuum format western immunoblotting system is protein dependent as we have observed poor detection of some larger complex glycoproteins. If detection of the protein is lower than expected, one can check whether the protein was effectively transferred to the membrane by post-staining the gel after the transfer step to determine if any proteins remain in the gel. Alternatively, if you used a stained molecular weight standard on your gel, you can monitor these to determine if there was a good transfer to the membrane at all molecular sizes.
16. The mouse anti-6×HIS(C-term)-alkaline phosphatase conjugated antibody will only detect proteins where the HIS-tag has been added to the C-terminus and will not detect the HIS-tag added in N-terminus or internal positions of a protein; an alternate antibody is needed to detect the latter.
17. More time may be required for the sample to thaw following this process. Allow the sample to completely thaw prior to transferring the sample to a 1.5 mL tube. The thawing time varies depending on the buffer used.
18. In the elution step of endotoxin removal column, it is best to add more buffer to elute (approx 4.0 mL instead 1.5 mL), collect 1 mL fractions, and test each fraction to determine the fraction(s) that contain your target protein. In our experience we have found our ChIL-12^{HIS} protein present in the entire 4 mL effluent volume. If the final endotoxin level is above the desired endotoxin level, repeat the endotoxin removal procedure by reloading the sample onto a regenerated column. The column can be reused up to five times for the same protein. Do not reuse the column for samples containing different recombinant proteins. For endotoxin removal of larger sample sizes, consider the less costly resins from BioRad (FPLC-compatible Affi-prep Polymixin Cat#156-0010) or Sigma (Polymixin B-agarose, Cat#P1411).
19. When concentrating protein using a centrifugal filter device system, some protein becomes trapped in the membrane, which must be considered in final recovery estimates.

20. Reconstitute Limulus Amebocyte Lysate (LAL) and Chromogenic Substrate immediately before adding to your samples. The absorbance of the ToxinSensor™ Chromogenic LAL Endotoxin Assay is read at 545 nm in a spectrophotometer using either cuvettes or in a 96-well microplate format. In order to read samples in the latter format, at the last step of this assay transfer 300 μL of each reaction to a well in a 96-well plate.
21. Slices are used in order to provide better access of the trypsin to the protein in the gel. Additional details on in-gel trypsin digestion is provided in Chapter 4.
22. Manufacturer's guidance suggests using ZipTips for sample cleanup. We have found that this process can result in sample loss for some proteins so we often skip this step without an apparent loss in peptide detection. Also, for sample preparation, be aware of possible contaminants present (salts, buffer, glycerol). Maximum tolerated buffer concentrations are 20, 30, and 50 mM for phosphate, ammonium bicarbonate, and Tris–HCl buffers, respectively.
23. Ions produced by the laser could be $[M+H]^{+}$, $[M+2\ H]^{2+}$ (double charge), $[2\ M+H]^{+}$ (dimer). Peaks are isotopically resolved and appear as multiple peaks separated by a 1-Da mass difference. Use the mass of the first monoisotope for analysis.
24. Resolution of the instrument higher than 10,000 is highly recommended. This resolution is checked by dividing the MW of ACTH by its peak width at half the peak height. Having a higher peak resolution will result in better mass accuracy.
25. For sequencing analysis, samples need to be purified and run using PDVF membranes, and not nitrocellulose. Avoiding the introduction of contaminants greater than 10% is critical as this may result in an ambiguous sequence. For N-terminal sequencing it is important to use pre-cast gels to help eliminate any N-terminal blockage due to unpolymerized acrylamide in the gel. Samples can also be sent out for sequencing in liquid form if they are free of salts, buffers, amines, etc., and contain a single protein.
26. It is important to use high-grade methanol for this procedure.
27. For silver stain, the fixing step can be completed as quickly as 20 min to overnight with no adverse effects.
28. Purified protein samples are accurately quantified using this ELISA format. Conversely we have found that crude plant protein extracts produce high background and thus, this ELISA is not suitable for ChIL-12^{HIS} detection in crude extracts.
29. The coating step of an ELISA can be done as short as 2 h at RT; however, it is important to empirically establish this for each target protein to be tested.
30. Bring TMB Substrate Solution to room temperature 15–30 min before use.

31. It is better to read the ELISA plate at 450 nm within 30 min after addition of the stop solution.
32. To deglycosylate a native glycoprotein, longer incubation time as well as more enzyme may be required.

Acknowledgments

This work was supported by USDA SBIR Grant Phase I (2007-33610-17978) and Phase II (2008-33610-19482) to BioStrategies LC and a subaward to Arkansas Biosciences Institute (ABI) at Arkansas State University. Thanks to Nora Rubio, Jorge Ayala, and Nathan Stephens for technical support and help in standardization of direct ELISA, protein preparation for N-terminal sequencing, and endotoxin removal protocols, as part of BioStrategies LC projects. Thanks go as well to Drs. Brett Savary and Prasanna Vasu (ABI) for help in developing the MALDI-ToF procedures.

References

1. Liu C, Towler M, Medrano G, et al (2009) Production of mouse interleukin-12 is greater in tobacco hairy roots grown in a mist reactor than in an airlift reactor. *Biotechnol. Bioeng.* **102**, 1074–1086.
2. Medrano G, Reidy M, Liu J, et al (2009) Rapid system for evaluating bioproduction capacity of complex pharmaceutical proteins in plants. In: Faye L, and Gomord V (eds). Methods Mol Biol Humana Press, Totowa, NJ **483**, 51–67.
3. Medrano G, Dolan MC, Stephens NT, et al (2010) Efficient plant-based production of chicken interleukin-12 yields a strong immunostimulatory cytokine. *J. Interferon Cytokine Res.* **30**, 143–154.
4. Condori J, Medrano G, Sivakumar G, et al (2009) Functional characterization of a stilbene synthase gene using a transient expression system in planta. *Plant Cell Rep.* **28**, 589–599.
5. Cramer CL, Boothe JG, Oishi KK (1999) Transgenic plants for therapeutic proteins: linking upstream and downstream strategies. *Curr. Top. Microbiol. Immunol.* **240**, 95–118.
6. Cramer CL, Weissenborn DL, Oishi KK, et al (1996) Bioproduction of human enzymes in transgenic tobacco. *Ann. N.Y. Acad. Sci.* **792**, 62–71.
7. Liu J, Dolan MC, Reidy M, et al (2008) Expression of bioactive single-chain murine IL-12 in transgenic plants. *J. Interferon Cytokine Res.* **28**, 381–392.
8. Daniell H, Singh ND, Mason H, et al (2009) Plant-made vaccine antigens and biopharmaceuticals. *Trends Plant Sci.* **14**, 669–679.
9. Tremblay R, Wang D, Jevnikar AM, et al (2010) Tobacco, a highly efficient green bioreactor for production of therapeutic proteins. *Biotechnol. Adv.* **28**, 214–221.
10. Wang DJ, Brandsma M, Yin Z, et al (2008) A novel platform for biologically active recombinant human interleukin-13 production. *Plant Biotechnol. J.* **6**, 504–515.
11. Matoba N, Kajiura H, Cherni I, et al (2009) Biochemical and immunological characterization of the plant-derived candidate human immunodeficiency virus type 1 mucosal vaccine CTB-MPR. *Plant Biotechnol. J.* **7**, 129–145.
12. Conley AJ, Joensuu JJ, Menassa R, et al (2009) Induction of protein body formation in plant leaves by elastin-like polypeptide fusions. *BMC Biol.* **7**, 48.
13. Alvarez ML, Topal E, Martin F, et al (2010) Higher accumulation of F1-V fusion recombinant protein in plants after induction of protein body formation. *Plant Mol. Biol.* **72**, 75–89.
14. Hood EE, Bailey MR, Beifuss K, et al (2003) Criteria for high-level expression of a fungal laccase gene in transgenic maize. *Plant Biotechnol. J.* **1**, 129–140.

15. Torrent M, Llompart B, Lasserre-Ramassamy S, et al (2009) Eukaryotic protein production in designed storage organelles. *BMC Biol.* 7, 5.
16. Schillberg S, Twyman RM, Fischer R (2005) Opportunities for recombinant antigen and antibody expression in transgenic plants – technology assessment. *Vaccine* **23**, 1764–1769.
17. Reed DG, Nopo-Olazabal LH, Funk V, et al (2005) Expression of functional hexahistidine-tagged ricin B in tobacco. *Plant Cell Rep.* **24**, 15–24.
18. Hood E, Cramer CL, Medrano G, et al (2012) Protein targeting: strategic planning for optimizing protein products through plant biotechnology. In: Altman A, Hasegawa PM (eds). Plant Biotechnology and Agriculture: Prospects for the 21st Century. San Diego, CA: Elsevier Inc. doi:10.1016/B978-0-12-381466-1.00003-1.
19. Lambrecht B, Gonze M, Meulemans G, et al (2000) Production of antibodies against chicken interferon-gamma: demonstration of neutralizing activity and development of a quantitative ELISA. *Vet. Immunol. Immunopathol.* **74**, 137–144.
20. Shaaltiel Y, Bartfeld D, Hashmueli S, et al (2007) Production of glucocerebrosidase with terminal mannose glycans for enzyme replacement therapy of Gaucher's disease using a plant cell system. *Plant Biotechnol. J.* **5**, 579–590.
21. Jeoh T, Michener W, Himmel ME, et al (2008) Implications of cellobiohydrolase glycosylation for use in biomass conversion. *Biotechnol. Biofuels* **1**, 10.

Chapter 30

Cell-Free Protein Synthesis as a Promising Expression System for Recombinant Proteins

Xumeng Ge and Jianfeng Xu

Abstract

Cell-free protein synthesis (CFPS) has major advantages over traditional cell-based methods in the capability of high-throughput protein synthesis and special protein production. During recent decades, CFPS has become an alternative protein production platform for both fundamental and applied purposes. Using *Renilla luciferase* as model protein, we describe a typical process of CFPS in wheat germ extract system, including wheat germ extract preparation, expression vector construction, in vitro protein synthesis (transcription/translation), and target protein assay.

Key words: Cell-free protein synthesis, In vitro transcription, In vitro translation, Wheat germ extract, Renilla luciferase

1. Introduction

Cell-free protein synthesis (CFPS) is based on the demonstration that cell integrity is not necessary for protein translation (1). Instead of relying on living cells, CFPS is carried out in an artificial buffer supplemented with a crude cell lysate, an energy supply, amino acids, and an exogenously mRNA or DNA template.

Compared with cell-based systems, the biggest advantage is that CFPS offers the quickest way to link a phenotype (protein) to a genotype (gene); protein synthesis and functional assays can be carried out in a few hours (1). In the postgenomic era, CFPS has the potential to become one of the most important high-throughput technologies for functional genomics and proteomics (2–4). Moreover, CFPS is independent of host cells, thus able to produce special proteins, including membrane proteins (5), cell-toxic proteins (6),

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_30, © Springer Science+Business Media, LLC 2012

isotope-labeling proteins (7), and novel proteins with unnatural amino acids (8, 9). Furthermore, the recent advance in improving cell-free translation efficiency, such as the development of continuous-flow CFPS system, has resulted in protein yields that exceed 1 mg/mL, making CFPS a competitive alternative to the current cell-based methods for productive purposes (10–12).

CFPS systems are often prepared from the cell extracts of *Escherichia coli*, wheat germs, and rabbit reticulocytes (1); all these crude extracts containing the macromolecular components (70S or 80S ribosomes, tRNAs, aminoacyl-tRNA synthetases, initiation, elongation and termination factors, etc.) are necessary for translation. Alternatively, the whole translational mechanism could be reconstituted from individually purified macromolecular components (13). In this chapter, the wheat germ-based CFPS systems are used to express a model protein *Renilla luciferase* (Rluc) because (1) they are low cost; (2) they are eukaryotic systems, and thus more suitable for the expression of eukaryotic proteins (14). Four procedures necessary for CFPS are outlined, including: (1) preparation of wheat germ extract; (2) preparation of DNA template; (3) preparation of reactions for protein synthesis; and (4) characterization of expressed protein.

2. Materials

2.1. Preparation of Wheat Germ Extract (see Note 1)

1. Flotation solvent: Cyclohexane and carbon tetrachloride (240:600, v/v) (see Note 2).
2. Extraction buffer: 40 mM HEPES–KOH, pH 7.8, 100 mM KOAc, 1 mM $Mg(OAc)_2$, 2 mM $CaCl_2$, and 4 mM DTT.
3. Column buffer: 40 mM HEPES–KOH, pH 7.8, 100 mM KOAc, 5 mM $Mg(OAc)_2$, and 4 mM DTT.
4. 100 mM spermidine.
5. 500 mM creatine phosphate (CP).
6. 0.5% (w/v) Nonidet P-40.
7. G-25 (fine) column (1 × 30 cm).
8. Ultracentrifuge.
9. Sterile Milli-Q (or double distilled) H_2O.

2.2. Preparation of Expression Vector

1. *pIVEX1.3WG* expression vector with a T7 promoter (Roche).
2. *pRL-Null* plasmid containing Rluc cDNA (Promega).
3. DNA Miniprep Kit (QIAprep Spin Miniprep Kit, Qiagen).
4. DNA Gel Extraction Kit (QIAEXII Gel Extraction Kit, Qiagen).
5. PCR thermal cycler.

6. Thin-wall PCR tubes.
7. PCR 2× Master Mix (New England BioLabs).
8. Luria-Bertani (LB) medium: Add 1.5% (w/v) agar for solid medium.
9. Competent *E. coli* Top10 (Invitrogen).
10. T4 DNA Ligase: 2–3 U/μL (Promega).
11. Restriction enzymes: *Nco*I, *Xma*I, *Eco*RI (New England BioLabs).
12. 1,000× ampicillin: 50 mg/mL in water (use at a final concentration of 50 μg/mL).
13. 1% (w/v) agarose gel.
14. Ethidium bromide.
15. Electrophoresis apparatus.
16. TAE buffer 10× stock: 40 mM Tris–HCl, pH 8.0, 1 mM EDTA–NaOH, and 11.42 mL/L glacial acetic acid.
17. Microcentrifuge.
18. Phenol/chloroform extraction solution: Phenol:chloroform: isoamyl alcohol (25:24:1 v/v).
19. Nuclease-free H_2O.

2.3. Cell-Free Protein Synthesis

1. Transcription buffer (5×): 200 mM HEPES–KOH, pH 7.8, 30 mM $Mg(OAc)_2$, 10 mM spermidine, and 50 mM NaCl.
2. Energy-generating buffer (10×): 40 mM HEPES–KOH, pH 7.8, 4 mM spermidine, 12 mM ATP, 3 mM GTP, 180 mM CP, and 5.5 mg/mL creatine phosphokinase (CPK).
3. Translation premix (20×): 500 mM HEPES–KOH, pH 7.8, 30 mM DTT, 5 mM spermidine, 20 mM ATP, 8 mM GTP, and 320 mM CP.
4. Transcription–translation premix (20×): 500 mM HEPES–KOH, pH 7.8, 30 mM DTT, 9 mM spermidine, 20 mM ATP, 8 mM GTP, 8 mM CTP, 8 mM UTP, and 320 mM CP.
5. Amino acids mixture: 1 mM each of 20 amino acids.
6. Total yeast tRNA (or deacylated tRNA prepared from wheat embryos): 2 mg/mL in 25 mM HEPES–KOH, pH 7.8.
7. Ribonucleotide solution mixture (rNTP): 80 mM each of four ribonucleotides in 20 mM Tris–HCl, pH 7.8, 10 mM DTT.
8. T7 RNA polymerase: 50 U/μL.
9. Ribonuclease inhibitor (RNasin): 40 U/μL (Promega).
10. CPK: 10 mg/mL in 40 mM potassium phosphate buffer, pH 7.5, 10 mM 2-mercaptoethanol, and 50% glycerol (v/v).
11. Nuclease-free H_2O.

2.4. Characterization of Protein

1. *Renilla* Luciferase Assay System kit (Promega). Prepare the Rluc assay reagent by adding 1 volume of Rluc substrate (100×) to 100 volumes of Rluc assay buffer (see Note 3).
2. Siliconized polypropylene tube.
3. Luminometer: Modulus™ System with Luminescence Module (Turner Biosystems).
4. Mini-PROTEAN gel-casting chamber (Bio-Rad).
5. Vertical mini gel electrophoresis apparatus (Bio-Rad).
6. Mini Trans-Blot electrophoretic transfer cell (Bio-Rad).
7. Acrylamide/bis solution (see Note 4): 30% (w/v) acrylamide, 0.8% (w/v) bisacrylamide. Filter the solution with a 0.45-μm filter and store at 4°C in the dark.
8. Separating buffer (4×): 1.5 M Tris–HCl, pH 8.8, and 0.4% SDS. Store at 4°C.
9. Stacking buffer (4×): 0.5 M Tris–HCl, pH 6.8, and 0.4% SDS. Store at 4°C.
10. H_2O-saturated isobutanol (see Note 5).
11. Running buffer (5×): 125 mM Tris, 960 mM glycine, 0.5% (w/v) SDS. Store at room temperature (RT).
12. SDS loading buffer (5×): 125 mM Tris–HCl, pH 6.8, 500 mM DTT, 5% (w/v) SDS, 25% (v/v) glycerol, and 0.1% (w/v) bromophenol blue.
13. *N*,*N*,*N*′,*N*′-Tetramethyl-ethylenediamine (TEMED).
14. Ammonium persulfate: 10% (w/v) (see Note 6).
15. Prestained protein ladder.
16. Coomassie Blue stain solution (1×): 0.1% (w/v) Coomassie Brilliant Blue R-250, 50% (v/v) methanol, and 10% (v/v) glacial acetic acid.
17. Destain solution: 10% (v/v) acetic acid.
18. Transfer buffer: 25 mM Tris, 192 mM glycine, and 20% (v/v) methanol. Store at 4°C.
19. Tris-buffered saline with Tween (TBS-T, 10×): 200 mM Tris–HCl, pH 7.6, 1.37 M NaCl, and 0.5% (v/v) Tween-20.
20. Blocking buffer: 5% (w/v) bovine serum albumin (BSA) in 1× TBS-T.
21. Nitrocellulose transfer membrane.
22. Antibodies: Primary antibody, rabbit anti-Rluc (MBL International), and secondary antibody, goat anti-rabbit IgG (H + L), conjugated with peroxidase (Jackson Immuno Research Laboratories).
23. Rluc standard (NanoLight Technology).

24. SuperSignal® West Pico Chemiluminescent Substrate (Thermo Scientific).
25. X-ray film; darkroom and film developer.
26. Filter paper.
27. Polyvinyl-chloride wrap.

3. Methods

3.1. Preparation of Wheat Germ Extract

Procedures described here are based on the early publication of Erickson and Blobel (15) with modifications suggested by Madin et al. and Shirokov et al. (14, 16).

3.1.1. Preparation of Wheat Embryos

1. Grind wheat seeds using a blender with four grinding pulses of 2 s each for every portion of seeds (2–3 g of embryos are obtained from 1 kg of seeds).
2. Sieve the ground material to select embryo fractions (710- to 850-mm mesh).
3. Add the sieved embryos into the flotation solvent at a ratio of 50 mg/mL in a fume hood. Stir gently with a glass rod. The damaged embryos and contaminants should sediment while intact embryos float. After a 2–3-min separation, collect the floating embryos and dry them overnight in the fume hood at RT.
4. Weigh up the embryos and wash them three times with sterile Milli-Q H_2O (10 mL per gram of embryos) under vigorous stirring to remove contaminating endosperm (see Note 7).
5. Sonicate the embryos for 3 min in the Nonidet P-40 solution.
6. Wash once more in the sonicator with sterile Milli-Q H_2O.
7. Immediately transfer the washed embryos into liquid nitrogen or store them at −80°C as small aliquots.

3.1.2. Preparation of Wheat Germ Extract (Fig. 1) (see Note 8)

1. Grind 5 g of the embryos into a fine powder in liquid nitrogen (see Note 9).
2. Add 5 mL of extraction buffer to the frozen powder and thoroughly mix to saturate powder with buffer. Allow frozen powder/buffer mix to thaw before further grinding it to a homogeneous paste.
3. Transfer the mixture into centrifuge tubes, vortex briefly, and centrifuge at 30,000 × g at 4°C for 30 min.
4. Collect the liquid phase between the floating lipid layer and pellet (see Note 10).

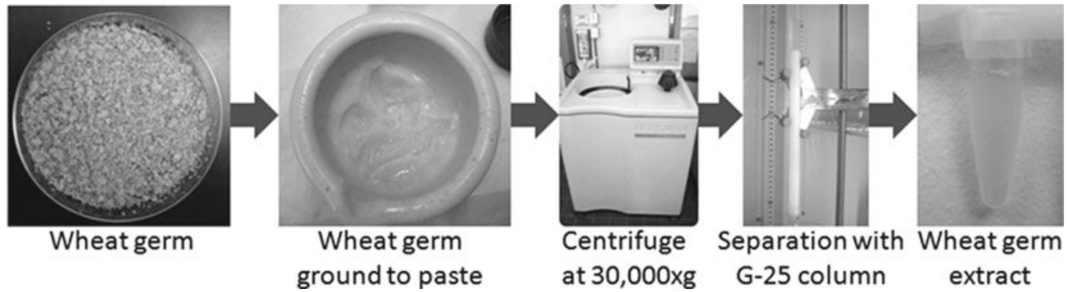


Fig. 1. Flowchart of the procedure to prepare wheat germ extract.

5. Measure the volume of the extract obtained and load it onto the G-25 (fine) column equilibrated with column buffer. After sample absorption, elute with two volumes of column buffer and collect the eluate in about 15 fractions in Eppendorf tubes. Combine the most opaque fractions to obtain a solution with approximately equal volume to the extract loaded onto the column.
6. Centrifuge the solution at 30,000 × *g* at 4°C for 10 min and collect the supernatant.
7. Adjust the concentration of the extract to 200 A_{260}/mL with column buffer, aliquot, and store at −80°C immediately.

3.2. Preparation of Expression Vector

The *pIVEX1.3WG* vector containing optimized 5′- and 3′-UTRs for protein expression in wheat germ-based CFPS systems is used (Fig. 2). Standard molecular cloning procedures are followed to subclone the *Rluc* gene into the *pIVEX1.3WG* vector, which are not described in detail here due to space limitations. Briefly, the *Rluc* gene is amplified from plasmid *pRL-null* by PCR, and then subcloned into the *pIVEX1.3WG* vector as a *Nco*I/*Xma*I fragment to generate *pIVEX1.3RL* (10, 11). The plasmid *pIVEX1.3RL* is then linearized with *Eco*RI and purified with phenol/chloroform extraction before used as a template in CFPS systems.

3.3. Cell-Free Protein Synthesis

CFPS can be performed in three modes: (1) uncoupled transcription/translation (Fig. 3a); (2) coupled transcription/translation (Fig. 3b); and (3) continuous-exchange cell-free (CECF) system (Fig. 3c). The protocols for each mode of CFPS are shown below.

3.3.1. Uncoupled Transcription/Translation

1. Prepare the in vitro transcription mixture as directed in Table 1 and incubate at 37–40°C for 1 h.
2. Prepare the in vitro translation mixture as directed in Table 2 (see Note 11) and incubate at 25–30°C for 1–2 h.
3. Store the reaction mixture at −80°C for protein characterization.

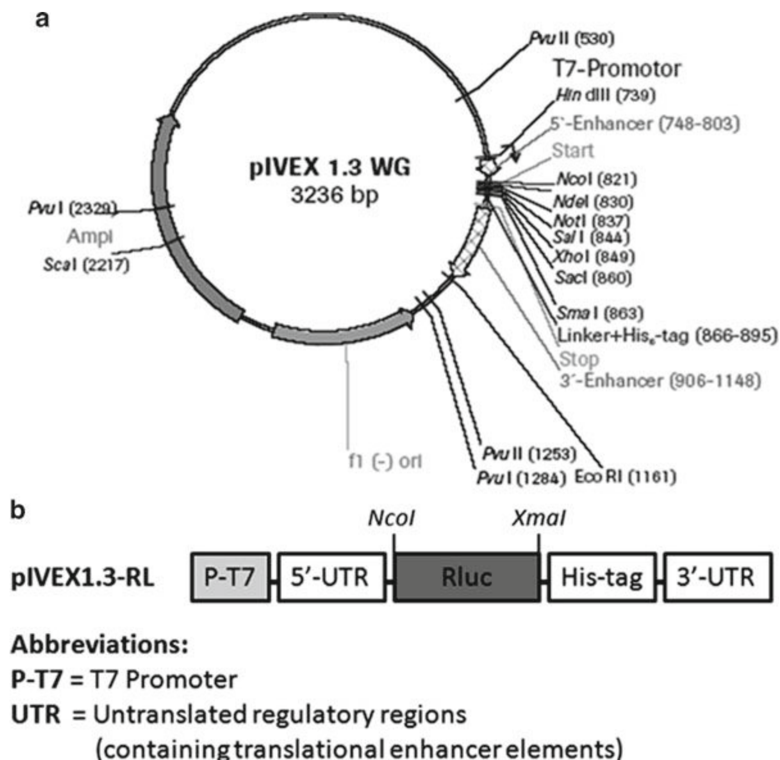


Fig. 2. Map of *pIVEX1.3WG* vector (**a**) and functional elements of *pIVEX1.3-RL* plasmid (**b**). The *pIVEX1.3-RL* plasmid is generated by subcloning the *Rluc* gene into the *pIVEX1.3WG* vector at the *Nco*I and *Xma*I sites.

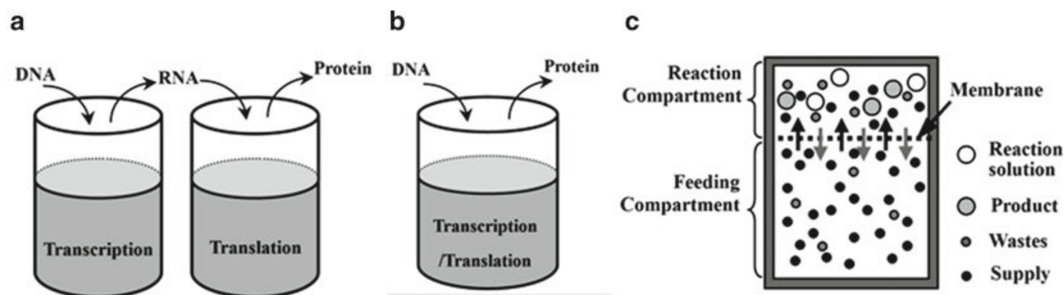


Fig. 3. Schematic illustrations of different CFPS modes. (**a**) Uncoupled in vitro transcription/translation; (**b**) coupled in vitro transcription/translation system; (**c**) CECF mode.

3.3.2. Coupled Transcription/Translation

1. Prepare the transcription/translation mixture as directed in Table 3 (see Note 11) and incubate at 30°C for 1–2 h.
2. Store the reaction mixture at −80°C for protein characterization.

3.3.3. Continuous-Exchange Cell-Free System

The main limitation of the batch-formatted reactions is their short lifetimes (<1 h) and consequently low yields. This is due to the rapid depletion of the high-energy phosphate pool and the accumulation of free phosphates which inhibit protein synthesis (17).

Table 1
In vitro transcription reaction

Components	Volume (μL)
5× transcription buffer	4
100 mM DTT	2
40 U/μL RNasin	1
8 mM rNTP	1.25
0.5–1 μg/μL linearized plasmid	1
50 U/μL T7 RNA polymerase	0.4
Nuclease-free H_2O	10.35
Total	20

Table 2
In vitro translation reaction

Components	Volume (μL)
Wheat germ extract	10
10× energy-generating buffer	2
1 mM amino acid mixture	3.2
40 U/μL ribonuclease inhibitor	1
2 mg/mL tRNA	0.5
0.1–0.5 μg/mL mRNA	1
Nuclease-free H_2O	2.3
Total	20

Table 3
Coupled transcription/translation reaction

Components	Volume (μL)
Wheat germ extract	10
10× energy-generating buffer	2
50 U/μL T7 RNA polymerase	0.5
1 mM amino acid mixture	1.6
40 U/μL RNasin	0.4
0.5–1 μg/μL linearized plasmid	1
Nuclease-free H_2O	4.5
Total	20

Table 4
Continuous-exchange cell-free (CECF) reaction

Compartments	Components	Volume (μL)
Feeding solution	20× transcription–translation premix	50
	100 mM $Mg(OAc)_2$	30
	2.5 M KOAc	28
	1 mM amino acid mixture	100
	Nuclease-free H_2O	792
	Total	1,000
Reaction solution	20× transcription–translation premix	2.5
	100 mM $Mg(OAc)_2$	1
	2.5 M KOAc	0.8
	1 mM amino acid mixture	5.0
	40 U/μL RNasin	1
	10 mg/mL CPK	0.5
	50 U/μL T7 RNA polymerase	2.4
	0.5–1 μg/μL linearized plasmid	2.5
	Wheat germ extract	25
	2 mg/mL tRNA	1.5
	Nuclease-free H_2O	7.8
	Total	50

CECF system overcomes this problem with continuous exchange of substrates and by-products (Fig. 3c) (1).

1. Prepare the reaction solution and feeding solution as directed in Table 4 (see Note 12).
2. Add the feeding solution into the feeding compartment. Transfer the reaction solution into the reaction compartment. Close the compartments with adhesive film.
3. Shake at 900 rpm and 24°C for up to 24 h.
4. Store the reaction mixture at −80°C for protein characterization.

3.4. Characterization of Protein

3.4.1. Rluc Functional Assay

1. Dilute the sample (1:10 or 1:100) with the lysis solution of the *Renilla* Luciferase Assay System kit.
2. Add 5 μL of diluted sample into a siliconized polypropylene tube.
3. Add 45 μL of Rluc assay reagent to the tube and mix quickly by flicking the tube with a finger (1–2 s).
4. Place the tube in a luminometer to measure the relative luminescence units (RLUs).

3.4.2. SDS-Polyacrylamide Gel Electrophoresis

1. Assemble the glass-plate sandwich of the Mini-Protean gel-casting chamber according to manufacturer's instructions (see Note 13).

2. Prepare separating gel solution (see Note 14) by mixing 4.8 mL of acrylamide/bis solution, 3 mL of separating buffer, 4.2 mL of H_2O, 40 μL of ammonium persulfate, and 10 μL of TEMED in a 50-mL centrifuge tube. Swirl gently to mix and use immediately.
3. Pour the separating gel solution to the sandwich, leaving space for a stacking gel (~0.5 cm), and gently cover the top of the gel with a layer (~0.5 cm thick) of H_2O-saturated isobutanol. Allow the gel to polymerize in about 30 min at RT.
4. Pour off the isobutanol and rinse the top of the gel extensively with distilled water (see Note 15).
5. Prepare stacking gel solution (see Note 14) by mixing 520 μL of acrylamide/bis solution, 1 mL of stacking buffer, 2.44 mL of H_2O, 20 μL of ammonium persulfate, and 5 μL of TEMED in a 15-mL centrifuge tube. Swirl gently to mix and use immediately.
6. Pour the stacking gel solution onto the separating gel until overflows and insert the comb. Allow the stacking gel solution to polymerize in about 30 min at RT (see Note 16).
7. Carefully remove the comb without tearing the edges of the polyacrylamide wells. Rinse wells with 1× running buffer (see Note 17).
8. Arrange the gel sandwich in the vertical mini gel electrophoresis apparatus. Add the running buffer to the upper and lower chambers of the gel unit.
9. Prepare protein samples by mixing 12 μL of the reaction mixtures with 3 μL of 5× SDS loading buffer in 1.5-mL centrifuge tubes and placing the tubes in the boiling water bath for 5 min.
10. Load the samples on the polyacrylamide gel. Include one well for prestained molecular weight markers.
11. Run the gel at 100 V until the bromophenol blue tracking dye enters the separating gel. Then, increase the voltage to 150 V until the bromophenol blue tracking dye reaches the bottom of the separating gel.
12. Stain the gel in 10 volumes of Coomassie Blue stain solution for at least 30 min at RT.
13. Destain the gel in 10 volumes of destain solution at room temperature long enough so that the background is clear (see Note 18).
14. Take a picture of the gel image (Fig. 4) or dry the gel to keep a record.

3.4.3. Western Blotting

1. Cut the nitrocellulose transfer membrane and filter paper to the dimensions of the gel (see Note 19). Submerge the gel, membrane, filter paper, and fiber pads in transfer buffer for 15 min to 1 h.
2. Transfer the proteins from the polyacrylamide gel to the nitrocellulose membrane at 100 V for 1 h using a Mini Trans-Blot electrophoretic transfer cell following manufacturer's procedures.
3. Incubate the membrane in 10 mL of blocking buffer for 1 h at RT on a rocking platform.
4. Discard the blocking buffer. Add 8 mL of TBS-T, 2 mL of fresh blocking buffer, and the primary antibody (1:5,000 v/v), mix thoroughly, and incubate for 1 ~ 3 h at RT on the rocking platform.
5. Discard the solution and wash the membrane three times for 15 min each time with 100 mL of TBS-T solution.
6. Add 8 mL of TBS-T, 2 mL of fresh blocking buffer, and the secondary antibody (1:15,000 v/v), mix thoroughly, and incubate for 1 h at RT on the rocking platform.
7. Discard the solution and wash the membrane three times for 15 min each time with 100 mL of TBS-T solution.
8. Place the membrane with the protein side facing up on a polyvinylchloride wrap. Mix the two chemiluminescent substrate solutions with equal volume (each 750 μL) in a 1.5-mL centrifuge tube. Apply the mixture evenly on to the membrane. Immediately cover the membrane with the second sheet of the polyvinyl-chloride film to spread the substrate evenly and without air bubbles over the membrane. Incubate for 5 min at RT.
9. Squeeze out excess liquid and seal the edges of the film.
10. Expose to the X-ray film for 10 s to 10 min (Fig. 4) (see Note 20).

3.4.4. Quantification of Expressed Protein

The total amount of produced Rluc is measured by Western blotting. The Rluc standard is used to construct the calibration curve, and electrophoresed at the same gel with the samples. Then, the concentration of the expressed Rluc in the samples can be determined by the comparison of band intensities and the calibration curve (11). The active amount of Rluc is determined as in Subheading 3.4.1. The measured RLU is converted to weight concentration according to a calibration curve which can be obtained with the Rluc standard (11). The yields of Rluc produced from three CFPS modes are summarized in Table 5. CECF can produce protein yields 7- to 16-fold higher than those produced from the batch-formatted reactions.

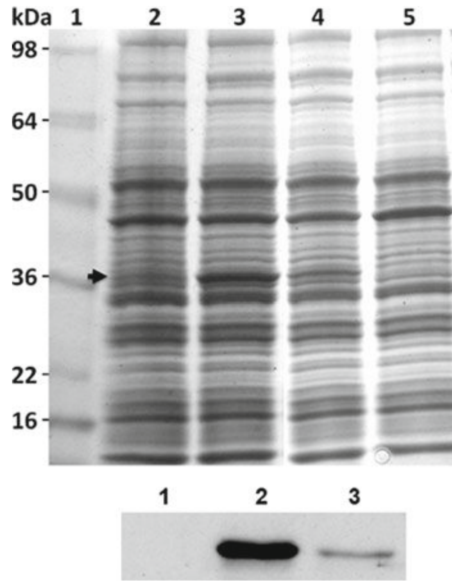


Fig. 4. SDS-PAGE (*top*) and Western blotting (*bottom*) analysis of the Rluc protein expressed in wheat germ-based CFPS systems. In the SDS-PAGE image (*top*): *Lane 1*, molecular marker; *Lane 2*, Rluc expressed in *uncoupled* transcription/translation mode; *Lane 3*, Rluc expressed in CECF mode; *Lane 4*, Rluc expressed in *coupled* transcription/translation mode; *Lane 5*, negative control (*coupled* transcription/translation solution without DNA template). In the Western blotting image (*bottom*): *Lane 1*, negative control; Lane 2, Rluc expressed in CECF mode; *Lane 3*, Rluc expressed in *uncoupled* transcription/translation system.

Table 5
Yields of Rluc expressed in different CFPS modes

CFPS modes	Total protein (mg/mL)	Active protein (mg/mL)	Percentage of active protein
Uncoupled transcription/translation	0.065	0.024	36
Coupled transcription/translation	0.15	0.052	35
CECF	1.08	0.32	30

4. Notes

1. Make solutions with Milli-Q (or double distilled) H_2O and sterilize by filtration through 0.22-μm filter. Store at –20°C.
2. Stir until no schlieren mixing lines are visible.
3. The Rluc assay reagent is stable for 12 h at RT, 2 weeks at –20°C, or 1 month at –70°C.
4. Acrylamide monomer is neurotoxic. Wear mask when weighing acrylamide powder. Wear gloves while handling the solution.

5. H_2O-saturated isobutanol is made by shaking equal volumes of H_2O and isobutanol in a glass bottle and allowing the solution to separate. Collect the top layer and store at RT.
6. This solution should be frozen in single-use aliquots at −20°C.
7. Mechanical or manual stirring rather than magnetic stirring bar should be used to avoid damaging the embryos.
8. All these procedures should be carried out in the minimum amount of time in a 4°C cold room; all the solutions and glassware are required to be sterilized to minimize contamination of the extract by RNases.
9. The mortar and pestle should be chilled with liquid nitrogen or in −80°C freezer before use. Do NOT let the powder thaw during grinding process.
10. Care should be taken to avoid transferring the floating lipid layer into the tube.
11. Additional potassium can be added into the translation mixture. The optimal concentrations of potassium vary from 50 to 200 mM, corresponding to different wheat germ extract preparations or different DNA or mRNA used.
12. The optimal concentrations of magnesium and potassium may vary with different wheat germ extract preparations or different DNA or mRNA used, which can be determined by carrying out a set of batch reactions at different salt concentrations.
13. It is critical that the glass plates for the gels are scrubbed clean with a rinsable detergent after use and rinsed extensively with distilled water. They can be kept clean until use in a plastic rack in 30% nitric acid (use caution when removing). They just need rinsing (distilled water and then 95% ethanol) to remove the acid and air dry.
14. The recipes are adequate for producing two gels of dimensions 7 cm × 8 cm × 0.75 mm, each gel consisting of 12 mL of separating gel and 4 mL of stacking gel.
15. Residual isobutanol can reduce resolution of the protein bands; therefore, it must be completely removed. The isobutanol overlay should not be left on the gel longer than 2 h.
16. A sharp optical discontinuity is visible around wells on polymerization.
17. The rinse removes unpolymerized monomer; otherwise, the monomer will continue to polymerize after the comb is removed, creating uneven wells that interfere with sample loading and subsequent separation.
18. Destaining can be lasted for long periods as long as there is solution remaining in the plate. A piece of foam or paper towel may be added to the solution to speed up the destaining process.

19. Always wear gloves and use tweezers when handling membranes to prevent contamination.
20. Multiple exposures can be taken to achieve the desired signal strength.

Acknowledgments

This work is supported by Arkansas NSF-EPSCoR funds and the Arkansas Biosciences Institute.

References

1. Katzen, F., Chang, G., and Kudlicki, W. (2005) The past, present and future of cell-free protein synthesis, *Trends Biotechnol.* 23, 150–156.
2. He, M. Y. (2008) Cell-free protein synthesis: applications in proteomics and biotechnology, *N. Biotechnol. 25*, 126–132.
3. Sawasaki, T., Ogasawara, T., Morishita, R., et al., (2002) A cell-free protein synthesis system for high-throughput proteomics, *Proc. Natl. Acad. Sci. USA* 99, 14652–14657.
4. Jackson, A. M., Boutell, J., Cooley, N., et al., (2004) Cell-free protein synthesis for proteomics, *Brief Funct. Genomic. Proteomic.* 2, 308–319.
5. Klammt, C., Schwarz, D., Lohr, F., et al., (2006) Cell-free expression as an emerging technique for the large scale production of integral membrane protein, *FEBS J.* 273, 4141–4153.
6. Xu, Z. N., Chen, H. Q., Yin, X. F., et al., (2005) High-level expression of soluble human beta-defensin-2 fused with green fluorescent protein in Escherichia coli cell-free system, *Appl. Biochem. Biotech.* 127, 53–62.
7. Vinarov, D. A., Lytle, B. L., Peterson, F. C., et al., (2004) Cell-free protein production and labeling protocol for NMR-based structural proteomics, *Nat. Methods* 1, 149–153.
8. Bundy, B. C., and Swartz, J. R. (2010) Site-Specific Incorporation of p-Propargyloxyphenylalanine in a Cell-Free Environment for Direct Protein-Protein Click Conjugation, *Bioconjug. Chem.* 21, 255–263.
9. Chattopadhaya, S., Tan, L. P., and Yao, S. Q. (2006) Strategies for site-specific protein biotinylation using in vitro, in vivo and cell-free systems: toward functional protein arrays, *Nat. Protoc.* 1, 2386–2398.
10. Park, N., Kahn, J. S., Rice, E. J., et al., (2009) High-yield cell-free protein production from P-gel, *Nat. Protoc.* 4, 1759–1770.
11. Park, N., Um, S. H., Funabashi, H., et al., (2009) A cell-free protein-producing gel, *Nat. Mater.* 8, 432–437.
12. Kim, T. W., Kim, H. C., Oh, I. S., et al.,(2008) A highly efficient and economical cell-free protein synthesis system using the S12 extract of Escherichia coli, *Biotechnol. Bioprocess Eng.* 13, 464–469.
13. Shimizu, Y., Inoue, A., Tomari, Y., et al., (2001) Cell-free translation reconstituted with purified components, *Nat. Biotechnol.* 19, 751–755.
14. Madin, K., Sawasaki, T., Ogasawara, T., et al.,(2000) A highly efficient and robust cell-free protein synthesis system prepared from wheat embryos: plants apparently contain a suicide system directed at ribosomes, *Proc. Natl. Acad. Sci. USA* 97, 559–564.
15. Erickson, A. H., and Blobel, G. (1983) Cell-free translation of messenger RNA in a wheat germ system, *Meth. Enzymol.* 96, 38–50.
16. Shirokov, V.A., Kommer, A., Kolb, V.A., et al., (2007) Continuous-exchange protein-synthesizing systems, *In Vitro Transcription and Translation Protocols* (Second Edition), 19–55. Humana Press.
17. Kim, D.M., and Swartz, J.R. (1999) Prolonging cell-free protein synthesis with a novel ATP regeneration system, *Biotechnol. Bioeng.* 66, 180–188.

Part VI

Animals and Animal Cells

Chapter 31

The Use of Bacterial Artificial Chromosomes for Recombinant Protein Production in Mammalian Cell Lines

Leander Blaas, Monica Musteanu, Beatrice Grabner, Robert Eferl, Anton Bauer, and Emilio Casanova

Abstract

The choice of an expression vector is a critical step in the field of recombinant protein production in mammalian cells lines. Most expression vectors used in the field are sensitive to the surrounding chromatin to their integration site into the host genome cell. This so-called chromatin positional effects influences the expression levels of the transgene and tends to silence its expression over time. Bacterial artificial chromosomes (BACs) are vectors that can accommodate inserts of up to 400 kb. Due to the large cloning capacity, BACs can harbour an entire locus with all or most of the regulatory elements controlling the expression of a gene. Therefore, BACs contain their own natural chromatin domain and are subjected to chromatin positional effects to a lesser extend or not at all. This makes cell lines generated with BAC-based expression vectors more predictable in terms of protein production and stability. In this chapter, we explore the use of BACs as expression vectors for recombinant protein production in mammalian cells.

Key words: Rosa26 locus, Chromatin positional effects, Homologous recombination, Cassette exchange

1. Introduction

Recombinant protein production in mammalian cells constitutes an important topic in biology and economics, consequently a lot of research has been undertaken to develop efficient methods for protein production. Several key aspects, such as cell hosts, expression vectors, transfection, selection, and metabolic engineering, have been investigated (1, 2). In this chapter, we focus on the design and use of bacterial artificial chromosomes (BACs) as expression vectors for recombinant protein production in mammalian cells.

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_31, © Springer Science+Business Media, LLC 2012

Generation of stable mammalian cell lines producing high levels of a specific protein is typically a cumbersome and time-consuming procedure (2). In the simplest approaches, stable cell lines are generated by random integration into the genome of the host cell of an expression vector containing a promoter, a gene of interest, a polyadenylation signal, and a selection marker. Although this method is straightforward, stochastically integrated vectors in the host cell genome are greatly affected by the so-called chromatin positional effects (3). Chromatin surrounding the integration site may negatively influence expression levels of the transgene or silence its expression over time. This makes the screening for clones with high and stable expression an onerous process. To overcome this problem, expression vectors can be flanked by DNA sequences that may protect against undesired chromatin effects, such as insulators (4), matrix attachment regions (5), chromatin opening elements (6), antirepressor elements (7), and others. Another attractive strategy relies on targeting the expression vector to an endogenous locus, which is not subjected to chromatin effects (8). However, since homologous recombination is a rare event in somatic cells, the target cell line needs to be pre-engineering (for example, targeting the locus with the necessary elements to perform cassette exchange with the gene of interest), thus limiting the flexibility of this system. On the other hand, in this system only a single copy of the expression vector is integrated, thus limiting the level of expression prior to gene amplification.

Ideally, an expression vector should have three features: (1) it should not be affected by the chromatin surrounding its integration site, (2) expression levels should correlate with the number of integrated copies, and (3) expression levels should be maintained over time. Interestingly, scientists working in the mouse transgenic field have faced similar problems to those working on recombinant protein production (3). Transgenic mice generated by random integration of “conventional expression vectors” (promoter, gene of interest, and polyadenylation signal) can have mosaic expression (some cells express the transgene, others not), expression may be silenced over time, or they may not express the transgene at all (3). It is established that these problems are due to unwanted chromatin positional effects. Because of this, the mouse transgenic field progressively withdrawn from the use of “conventional expression vectors” and increasingly employs large vectors (up to several hundreds of kilobases), which are able to accommodate a desired locus. Such large vectors include most, if not all, of the regulatory elements required for the transcription of a gene, thus carrying their own chromatin context and leaving them protected against chromatin positional effects. Large vectors, such as yeast artificial chromosomes (YACs), BACs, and P1-derived artificial chromosomes (PACs), are currently widely use in the transgenic field. They confer faithful expression of a gene of interest, expression levels correlate to the number of copies integrated, expression is maintained over

time and independent of the genomic integration site (i.e. they are not influenced by the surrounding chromatin) (3). Therefore, these vectors fulfil the above mentioned criteria of an ideal expression vector applicable to recombinant protein production.

Although YACs, BACs, and PACs have been used in the mouse transgenic field, BACs have obvious advantages for the recombinant protein expression field: well-characterized BAC libraries exist, BACs can accommodate up to 400 kb and, in contrast to YACs, BACs are very simple to manipulate. The advantage of BACs as expression vectors lies in their capacity to carry on their own chromatin environment, thus in terms of designing an expression vector for mammalian cells, it is very important to choose the right locus present in the BAC. In this context, genes, which are highly transcribed in the host cells, or the so-called open chromatin loci are good candidates to drive the expression of the gene of interest. In our earlier work, we have shown that the combination of a BAC containing the Rosa26 locus and an artificial promoter (CAGGS) increased recombinant protein production by a factor of 10 (9). Recently, another report using a BAC containing the dihydrofolate reductase locus has confirmed and extended our results (10). The Rosa26 locus contains a gene of unknown function which is accepted to be expressed ubiquitously, it is widely used to express transgenes in mice and is considered to be an open chromatin region (11). This Rosa26 locus-containing BAC improves protein production due to its position-independent and copy number-dependent expression. Other open chromatin loci containing BACs, which might be a good candidate for expression vectors are the β-actin gene of ribosomal proteins, such as Rpl23a (a gene that is high expressed in HEK293 cells (12)).

Possibly, for scientists who are not familiar with the BAC field the biggest challenge would be the modification of a BAC in order to express a gene of interest. Due to their large size, BACs cannot be modified by classical cloning procedures like digestion with a restriction enzyme, ligation, etc., because restriction enzymes like *Eco*RI may have numerous recognition sites within a BAC. BACs can be modified in one of two ways: by homologous recombination in *Escherichia coli* (“recombineering”) (13, 14) or by recombinase-meditated cassette exchange in a previously customized BAC (15). Either way, modification of BAC is straightforward and can be performed in a regular molecular biology laboratory. Nowadays, it is even performed in high-throughput approaches (16, 17).

2. Materials

1. Luria-Bertani medium (LB, LB agar).
2. 50 mg/mL Ampicillin stock in water. Store at −20°C.
3. 50 mg/mL Kanamycin stock in water. Store at −20°C.

4. 12.5 mg/mL Chloramphenicol stock in ethanol 96%. Store at –20°C.
5. 2 mg/mL Neomycin (G418) stock in water. Store at –20°C.
6. 5 mg/mL Tetracycline stock in ethanol 96%. Store at –20°C protected from light.
7. Electroporation cuvettes, 0.2 cm gap.
8. Gene pulser apparatus (BioRad).
9. 20% L(+)-Arabinose stock in water, freshly prepared.
10. 10% Glycerol.
11. P1 buffer: 50 mM Tris–HCl pH 8.0, 10 mM EDTA, 100 μg/mL RNase A. Store at 4°C.
12. P2 buffer: 200 mM NaOH, 1% SDS. Store at room temperature (RT).
13. P3 buffer: 3 M Potassium acetate pH 5.5. Store at 4°C.
14. TE buffer: 10 mM Tris–HCl, pH 7.5, and 1 mM EDTA.
15. Isopropanol.
16. 70% Ethanol.
17. 3 M sodium acetate, pH 5.2.
18. Lipofectamine 2000 (Invitrogen).
19. HEK cells (American Type Culture Collection).
20. DH10B *E. coli* strain cells (Invitrogen).

3. Methods

Modification of a BAC with a gene of interest: The first step consists of the identification of a BAC containing an appropriate locus. In our previous work, we have chosen a 200 kb murine BAC containing the Rosa26 gene (*Thumpd3*) because this locus is considered to be an open chromatin region. As mentioned before, other interesting genes are β-actin or ribosomal proteins. Murine and human BACs can be mapped and identified in the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/) or Ensembl (http://www.ensembl.org/), and they can be purchased in the BACPAC Resource Center (BPRC) at the Children's Hospital Oakland Research Institute in Oakland (CHORI, http://bacpac.chori.org/). For a detailed explanation on how to identify and purchase BACs, see ref. 18. When using a BAC as an expression vector for recombinant protein production there are at least two options: (1) Insertion of the gene of interest (GOI) under the control of an "ectopic" promoter (CAGGS, SV40, etc.) in the BAC, i.e. a Rosa26 locus containing BAC, thus taking only the advantage of the BAC chromatin context or

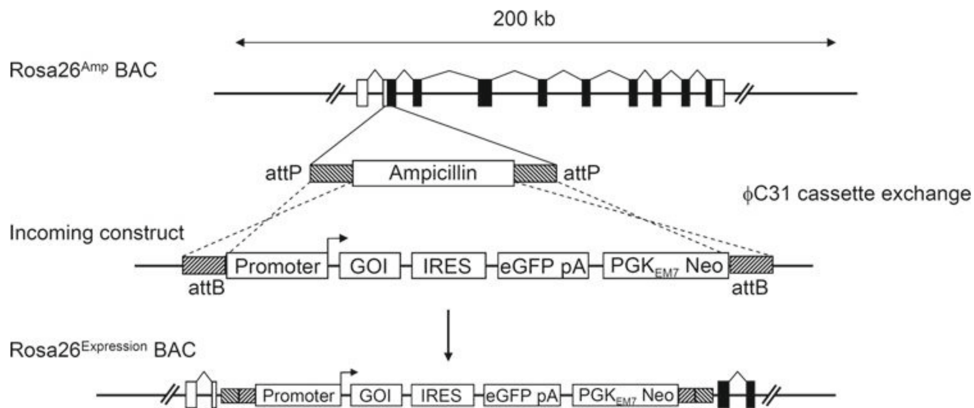


Fig. 1. Schematic representation of ΦC31-mediated cassette exchange in the Rosa26amp BAC. A BAC containing the Rosa26 locus has been previously adapted for ΦC31-mediated cassette exchange by introducing an ampicillin cassette flanked by attP sites in the exon 2. The incoming construct contains a promoter (e.g. CAGGS promoter) a gene of interest (GOI), an internal ribosomal binding sequence (IRES), an eGFP followed by and SV40 polyadenylation signal and a selection cassette containing the phosphoglycerate kinase promoter (PGK), a bacterial promoter (EM7), and the neomycin resistance gene. This selection cassette confers resistance to kanamycin in prokaryotes and to neomycin in eukaryotes. The incoming construct is flanked by two attB sites and upon expression of the ΦC31 integrase, the attP and attB sites recombine, exchanging the ampicillin gene for the incoming construct into the BAC. *White rectangles* describe non-coding regions of the Rosa26 locus (*Thumpd3*) and *black rectangles* correspond to coding regions.

(2) placement of the GOI into the BAC under the control of an endogenous promoter of a highly transcribed gene present in the BAC, for example the β-*actin* gene/promoter. Although we have only explored the first possibility experimentally, it may be useful to discuss both options. We describe a method for the modification of a BAC containing the Rosa26 locus with an ectopic promoter using ΦC31-mediated cassette exchange, and the modification of a BAC containing the β-actin locus using homologous recombination in *E. coli* ("recombineering") to express the GOI under the endogenous β-actin promoter.

3.1. Generation of a Rosa26 BAC-Based Expression Vector Using ΦC31-Mediated Cassette Exchange

3.1.1. Generation of Rosa26Amp BAC DH10B Electrocompetent Cells and Electroporation of the pSC101-BAD-ΦC31-int Plasmid

1. Streak *E. coli* DH10B harbouring the Rosa26Amp BAC (15) (Fig. 1) on LB-agar plates containing 12.5 μg/mL chloramphenicol and 50 μg/mL ampicillin. Incubate overnight at 37°C (see Note 1).
2. Pick a single colony and grow it overnight at 37°C in 2 mL of LB medium containing 12.5 μg/mL chloramphenicol and 50 μg/mL ampicillin.
3. Transfer 1 mL of the culture into 50 mL of fresh LB media supplemented with 12.5 μg/mL chloramphenicol and 50 μg/mL ampicillin. Grow at 37°C until the optical density at 600 nm (OD_{600}) reaches 0.5.
4. Incubate the cells on ice 15 min.
5. Centrifuge the cell suspension for 10 min at 2,000 × *g* at 4°C

6. Discard the supernatant and resuspend the pellet in 50 mL of ice cold water.
7. Repeat steps 5 and 6 twice. Resuspend the last pellet in 1 mL of ice cold water (see Note 2).
8. Use 50 ng of the pSC101-BAD-ΦC31-int plasmid (15) to electroporate 100 μL of cells in a BioRad Gene Pulser (conditions: 2.5 kV, 25 μF and 200 Ω using 0.2 cm gap electroporation cuvettes).
9. After cell electroporation, add 1 mL of LB medium without antibiotics, incubate 90 min at 30°C and plate the cells on LB-agar plates containing 12.5 μg/mL chloramphenicol, 50 μg/mL ampicillin, and 5 μg/mL tetracycline (see Note 3).

3.1.2. Performing ΦC31-Mediated Cassette Exchange into the Rosa26^{Amp} BAC

1. Grow a colony harbouring the Rosa26^{Amp} BAC and the pSC101-BAD-ΦC31-int plasmid in 2 mL of LB media containing 12.5 μg/mL chloramphenicol, 50 μg/mL ampicillin and 5 μg/mL tetracycline overnight, at 30°C.
2. Transfer 1 mL of the overnight culture into 50 mL of LB medium containing 12.5 μg/mL chloramphenicol, 50 μg/mL ampicillin, and 5 μg/mL tetracycline. Grow the culture at 30°C until the OD_{600} reaches 0.2. Add L-arabinose at 0.2% final concentration and shift the culture at 37°C until the OD_{600} reaches 0.5. Prepare electrocompetent cells as described above (Subheading 3.1.1, steps 4–7).
3. Electroporate 100 ng of linearized incoming construct (Fig. 1) using the same conditions as in Subheading 3.1.1, step 8 (see Note 4).
4. Add 1 mL of LB medium to the electroporated cells (without antibiotics), incubate 90 min at 37°C, plate on LB agar plates containing 12.5 μg/mL chloramphenicol, 25 μg/mL kanamycin, and incubate overnight at 37°C (see Note 5).
5. Screen the colonies for cassette exchange performing replica plates. Pick 20–30 colonies and streak each one on an LB-agar plate containing 12.5 μg/mL chloramphenicol, 50 μg/mL ampicillin and on a replica LB-agar plate containing 12.5 μg/mL chloramphenicol, 25 μg/mL kanamycin. Clones that underwent cassette exchange and harbour the recombined Rosa26 BAC should be resistant to kanamycin and sensitive to ampicillin (see Note 6). A flow chart of ΦC31-mediated cassette exchange into the Rosa26^{Amp} BAC is described in Fig. 2.

3.1.3. Confirm the ΦC31-Mediated Cassette Exchange into the Rosa26 BAC at the Molecular Level

BAC DNA extraction (alkaline lysis mini-prep).

1. Grow 6–12 colonies (kanamycin resistant, ampicillin sensitive) in 2 mL LB medium containing 12.5 μg/mL chloramphenicol, 25 μg/mL kanamycin at 37°C overnight (see Note 7).
2. Transfer the culture into an Eppendorf tube; centrifuge 5 min at 1,000 × *g*.

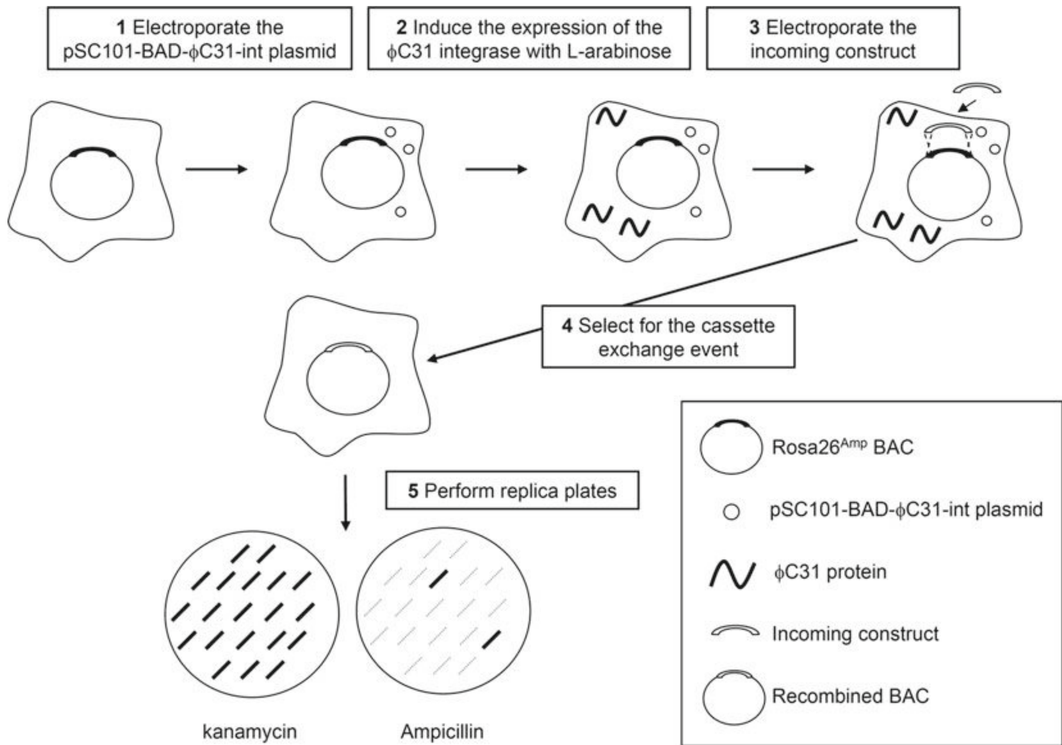


Fig. 2. Flow chart of the ΦC31-mediated cassette exchange in a BAC. Step 1: Electroporate the plasmid encoding the ΦC31 integrase into *E. coli* harbouring the Rosa26^{amp} BAC. Step 2: Induce the expression of the ΦC31 integrase with L-arabinose and make electrocompetent cells. Step 3: Electroporate the linearized incoming construct. Step 4: Select for colonies that underwent cassette exchange in chloramphenicol/kanamycin plates. Step 5: Confirm the cassette exchange event by performing replica plates. Bacteria, which have recombined the incoming construct into the BAC should be resistant to kanamycin and sensitive to ampicillin.

3. Remove the supernatant; resuspend the pellet by vortexing in 300 μL of P1 buffer.
4. Add 300 μL of P2 buffer, mix gently, and incubate at room temperature for 5 min.
5. Add 300 μL of ice-cold P3 buffer mix gently, incubate 5–10 min on ice; at this point, some debris should be visible.
6. Centrifuge at 13,500 × *g* for 15 min at 4°C.
7. Transfer the supernatant (approximately 700 μL) to a fresh Eppendorf tube.
8. Add 1 volume (700 μL) of ice cold isopropanol, mix gently, and incubate 15 min at −20°C.
9. Centrifuge at 13,500 × *g*, 15 min at 4°C.
10. Discard the supernatant and wash the pellet with 1 mL of 70% ethanol.
11. Discard the supernatant, let the pellet dry 5–10 min at room temperature, and resuspend in 20 μL of TE (see Note 8).

3.1.4. Confirm the Presence of the Incoming Construct in the BAC Using a PCR Strategy or Southern Blot

Design two primers hybridizing to the Rosa26 BAC and to the incoming construct as described in Fig. 3a (P1 and P2). Perform a PCR reaction using 1 μL of BAC DNA described above. The presence of a PCR product will indicate that the incoming construct is correctly recombined in the Rosa26 BAC. Alternatively, the correct recombination of the incoming construct in the Rosa26 BAC can be confirmed using a Southern blot approach as described in the Fig. 3b (see Notes 9 and 10).

3.2. Generation of a β-Actin BAC-Based Expression Vector Using Homologous Recombination in E. Coli

1. Transform the pSC101-BAD-gba-tet plasmid (19) by electroporation in cells harbouring the β-actin BAC as described in Subheading 3.1.1 (see Notes 11 and 12).

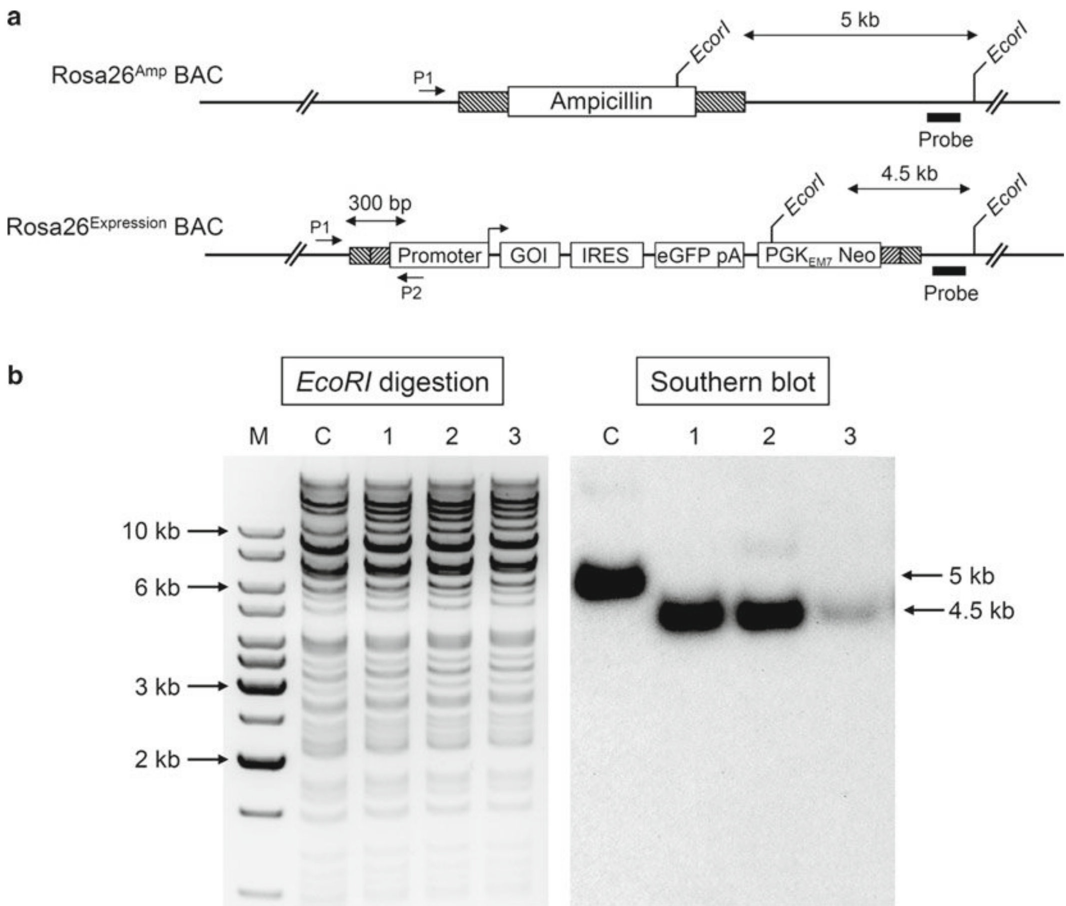


Fig. 3. Confirmation at the molecular level of the ΦC31-mediated cassette exchange into the Rosa26 BAC. (**a**) Strategy used to analyze the correct recombination events. A simple PCR analysis using two pairs of primers, one annealing to the Rosa26 BAC (P1) and a second one to the incoming construct (P2) will rapidly identify positive recombined clones. Only clones that underwent recombination of the incoming construct into the BAC should give a band. Alternatively, a Southern blot strategy can be designed. Perform a Southern blot with a probe located in the Rosa26 BAC (*black rectangle*). Digest the BAC DNA with an appropriate enzyme (in the example, *Eco*RI) that generate DNA fragments of different sizes in the case of the original BAC (Rosa26^{amp} BAC, 5 kb in the example) and the recombined BAC (Rosa26^{Expression} BAC, 4.5 kb, in the example). Fractionate the DNA in an agarose gel, blot it, and hybridize with a probe in the Rosa26 BAC. (**b**) Example of a Southern blot identifying a cassette exchange event. DNA from BACs was digested with *Eco*RI, fractionated in a 0.8% agarose gel, blotted, and hybridized with a probe in the Rosa26 BAC. The original BAC shows a corresponding band at 5 kb (c) while BACs that underwent cassette exchange show a band at 4.5 kb (1, 2 and 3). M: DNA molecular marker.

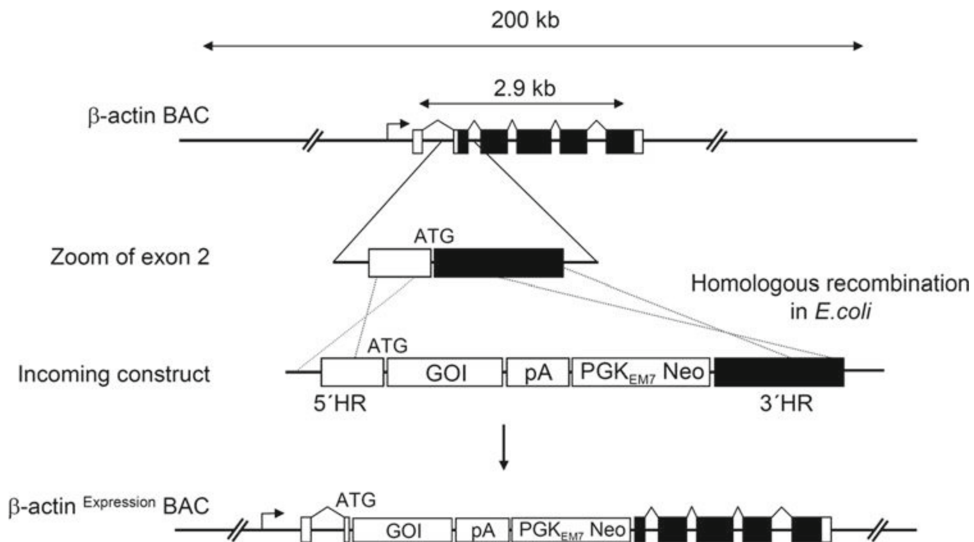


Fig. 4. Generation of a *β*-actin BAC-based expression vector using homologous recombination in *E. coli.* A BAC containing the β-actin locus is modified via homologous recombination in *E. coli* ("recombineering") in order to express the gene of interest under the *β-actin* endogenous promoter. An incoming construct containing 5′and 3′ 200 bp homologous regions (5′HR, 3′HR) to the exon 2 harbouring the initiation codon of the *β-actin* gene, a gene of interest (GOI), a polyadenylation signal (pA) and a PGK_{EM7} neo selection cassette is recombined into the *β*-actin locus. The 5′homology region design allows translation of the gene of interest from the endogenous initiation codon of the *β-actin* gene, thus allowing the control of transcription and translation of the GOI under the regulatory elements of the *β-actin* gene. *White rectangles* describe non-coding regions and *black rectangles* correspond to coding regions of the *β-actin* gene.

2. Grow a culture from a colony containing the β-actin BAC and the pSC101-BAD-gba-tet plasmid, induce the expression of the plasmid with arabinose and perform electrocompetent cells as in Subheading 3.1.2.
3. Electroporate 100 ng of the linearized incoming construct (Fig. 4, see Note 13).
4. After cell electroporation add 1 mL of LB medium without antibiotics, incubate 90 min at 37°C and plate the cells on LB-agar plates containing 12.5 μg/mL chloramphenicol, and 25 μg/mL kanamycin. Incubate the plates at 37°C overnight. Check the correct homologous recombination event by PCR or Southern blot as described in Subheading 3.1.3.

3.3. Generation of an HEK-293 Stable Culture Using BAC-Based Vectors

1. Prepare BAC DNA using NucleoBond BAC 100 Kit (Clontech) according to manufactures instructions (see Note 14).
2. For the transfection of a 10 cm Petri dish digest 24 μg of BAC DNA with *Not*I, use 100 U in a final volume of 300 μL, digest 6 h or overnight (see Note 15).
3. Precipitate the DNA by adding 30 μL of 3 M sodium acetate, pH 5.2 and 750 μL of 96% ethanol. Mix gently. At this point, the DNA pellet should be visible. Let the pellet sink to the bottom of the tube, remove the supernatant, wash the pellet

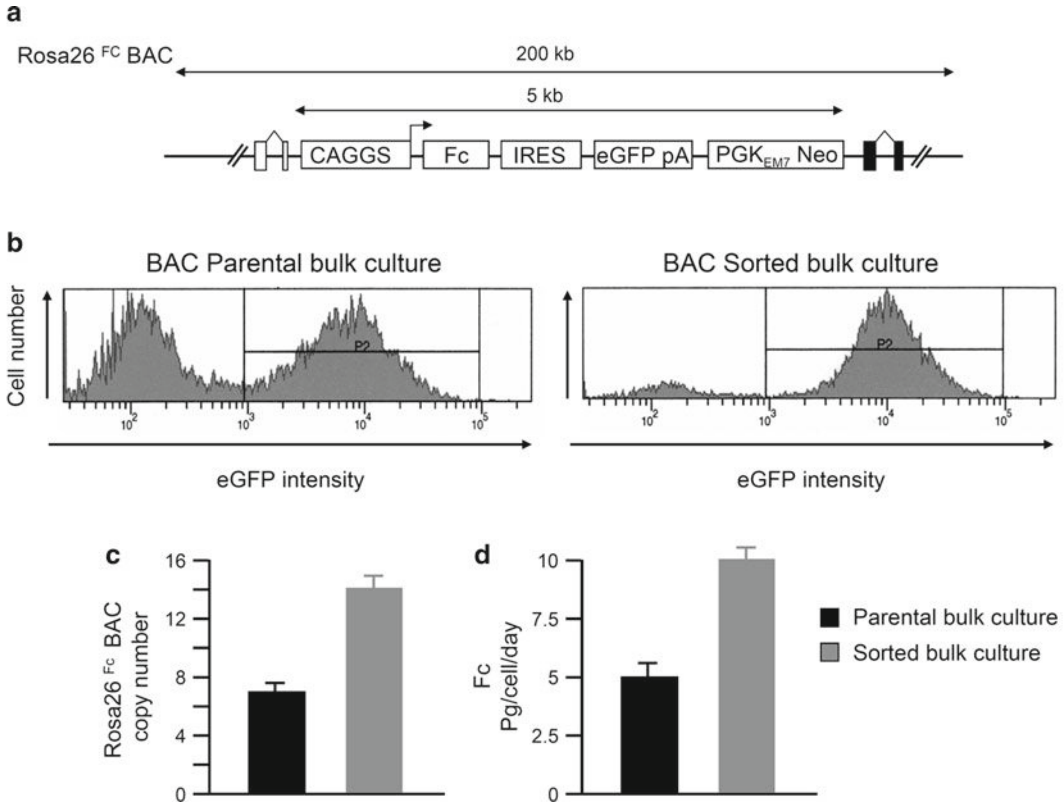


Fig. 5. Example of a stable culture generated from a BAC-based vector. (**a**) HEK 293-cells were transfected with a Rosa26 BAC expressing the Fc fragment of the constant region of human IgG1 as gene of interest from the CAGGS promoter (Rosa26 ^{Fc} BAC). (**b**) After neomycin selection, we established two bulk cultures: Parental culture (initial bulk culture without any further manipulation) and sorted culture (eGFP FACS sorted culture enriched twice for eGFP expression). (**c**) Copy number analysis of the Rosa26 ^{Fc} BAC. We measured the number of BAC integrated copies in both the cultures, parental and sorted (enriched twice according to the eGFP expression). The average number of Rosa26 ^{Fc} BAC copies was 7 in the parental versus 14 in the sorted culture, thus suggesting that an increase in the BAC copy number correlates with the increase in eGFP expression and more importantly, with the levels of expression of Fc: 5 pcd (picogram/cell/day) in the parental culture versus 10 pcd in the sorted culture (**d**).

with 1 mL of 70% ethanol, remove the supernatant, let it dry at room temperature and resuspend with 50 μL of TE buffer (see Note 16).

4. Transfect a 10 cm Petri dish containing 80% confluence HEK-293 cells with 50 μL (24 μg) of linearized BAC DNA with Lipofectamine 2000, according to the manufacturer's instructions (see Note 17).
5. Two days later, start the antibiotic selection by adding to the medium neomycin (800 μg/mL, final concentration); 1 or 2 weeks later neomycin-resistant colonies should appear. Expand the culture and analyze the expression of your GOI as illustrated in Fig. 5 (see Note 18).

4. Notes

1. The Rosa26^{Amp} BAC has previously been prepared for ΦC31-mediated cassette exchange. It is resistant to chloramphenicol and contains an ampicillin cassette flanked by two attP sites.
2. Electrocompetent cells can be stored in aliquots at −80°C. For this purpose, use ice cold glycerol 10% (v/v) in the last wash instead of water, resuspend the pellet in 10% glycerol, make aliquots, and snap-freeze in liquid N_2.
3. The pSC101-BAD-ΦC31-int is a tetracycline resistance plasmid that expresses the ΦC31 integrase upon induction with L-arabinose. It contains an origin of replication, which is temperature sensitive and should be grown at 30°C.
4. The incoming construct in this example consists of an ectopic promoter (CAGGS, Sv40, etc.), a gene of interest, an internal ribosome entry sequence, an eGFP as fluorescent marker and a PGK/EM7 neomycin cassette conferring resistance to kanamycin from the EM7 promoter in prokaryotes and to neomycin (G418) from the PGK promoter in eukaryotes. These elements are flanked by two attB. In the presence of ΦC31 integrase, the attB sites from the incoming construct will recombine with the attP sites present in the Rosa26^{amp} BAC and the incoming construct will be integrated into the BAC. Finally, our incoming construct contains a unique *Not*I recognition site that allows the linearization of the plasmid.
5. The pSC101-BAD-ΦC31-int plasmid does not replicate at 37°C, and it will be lost at this step in the absence of tetracycline. Cells, which underwent ΦC31-mediated cassette exchange, lose ampicillin resistance (originally present in the BAC) and become kanamycin resistant (present in the incoming construct).
6. Routinely, we get 80–90% kanamycin resistant/ampicillin-sensitive colonies; however, in some experiments we have found a significant number of colonies resistant to both antibiotics, this is mainly due to incomplete linearization of the incoming construct.
7. If needed, volumes can be scale up.
8. The alkaline lyses DNA mini-prep is a rather crude method that yields BAC DNA with enough purity for PCR analysis and enzymatic restriction analysis. For sequencing BAC DNA, we suggest the use of NucleoBond BAC 100 Kit (Clontech).
9. Initially, we always confirmed the correct insertion of the incoming construct using both PCR and Southern blot approaches. However, we observed that all the colonies resistant to kanamycin and sensitive to ampicillin were correctly recombined, thus we now only perform a PCR analysis.

10. Positive colonies can be stored at −80°C as glycerol stocks.
11. For a detailed description of finding and purchasing the BACs, see ref. 18. In general, we tend to choose BACs in which the GOI is located in the middle, harbouring enough upstream and downstream sequence to ensure the presence of most regulatory elements of the locus. In case of the β-*actin* gene, the gene is rather compact (2.9 kb, Fig. 4) and for example the BAC RP23-9701 might be a good candidate.
12. The pSC101-BAD-gba-tet is a plasmid containing an origin of replication, which is temperature sensitive (grows at 30°C but not a 37°C), and allows to perform homologous recombination in *E. coli* upon addition of L-arabinose and is tetracycline resistant. It should be handled as mentioned for the pSC101-BAD-ΦC31-int. Antibiotic selection is done in the presence of chloramphenicol and tetracycline (but not ampicillin).
13. The incoming construct contains 200 bp upstream and downstream homologous regions to the target location into the BAC (Fig. 4). As 5′homology region, we usually use 200 bp upstream of the initiation codon fused to the gene of interest. As 3′homology region, we use 200 bp downstream of the initiation codon. The aim is that the GOI is translated from the endogenous initiation codon present in the BAC, in order to remain as close as possible to the endogenous configuration. The minimal incoming construct contains the gene of interest, a polyadenylation signal and a resistance cassette that allows double selection in eukaryotes and prokaryotes, for example the PGK_{EM7} neomycin (neomycin in eukaryotes, kanamycin in prokaryotes), other features can be also included (IRES-eGFP, etc.). A more detailed protocol describing the modification of a BAC via homologous recombination can be found in (20).
14. In our experience, the NucleoBond BAC 100 Kit gives satisfactory yields and high-quality BAC DNA. From a 500 mL overnight culture, 50–100 μg of BAC DNA are expected.
15. In order to generate stable cell lines, it is convenient to linearize the BAC; linearization increases the number of stably transfected cells. The BAC vector backbone contains two *Not*I recognition sites flanking the insert, thus *Not*I digestion releases the insert. Alternatively, if *Not*I recognition sites are present in the insert, the BAC can be linearized with *PI-Sce*I, a homing endonuclease derived from yeast, whose recognition sequence (about 27 bp) is present in the BAC backbone.
16. Linearized BACs are fragile, treat them gently and do not freeze them, store at 4°C. Do not let the BAC DNA pellet dry too long; otherwise, it will be difficult to dissolve it.
17. Linear BAC DNA tends to be viscous. In order to mix it with the medium, it might be necessary to pipette up and down. Do it gently with a cut off 1 mL tip.

18. The use of BAC is not restricted to HEK cells; other cell lines like CHO can also be used. Other transfection methods, such as electroporation can also be used. From the transfection of a 10 cm Petri dish, we typically obtain 200–300 antibiotic resistant colonies.

References

1. Rita Costa, A., Elisa Rodrigues, M., Henriques, M., et al., (2010) Guidelines to cell engineering for monoclonal antibody production, *Eu.r J. Pharm. Biopharm.* **74**, 127–138.
2. Wurm, F. M. (2004) Production of recombinant protein therapeutics in cultivated mammalian cells, *Nature Biotechnol.* **22**, 1393–1398.
3. Giraldo, P., and Montoliu, L. (2001) Size matters: use of YACs, BACs and PACs in transgenic animals, *Transgenic Res.* **10**, 83–103.
4. Pikaart, M. J., Recillas-Targa, F., and Felsenfeld, G. (1998) Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators, *Genes Develop.* **12**, 2852–2862.
5. Girod, P. A., Nguyen, D. Q., Calabrese, D., et al., (2007) Genome-wide prediction of matrix attachment regions that increase gene expression in mammalian cells, *Nature Methods* **4,** 747–753.
6. Antoniou, M., Harland, L., Mustoe, T., et al., (2003) Transgenes encompassing dual-promoter CpG islands from the human TBP and HNRPA2B1 loci are resistant to heterochromatin-mediated silencing, *Genomics* **82**, 269–279.
7. Kwaks, T. H., Barnett, P., Hemrika, W., et al., (2003) Identification of anti-repressor elements that confer high and stable protein production in mammalian cells, *Nature Biotechnol.* **21,** 553–558.
8. Huang, Y., Li, Y., Wang, Y. G., et al., (2007) An efficient and targeted gene integration system for high-level antibody expression, *J. Immuno.l Methods* **322**, 28–39.
9. Blaas, L., Musteanu, M., Eferl, R., et al., (2009) Bacterial artificial chromosomes improve recombinant protein production in mammalian cells, *BMC Biotechnol.* **9,** 3.
10. Bian, Q., and Belmont, A. S. (2010) BAC TG-EMBED: one-step method for high-level, copy-number-dependent, position-independent transgene expression, *Nucleic Acids Research* **38**, e127.
11. Zambrowicz, B. P., Imamoto, A., Fiering, S., Herzenberg, L. A., et al., (1997) Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells, *Proc. Natl. Acad. Sci. USA* **94,** 3789–3794.
12. Zagranichnaya, T. K., Wu, X., Danos, A. M., et al., (2005) Gene expression profiles in HEK-293 cells with low or high store-operated calcium entry: can regulatory as well as regulated genes be identified?, *Physiological Genomics* **21**, 14–33.
13. Yang, X. W., Model, P., and Heintz, N. (1997) Homologous recombination based modification in Escherichia coli and germline transmission in transgenic mice of a bacterial artificial chromosome, *Nature Biotechnol.* **15**, 859–865.
14. Zhang, Y., Buchholz, F., Muyrers, J. P., et al., (1998) A new logic for DNA engineering using recombination in Escherichia coli, *Nature Genetics* **20**, 123–128.
15. Blaas, L., Musteanu, M., Zenz, R., et al., (2007) PhiC31-mediated cassette exchange into a bacterial artificial chromosome, *BioTechniques* **43**, 659–660, 662, 664.
16. Gong, S., Zheng, C., Doughty, M. L., et al., (2003) A gene expression atlas of the central nervous system based on bacterial artificial chromosomes, *Nature* **425,** 917–925.
17. Poser, I., Sarov, M., Hutchins, J. R., et al., (2008) BAC TransgeneOmics: a high-throughput method for exploration of protein function in mammals, *Nature Methods* **5,** 409–415.
18. Johansson, T., Broll, I., Frenz, T., et al., (2010) Building a zoo of mice for genetic analyses: a comprehensive protocol for the rapid generation of BAC transgenic mice, *Genesis* **48,** 264–280.
19. Zhang, Y., Muyrers, J. P., Rientjes, J., et al., (2003) Phage annealing proteins promote oligonucleotide-directed mutagenesis in *Escherichia coli* and mouse ES cells, *BMC Molec. Biol.* **4,** 1.
20. Testa, G., Vintersten, K., Zhang, Y., et al., (2004) BAC engineering for the generation of ES cell-targeting constructs and mouse transgenes, *Methods Molec. Biol.* **256**, 123–139.

Chapter 32

Engineering the Chaperone Network of CHO Cells for Optimal Recombinant Protein Production and Authenticity

Lyne Jossé, C. Mark Smales, and Mick F. Tuite

Abstract

All proteins fold into a defined three-dimensional shape that is compatible with the cellular role and/or biological activity of those proteins. Molecular chaperones are a family of proteins whose role is to assist the folding and targeting of proteins in both normal and stressed cells. The rational manipulation of chaperone levels in a cell line engineered to produce a defined recombinant protein (rP) can significantly improve both the achievable steady-state levels and authenticity of a wide range of recombinant proteins. Here, we describe the methodology associated with expressing a variety of molecular chaperones in Chinese hamster ovary (CHO) lines in order to improve their recombinant protein production capacity. These chaperones include both those that facilitate the folding of the polypeptide chain (i.e. Hsp70, Hsp40) and those that can re-fold proteins that have misfolded in the cell (i.e. ClpB/Hsp104). This latter property is particularly important given the propensity for highly expressed recombinant proteins to misfold in the "foreign" cellular environment.

Key words: Chinese hamster ovary cells, Molecular chaperone, Protein folding, Luciferase, Protein aggregation, Hsp104, Hsp70, Hsp40, Recombinant antibody

1. Introduction

The last few years have seen rapid progress being made in the field of recombinant protein (rP) production in a wide range of cell types, including cultured mammalian cells. This progress largely reflects the significant advances made in cell line optimisation and transfection methodologies. Yet a number of challenges remain before we can be satisfied that we have the best-possible expression options for a wide range of recombinant proteins. Foremost among these challenges is protein aggregation. The accumulation of aggregation-prone misfolded proteins in cells can be a consequence

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_32, © Springer Science+Business Media, LLC 2012

of exposure of the cells to physical or chemical stress, but also can be due to engineered overexpression of recombinant proteins that can compromise cellular fitness (1).

Insoluble protein aggregates are linked to human pathologies, such as Alzheimer's, Parkinson's, and Creutzfeldt Jakob's disease, cataracts (2), and cystis fibrosis, but are also a significant problem for the biotechnology industry because aggregation can influence properties, such as protein stability or immunogenicity (3). To safeguard protein homeostasis, cells have evolved a quality control network to ensure that newly synthesised proteins are folded into their correct three-dimensional shape. As part of this control, proteins unable to reach or retain their active conformation are routed to a degradation pathway or sequestered in discrete cellular structures, such as aggresomes or Russell bodies (4, 5).

While aggregated recombinant proteins can be refolded in vitro, this is costly and not an economical solution for large-scale protein production. Hence, we need to seek solutions that prevent protein aggregation in the cell. One such solution is to rationally manipulate the cellular levels of certain molecular chaperones. Molecular chaperones are a diverse group of proteins central to the quality control network of both prokaryotic and eukaryotic cells. They assist the folding of a wide range of "client proteins", but also more generally facilitate the re-folding of stress-denatured proteins and thus can prevent the accumulation of protein aggregates. Consequently, chaperones have become the focus of research strategies designed to combat aggregation disorders, but more recently in biotechnology as a strategy to increase recombinant protein production and authenticity in prokaryotic and eukaryotic cells.

Molecular chaperones are highly conserved at the amino acid sequence level through evolution and include the heat-shock proteins (Hsps) and the oxidoreductase protein families (Table 1). Hsp chaperones are categorised according to their molecular weight with the Hsp70 and Hsp90 families being the most abundant. They also include two families of high-molecular-weight proteins: the Hsp105/110 family members – a diverged subgroup of the Hsp70 family – and the Hsp104/ClpB multimeric chaperones, such as ClpB (in bacteria, such as *Escherichia coli*), Hsp104 (in fungi, such as the yeast *Saccharomyces cerevisiae*), and Hsp101 (in plants, such as *Arabidopsis thaliana*) (6–9). Members of the Hsp70, Hsp90, and Hsp104/ClpB families typically work synergistically with accessory factors called co-chaperones which largely belong to the Hsp40 and small Hsp (sHsp) families. Other important chaperones include the protein disulphide isomerases (PDIs) that can catalyse thiol-disulphide oxidation, reduction, and isomerisation in addition to a basic chaperone function, and the lectin chaperones calnexin and calreticulin which bind monoglucosylated N-glycans in the endoplasmic reticulum (ER), but also work in cooperation with oxidoreductases, such as Erp57 (10).

Table 1
The major classes of eukaryotic molecular chaperones and their cellular functions

Molecular chaperones	Function	References
Heat-shock proteins		
Hsp110/SSE family (Hsp105 α,β; Sse1-2)	Hsp105 interacts with BIP; nucleotide exchange factor for Ssa1 and Ssb1	(15, 16)
Hsp104/ClpB (Hsp104, ClpB)	Stress response, thermotolerance	(8)
Hsp90	Protein folding, cell signalling, tumour repression	(17)
Hsp70 (e.g. Ssa1p, Kar2p, BIP/Grp78)	mRNA translation, translocation, folding, degradation	(18, 19)
Hsp60 (CCT)	*De novo* folding, regulation of cytoskeletal, and cell cycle network	(19–21)
Hsp40 (Ydj1, Hdj1)	Co-chaperones of Hsp70s	(22)
Small Hsps (e.g. Hsp26, Hsp27, Hsp42)	Stabilise unfolded proteins in an ATP-independent manner and prevent their aggregation until the proteins are refolded by ATP-dependent chaperones	(23, 24)
Grp170/Lhs1	Protein import in ER, nucleotide exchange factor	(25, 26)
Oxidoreductases		
PDI, PDIp, ERp44, ERp57, ERp72	Catalyse disulphide bond formation, breakdown, and isomerisation	(27)
Lectin chaperones		
Calnexin, calreticulin	Calcium-binding proteins, glycoprotein maturation in the ER	(28)

The ever-growing demand for large-scale production of therapeutic recombinant proteins for the pharmaceutical industry has prompted further studies on protein folding and in particular on the optimisation of the chaperone network. These manipulations are in addition to optimisation of culture techniques, engineering high-efficiency expression vectors, screening for cell lines with enhanced production of target proteins, and improvement of transfection procedures. The need for co-expression of "modulating" factors, such as molecular chaperones, must now also be added to list.

Recent advances in recombinant protein production have led to improved efficiency of transcription and translation and increased solubility of recombinant proteins. Among eukaryotic host expression systems, Chinese hamster ovary (CHO) cells are generally the preferred option being widely used for transfection, expression, and large-scale recombinant protein production in a wide variety of contexts. Here, we provide an overview of the technology required to engineer the transient expression of both homologous and heterologous molecular chaperones as part of an overall strategy for increasing the productivity and authenticity of recombinant proteins in CHO cell lines.

2. Materials

2.1. Cell Cultures

1. 96-well cell culture plate, sterile, with lid (Greiner Bio-One CELLSTAR).
2. CHO cell lines: CHOK1, an adherent cell line obtained from Lonza Biologics. CHOD6, an adherent cell line with a stably integrated cDNA expressing firefly luciferase from the CMV promoter. LB01, a suspension cell line engineered to stably express IgG4. This is a high-yielding monoclonal antibody cell line derived from the CHOK1 variant, CHOK1SV, developed by Lonza Biologics, and all CHOK1SV clones carry the glutamine synthase system (11).
3. DMEM/F12 (1:1) nutrient mix (Invitrogen) supplemented with the following: 100 μM MEM non-essential amino acid solution (Invitrogen), 400 μM L-glutamic acid and L-asparagine, 20 μM adenosine, guanosine, cytidine and uridine, 10 μM thymidine, 6 mM L-glutamine, and 10% (v/v) heat inactivated foetal calf serum (Lonza Biologics).
4. CD CHO medium (Invitrogen).
5. Methionine sulphoximine (MSX) (Sigma–Aldrich).

2.2. Transfection of Chaperones in CHO Cell Lines

1. Chaperones cDNAs were cloned into the mammalian expression vector pcDNA4/myc-His A in the absence of the hexa-histidine tag (Invitrogen).
2. pCDNA3/enhanced green fluorescent protein (EGFP) (a gift from Dr Matt Smith, University of Kent, UK) is used as a transfection control.
3. Lipofectamine 2000 (Invitrogen).
4. 1× Opti-MEM I reduced serum medium with GlutaMAX I (Invitrogen).
5. FuGENE HD transfection reagent (Roche).
6. 1× CD-CHO medium (Invitrogen).

2.3. Preparation of Cell Extracts for Analysis

1. 10× PBS: 80 g NaCl, 2 g KCl, 14.4 g Na_2HPO_4, and 2.4 g KH_2PO_4; pH 7.4.
2. 1× Trypsin/EDTA solution (Invitrogen).
3. Reporter Lysis Buffer (RLB) (Promega).
4. Soniprep 150 ultrasonic disintegrator (MSE).

2.4. SDS-PAGE and Western Blotting Analysis of Chaperones

1. 5× SDS gel-loading buffer: 250 mM Tris–HCl, pH 6.8, 500 mM DTT, 10% SDS, 0.5% bromophenol blue, and 50% glycerol.
2. 5× Tris–glycine running buffer: 25 mM Tris, 250 mM glycine, and 0.1% SDS.
3. 4–20% Tris–glycine gel (Invitogen).

4. Gel running apparatus: XCell Sure Lock gel system (Invitrogen).
5. Nitrocellulose membrane (Immobilon-P^{SQ}) (Millipore).
6. Extra thick filter paper (Bio-Rad).
7. 1× transfer buffer: 48 mM Tris base, 39 mM glycine, 0.037% SDS, and 20% methanol; pH 8.3.
8. Gel transfer apparatus: Trans Blot SD, Semi dry transfer cell (BioRad).
9. Primary antibodies used: Anti-TorsinA rabbit polyclonal antibodies were a gift from Dr Lisa Swanton (University of Manchester, UK). Anti-Hsp104 polyclonal rabbit antibodies (SPA-400). Anti-Hsp70 polyclonal rabbit antibodies (SPA-812). Anti-Hsp40 polyclonal rabbit antibodies (SPA-10400). Anti-Hsp27 polyclonal rabbit antibodies (SPA-803). Anti-firefly luciferase polyclonal rabbit antibodies (L0159, Sigma). Anti-*Gaussia* luciferase polyclonal rabbit antibodies (E8023S, New England Biolabs).
10. Secondary antibodies: For ECL detection, the secondary antibodies were either horseradish peroxidase (HRP)-conjugated anti-rabbit (DAKO) or anti-mouse (DAKO). For immunofluorescence, the secondary antibodies used for fluorescence reading were either fluorescein isothiocyanate (FITC)-conjugated anti-rabbit (Sigma–Aldrich) or anti-mouse (Sigma–Aldrich) antibodies.
11. ECL detection reagents. Prepare mix 1 and 2. For mix 1 (10 mL): 100 μL luminol (250 mM in DMSO), 44 μL coumaric acid (90 mM in DMSO), 1 mL 1 M Tris–HCl, pH 8.5, and 8.856 mL H_2O. For mix 2 (10 mL): 6.4 μL H_2O_2, 1 mL 1 M Tris–HCl, pH 8.5, and 8.894 mL H_2O.
12. X-ray film (Amersham Hyperfilm ECL, GE Healthcare).

2.5. Luficerase Reporter Activity Assay

1. pGL3 vector (Promega, UK) was used for transient expression of firefly luciferase.
2. pCMV-GLUC-1 vector (Nanolight Technology) was used for transient expression of *Gaussia* luciferase.
3. For measurement of the firefly luciferase activity, we use the Steadylite Plus™ reagent (Perkin Elmer).
4. For measurement of the *Gaussia* luciferase activity, we use the *Gaussia* Luciferase Assay Kit (New England Biolabs).
5. White opaque 96-well assay plates (Perkin Elmer).
6. Wallac 1450 MicroBeta Trilux liquid scintillation counter (Perkin Elmer).

2.6. Quantifying Protein Levels

1. An anti-EGFP mouse monoclonal antibody (clone EGFP-3E1) was a gift from Professor Bill Gullick (University of Kent, UK).
2. Anti-β actin mouse monoclonal antibodies (clone AC-15) (Sigma–Aldrich).

3. Methods

Manipulation of the expression of the various and diverse classes of molecular chaperones has proved to be challenging mainly because the product titres can vary so greatly depending on the chaperone, host cell system, and targeted recombinant product (Table 2). A more recent approach is the co-expression of a panel of cooperative chaperones rather than individual chaperones (12). Aside from the choice of chaperone, a key consideration must be the choice of transfection and expression strategies. The type of cell line, scale of the transfection, and stability of the resulting transgenic line are also parameters that need to be considered prior to transfection. In addition, the composition of the DNA:reagent complex solution can greatly influence the transfection efficiency of CHO variants, be they adherent or suspension adapted. Commonly used transfection reagents rely on the addition of $CaPO_4$ complexes, polycations polyethylenimines (PEIs), and liposomes.

Table 2
Examples of the reported effect of overexpressing molecular chaperones on recombinant protein (rP) titre in CHO cells

Engineered molecular chaperone	rP/product titre[a]	References
Hsp27	Firefly luciferase/+	(12)
Hsp27	Human interferon-γ/+	(29)
Hsp70	Human interferon-γ/+	(29)
BIP	Antibody/–	(30)
Grp170	Antibody/–	L Jossé and CMSmales, unpublished data
Hsp104	Firefly luciferase/+	(12)
ERp57	Thrombopoietin/+	(31)
PDI	Antibody/+	(30)
PDI/BIP	Antibody/–	(30)
PDI/Ero1	Antibody/variable	(32)
Calnexin/calreticulin	Thrombopoietin/+	(33)
BIP/ERdj3	Antibody/–	Feary M and CM Smales, unpublished data
TorsinA	Antibody/+	(12)

[a]Product titre is indicated by +, increased levels: 0, no effect: –, reduced levels

Combining osmotic shock with polycation treatment and glucose or NaCl with PEIs can increase the uptake of DNA by the cell. PEIs work independently of serum, and for transient transfection and short-term recovery of CHO cell lines they result in greater DNA delivery compared to the polycation-based method (13). PEIs indirectly protect the DNA that is taken up by the cell, from degradation, by destabilising the endosome (14). Interestingly, while PEIs seem to deliver DNA to more cells than $CaPO_4$, the expression level per cell appears to be greater with the latter (13). DNA integration events occur more frequently in the presence of $CaPO_4$ and the specific productivity of the resulting recombinant cell lines is higher (13). However, whether the reagent is sourced commercially or is available as a laboratory stock reagent, the scale of the transfection reaction and the cell type usually determine the transfection method. It is, therefore, important to carry out one or more pilot experiments to determine the most suitable transfection method to use for a given application and cell line. Table 3 shows different transfection protocols based on our own findings and those of others.

Below, we describe in more detail the small-scale transfection procedure we use for CHO cells for the transient expression of cDNAs encoding molecular chaperones. The effect of heterologous chaperone expression is monitored by measuring the luminescent activity of cytoplasmic firefly or secreted *Gaussia* luciferase reporter proteins. The luciferase activity usually accurately reflects the overall cell transfection efficiency. Normalisation of the luminescence signal is described in the data handling (Subheading 3.6) and an illustration of typical Western blot analyses of chaperone expression in CHO-K1 cells is presented in Fig. 1.

3.1. Cell Cultures

1. CHOD6 and CHOK1 cell lines are routinely cultured in DMEM/F12 (1:1)-based nutrient mix.
2. LB01 cells are cultured in CD-CHO medium. MSX is added at a final concentration of 50 mM.
3. Cells are maintained as adherent (CHOK1) or suspension (LB01) cultures at 37°C in a 5% CO_2-humidified atmosphere.

3.2. Transfection of Chaperones in CHO Cell Lines

1. The CHOD6 and CHOK1 cells are seeded in 2 mL supplemented DMEM/F12 at 2.5×10^5 cells/mL (or 0.5×10^6 cells) 24 h prior to transfection. The growth medium is replaced by fresh medium prior to transfection.
2. To each 2 mL well, a standard transfection mixture typically containing 1 μg of DNA of each vector together with 10 μL of Lipofectamine 2000 in 500 μL of serum-reduced Opti-MEM I is added. Luciferase activity is measured 24 h post transfection using the assay described in Subheading 3.4.

Table 3
Small- and large-scale transfection procedure for adherent and suspension cell lines

Cell lines/format/density[a]	Transfection medium	Target protein	Transfection reagent/method	Manufacturer/ references
CHO K1/adherent/transient 6-well plates 0.5×10^6 cells/mL	DMEM/F12 based	GFP Luciferase	Lipofectamine 2000	Invitrogen (12)
CHO K1/suspension/transient 6-well plates 1×10^6 cells/mL	CD-CHO (Invitrogen)	GFP IgL, IgM	Fugene HD	Roche (12)
CHO-DG4/suspension/transient 12-well plates 2×10^6 cells/mL	DMEM/F12 based RPMI 1640 based ProCHOD5 CDM[b] (Cambrex)	GFP IgL/IgM	1. Lipofectamine 2000 2. PEI[c,d] linear 25 kDa. DNA:PEI ratio 1:2 to 1:4 (DNA 2.5 μg/mL). 150 mM NaCl	Invitrogen (20–23)
CHO-DG44/suspension/transient BR 3, 5, 20,150 L 2×10^6 cells/mL	DMEM/F12 based RPMI 1640 based ProCHOD5 CDM[b] (Cambrex)	IgL/IgM	PEI[c, d] linear 25 kDa. DNA:PEI ratio 1:2 to 1:4 (DNA 2.5 μg/mL). 150 mM NaCl	(20–23)
CHO-DG44/adherent/transient and integrated 12-well plates 3×10^5 cells/mL	DMEM/F12 based	GFP IgG4 (linearised DNA)	$CaPO_4$	(13)
CHO-DG44/suspension/semi-stable 60-mm Petri dishes 2×10^4 cells/mL	DMEM/F12 based	GFP	Microinjection	(24)[e]

PEI polyethylenimine, *BR* bioreactor
[a]Density on day of transfection
[b]ProCHOD5 CDM performed better in Muller et al. study
[c]PEI has been reported to be cytotoxic to endothelial cells and L929 fibroblasts (15, 16)
[d]The use of PEI is serum independent (17)
[e]22 days post injection, only 3–4% of cells were recombinant

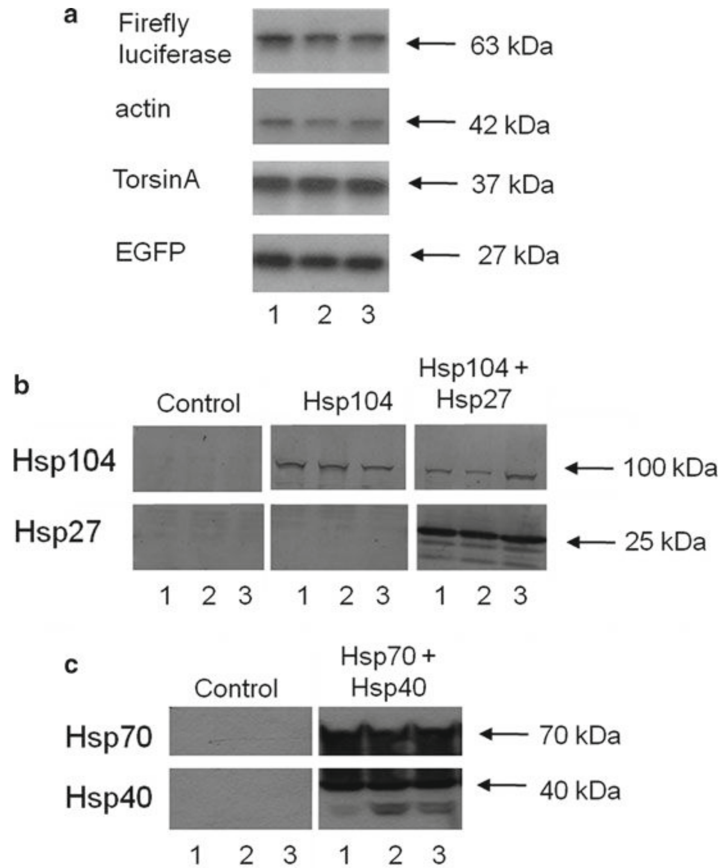


Fig. 1. Detection of various molecular chaperones and protein markers in CHO cells by Western blot analysis. (**a**) CHOD6, with integrated firefly luciferase, co-transformed with torsinA (chaperone), and enhanced green fluorescent protein (EGFP, transfection control) were probed first with mouse anti-β-actin and anti-EGFP mouse monoclonal antibodies. The membrane was then re-probed with anti-firefly luciferase and anti-torsinA rabbit polyclonal antibodies. *Lanes 1–3*: Biological replicates. (**b**, **c**) Detection of the indicated Hsps following transient transfection. Immunoblots show pairwise detection of chaperones using the relevant rabbit polyclonal antibodies. *Lanes 1–3*: Biological replicates.

3. LB01 cells are seeded in 2 mL CD-CHO at 2×10^6 cells/mL (or 4×10^6 cells) on the day of transfection. FuGENE HD Transfection Reagent:DNA ratios of 3 μL to 1 μg of DNA are diluted in 100 μL Opti-MEM I and added to each 2-mL well and samples processed within 24 h post transfection.
4. Successful transfection can be readily verified by assessing co-transfection of the marker EGFP. This is done using fluorescence microscopy by looking in each well at the percentage of cells that display green fluorescence. Quantitation of the level of GFP is further refined using Western blot analysis (described in Subheading 3.6).

3.3. Preparation of Cell Extracts for Analysis

1. CHOD6 or CHOK1 adherent cells are detached as follows. Each well is washed once with 1× PBS, and then 200 μL of trypsin/EDTA is added and left for 2–3 min at 37°C. The trypsin can then be inactivated by the addition of 1 mL DMEM/F12-based medium.
2. Detached or suspension cells (LB01) should be transferred to a 1.5-mL Eppendorf tube, the cells harvested by centrifugation at $1.06 \times g$ for 1 min at 4°C, and then washed once in 1× PBS.
3. Cell lysates can be prepared using RLB. This buffer contains detergent, reducing and chelating agents, and is compatible with the measurement of both firefly and *Gaussia* luciferase activity (see Note 1). One Complete Protease Inhibitor Cocktail mini-tablet should be added per 5 mL RLB. To shear DNA contaminants and reduce viscosity, cells are typically sonicated twice for 10-s bursts at full power (see Note 2).

3.4. SDS-PAGE and Western Blotting Analysis of Chaperones

Expressed chaperones or luciferases can be readily detected using standard SDS-PAGE in conjunction with Western blot analysis, as follows.

1. Approximately 8×10^5 CHO cells should be harvested (~200 μg of total protein) and resuspended in 50 μL of lysis buffer. 10 μL of 5× SDS gel-loading buffer is then added and the samples can be frozen at −20°C at this stage. We suggest using 6 μL (~20 μg of total protein), where one wants to detect high-abundance proteins, and 15 μL (~50 μg of total protein), where one is looking for low-abundance proteins. Protein samples are separated by SDS-PAGE on a 4–20% Tris–glycine gradient gel in 1× running buffer for 90 min at 150 V.
2. Following electrophoresis, proteins should be transferred to an Immobilon-P^{SQ} (Millipore) membrane using transfer buffer and then blocked for 15 min with 5% milk in 1× PBS. The blocked filter is then incubated overnight (12–16 h) with the required primary antibody, followed by five sequential washes in 1× PBS (see Note 3). The membrane is then incubated with a suitable secondary antibody for 1 h and then washed five times in 1× PBS.
3. Chaperone expression is assessed using either enhanced luminol-based chemiluminescent (ECL) or immunofluorescence (see Note 4).
4. For ECL analysis, 2 mL of mix 1 and 2 are combined just before and added to the membrane just before exposing the membrane to X-ray film. For immunofluorescence, the signal can be detected directly by a fluorescence scanner, such as the FujiFilm FLA-5100 plus.

5. To confirm that chaperones are expressed post transfection, the only option is to use Western blot analysis (see Note 5), whereas for luciferase protein levels these can be measured either by Western blot or by an activity assay (see Subheading 3.5).

3.5. Luciferase Reporter Activity Assay

Luciferase activity should be determined in biological triplicate using white, opaque, 96-well assay plates in the following protocol. For measurement of the firefly luciferase activity, the Steadylite Plus™ reagent should be used according to the manufacturer's instructions. 50 μL of the Steadylite Plus™ reagent is added to the culture supernatant obtained from the equivalent of 5×10^4 to 1×10^5 cells.

1. *Gaussia* luciferase activity is measured using the *Gaussia* Luciferase Assay Kit, but with the following modification according to the manufacturer's protocol: The reconstituted *Gaussia* Luciferase solution (1×) is further diluted 1:10 v/v with 1× PBS and 25 μL of this added to the cell culture supernatant (equivalent to 2,500 cells).
2. For both luciferase assays, the luminescence is best measured after a 10-min equilibration period using a Wallac 1450 MicroBeta Trilux liquid scintillation counter.

3.6. Quantifying Protein Levels

1. The protein GFP expressed from the co-transfected plasmid pCDNA3/GFP can be used as a transfection marker. The levels of β-actin detected by Western blotting can be used as an indirect measure of the cell number (see Note 6).
2. For each combination of chaperone tested, the activity of firefly luciferase can be calculated relative to the actual level of EGFP (transfection) and β-actin (cell density). The levels of EGFP and β-actin are determined by Western analysis in conjunction with the application of Image J software (http://rsbweb.nih.gov/ij) and the intensity of the luciferase activity (Wallac 1450 MicroBeta Trilux liquid scintillation counter) corrected for the EGFP and β-actin values. For each chaperone combination tested, triplicate biological samples should be analysed in order to assess how their expression affects recombinant protein production. Results are compared to cells carrying the backbone vector pcDNA4.
3. Statistical significance is evaluated using a one-way analysis variance (ANOVA) test and data plotted using the Minitab statistical analysis software. A P-value ≤ 0.05 is usually accepted as statistically significant.

4. Notes

1. Alternatively, the following buffers can be used. RIPA buffer: 50 mM Tris–HCl, pH 7.5, 150 mM NaCl, 1% Nonidet P-40, 0.5% sodium deoxycholate, and 0.1% SDS. L buffer: 250 μL 0.1% Triton X100, 20 mM HEPES, pH 7.2, 100 mM NaCl, one Complete Protease Inhibitor Cocktail mini-tablet per 5 mL buffer, and 0.2 mM PMSF.
2. Alternatively, cells can be shredded using a cell preparation kit, such as the QIAshredder homogenizer (Qiagen).
3. In order to reduce non-specific binding, an extra high-salt wash (10 mM Tris–HCl, pH 7.4, 1 M NaCl, and 0.1% Tween 20) can be performed.
4. Quantification of the chemiluminescent signal can be done by scanning densitometry of the films. Alternatively, the protein levels can be measured using immunofluorescence using FITC-conjugated secondary antibodies and capturing the signal digitally with an instrument, such as the FujiFilm FLA-5100 plus.
5. Where possible, we aim to detect two different proteins on the same blot using the required combination of antibodies (see Fig. 1 for an example). For membrane re-probing, the blotting membrane is stripped of bound antibodies using Restore™ Plus Western Blot Stripping buffer (Thermo Scientific).
6. Alternatively, cell density can also be assessed by counting cell number manually (10 μL of cell suspension) using a haemocytometer (Neubauer, depth 0.1 mm, 1/400 mm^2) or semi-automatically (100 μL of cells) using a cell analyser, such as the Vi-Cell (Beckman Coulter). Both methods rely on using trypan blue dye exclusion on detached cells to determine cell viability. One volume of 0.4% trypan blue is added to one volume of cells. Only dead cells take up the dye.

Acknowledgments

The work on molecular chaperones carried out in the authors' laboratory is funded by the Biotechnology and Biological Sciences Research Council (BBSRC) and the Engineering and Physical Sciences Research Council (EPSRC).

References

1. Liberek, K., Lewandowska, A., and Zietkiewicz, S. (2008) Chaperones in control of protein disaggregation, *EMBO J 27*, 328–335.
2. Vicart, P., Caron, A., Guicheney, P., et al., (1998) A missense mutation in the alphaB-crystallin chaperone gene causes a desmin-related myopathy, *Nat Genet 20*, 92–95.
3. Cromwell, M. E., Hilario, E., and Jacobson, F. (2006) Protein aggregation and bioprocessing, *Aaps J 8*, E572-579.
4. Kopito, R. R., and Sitia, R. (2000) Aggresomes and Russell bodies. Symptoms of cellular indigestion?, *EMBO Rep 1*, 225–231.
5. Meusser, B., Hirsch, C., Jarosch, E., et al., (2005) ERAD: the long road to destruction, *Nat Cell Biol 7*, 766–772.
6. Mogk, A., Tomoyasu, T., Goloubinoff, P., et al., (1999) Identification of thermolabile Escherichia coli proteins: prevention and reversion of aggregation by DnaK and ClpB, *Embo J 18*, 6934–6949.
7. Queitsch, C., Hong, S. W., Vierling, E., et al., (2000) Heat shock protein 101 plays a crucial role in thermotolerance in Arabidopsis, *Plant Cell 12*, 479–492.
8. Sanchez, Y., and Lindquist, S. L. (1990) HSP104 required for induced thermotolerance, *Science 248*, 1112–1115.
9. Sanchez, Y., Taulien, J., Borkovich, K. A., et al., (1992) Hsp104 is required for tolerance to many forms of stress, *Embo J 11*, 2357–2364.
10. Elliott, J. G., Oliver, J. D., and High, S. (1997) The thiol-dependent reductase ERp57 interacts specifically with N-glycosylated integral membrane proteins, *J Biol Chem 272*, 13849–13855.
11. Bebbington, C. R., Renner, G., Thomson, S., et al., (1992) High-level expression of a recombinant antibody from myeloma cells using a glutamine synthetase gene as an amplifiable selectable marker, *Biotechnology (N Y) 10*, 169–175.
12. Josse, L., Smales, C. M., and Tuite, M. F. (2010) Transient expression of human TorsinA enhances secretion of two functionally distinct proteins in cultured Chinese hamster ovary (CHO) cells, *Biotechnol Bioeng 105*, 556–566.
13. Chenuet, S., Martinet, D., Besuchet-Schmutz, N., et al., (2008) Calcium phosphate transfection generates mammalian recombinant cell lines with higher specific productivity than polyfection, *Biotechnol Bioeng 101*, 937–945.
14. Godbey, W. T., Wu, K. K., and Mikos, A. G. (1999) Tracking the intracellular path of poly(ethylenimine)/DNA complexes for gene delivery, *Proc Natl Acad Sci USA 96*, 5177–5181.
15. Meares, G. P., Zmijewska, A. A., and Jope, R. S. (2008) HSP105 interacts with GRP78 and GSK3 and promotes ER stress-induced caspase-3 activation, *Cell Signal 20*, 347–358.
16. Raviol, H., Sadlish, H., Rodriguez, F., et al., (2006) Chaperone network in the yeast cytosol: Hsp110 is revealed as an Hsp70 nucleotide exchange factor, *Embo J 25*, 2510–2518.
17. Prodromou, C., and Pearl, L. H. (2003) Structure and functional relationships of Hsp90, *Curr Cancer Drug Targets 3*, 301–323.
18. Hartl, F. U. (1996) Molecular chaperones in cellular protein folding, *Nature 381*, 571–579.
19. Bukau, B., and Horwich, A. L. (1998) The Hsp70 and Hsp60 chaperone machines, *Cell 92*, 351–366.
20. Albanese, V., Yam, A. Y., Baughman, J., et al., (2006) Systems analyses reveal two chaperone networks with distinct functions in eukaryotic cells, *Cell 124*, 75–88.
21. Dekker, C., Stirling, P. C., McCormack, E. A., et al., (2008) The interaction network of the chaperonin CCT, *Embo J 27*, 1827–1839.
22. Qiu, X. B., Shao, Y. M., Miao, S., et al., (2006) The diversity of the DnaJ/Hsp40 family, the crucial partners for Hsp70 chaperones, *Cell Mol Life Sci 63*, 2560–2570.
23. Arrigo, A. P., Simon, S., Gibert, B., et al.,(2007) Hsp27 (HspB1) and alphaB-crystallin (HspB5) as therapeutic targets, *FEBS Lett 581*, 3665–3674.
24. Hayes, D., Napoli, V., Mazurkie, A., et al., (2009) Phosphorylation dependence of hsp27 multimeric size and molecular chaperone function, *J Biol Chem 284*, 18801–18807.
25. Craven, R. A., Egerton, M., and Stirling, C. J. (1996) A novel Hsp70 of the yeast ER lumen is required for the efficient translocation of a number of protein precursors, *Embo J 15*, 2640–2650.
26. Steel, G. J., Fullerton, D. M., Tyson, J. R., and Stirling, C. J. (2004) Coordinated activation of Hsp70 chaperones, *Science 303*, 98–101.
27. Ellgaard, L., and Ruddock, L. W. (2005) The human protein disulphide isomerase family: substrate interactions and functional properties, *EMBO Rep 6*, 28–32.
28. Kleizen, B., and Braakman, I. (2004) Protein folding and quality control in the endoplasmic reticulum, *Curr Opin Cell Biol 16*, 343–349.
29. Lee, Y. Y., Wong, K. T., Tan, J., et al., (2009) Overexpression of heat shock proteins (HSPs)

in CHO cells for extended culture viability and improved recombinant protein production, *J Biotechnol 143*, 34–43.

30. Borth, N., Mattanovich, D., Kunert, R., et al., (2005) Effect of increased expression of protein disulfide isomerase and heavy chain binding protein on antibody secretion in a recombinant CHO cell line, *Biotechnol Prog 21*, 106–111.
31. Hwang, S. O., Chung, J. Y., and Lee, G. M. (2003) Effect of doxycycline-regulated ERp57 expression on specific thrombopoietin productivity of recombinant CHO cells, *Biotechnol Prog 19*, 179–184.
32. Mohan, C., and Lee, G. M. (2010) Effect of inducible co-overexpression of protein disulfide isomerase and endoplasmic reticulum oxidoreductase on the specific antibody productivity of recombinant Chinese hamster ovary cells, *Biotechnol Bioeng*. 107: 337–346.
33. Chung, J. Y., Lim, S. W., Hong, Y. J., et al., (2004) Effect of doxycycline-regulated calnexin and calreticulin expression on specific thrombopoietin productivity of recombinant Chinese hamster ovary cells, *Biotechnol Bioeng 85*, 539–546.

Chapter 33

High-Throughput Baculovirus Expression in Insect Cells

Richard B. Hitchman, Robert D. Possee, and Linda A. King

Abstract

Historically, it has been proved difficult to adapt the traditional baculovirus expression systems to an automated platform because of the complexity of the processes involved. One of the major bottlenecks is the selection of recombinant from parental viruses. We have developed a bacmid vector (*flash*BAC™) that does not require any form of selection pressure to separate recombinant virus from nonrecombinant parental virus. The method relies on homologous recombination in insect cells between a transfer plasmid containing the gene of interest and a replication-deficient bacmid. The gene of interest replaces the bacterial replicon at the polyhedrin locus, simultaneously restoring a virus gene essential for replication, and as only recombinant virus can replicate, no further separation techniques are required. This chapter describes methods for producing and expression testing multiple recombinant baculoviruses on automated platforms using the *flash*BAC system.

Key words: Baculovirus, High-throughput, Insect cells, Protein expression, *flash*BAC, Automated, Robot

1. Introduction

Automation of the baculovirus expression system has been described previously in the literature, but in these instances they almost always refer to bacmid DNA transformation of bacterial cells (1–7). The advantage of this method is that it can utilize well-established bacterial expression protocols for robotic systems, which are easily adapted and modified toward making recombinant baculoviruses. Implementing an automated system for making recombinant viruses by homologous recombination in insect cells has traditionally been much more challenging. This is mainly due to the purification process required for the selection of recombinant viruses away from a parental background. Plaque purification is a skilled and time consuming technique that is difficult to adapt to an automated platform, although several companies have commercial

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_33, © Springer Science+Business Media, LLC 2012

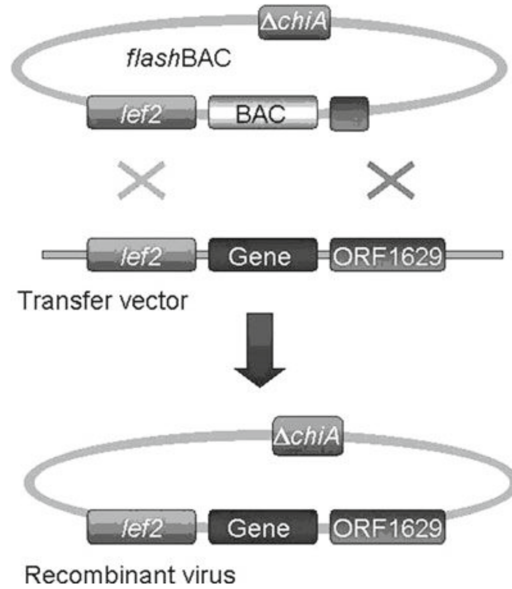


Fig. 1. Schematic showing steps in making a recombinant virus using *flash*BAC™. The *flash*BAC DNA (1) is mixed with a transfer vector (2) to generate a recombinant virus (3) via homologous recombination in insect cells. Reproduced from *Expression Systems*, *Methods Express* 2007 Chapter 9, 149–168 with permission from *Scion Publishing Ltd.*

systems that can do this. However, the costs of these robotic platforms are generally outside the range of most academic research and also many industrial laboratories. Recently, we have developed a system of producing recombinant viruses in insect cells by homologous recombination using a bacmid vector (8). This unique system (*flash*BAC™) combines the advantages of rapid virus production in insect cells with the flexibility of a bacmid-based system, removing the need for plaque purification and facilitating the development of an automated process that can be carried out using a standard liquid handling robot (Fig. 1). This chapter describes the use of *flash*BAC and associated technologies, for the production of multiple recombinant viruses on robotic platforms.

2. Materials

2.1. Insect Cell Culture

1. *Spodoptera frugiperda* 9 (Sf9) insect cells: These can be obtained from multiple sources including the Health Protection Agency in the UK, Sigma, Invitrogen, or the American Type Culture Collection.
2. EX-CELL™ 420 serum-free insect cell culture medium (SAFC Biosciences).
3. 125-mL Shaker flasks and 5-mL T-flasks (Corning Life Sciences).

4. Incubator shaker (110–130 rpm) set at 27–28°C.
5. Tissue culture hood.
6. Inverted phase-contrast microscope.
7. Neubauer counting chamber or automated cell counter (e.g., Countess and disposable counting chambers) (Invitrogen).
8. 0.4% (w/v) Trypan blue (Invitrogen).
9. Sodium dodecyl sulfate (SDS) (Sigma).

2.2. Recombinant Baculovirus Production in 24-Well Plates

1. CAS 1200 (Qiagen) or similar robotic workstation situated within a sterile environment, e.g., a Class II microbiological cabinet.
2. Carbon coated filtered tips for liquid level sensing (50 and 200 μL volume) (Alpha Laboratories).
3. 24-well tissue culture plates (Fisher).
4. Inverted phase-contrast microscope.
5. *flash*BAC kit (Oxford Expression Technologies) containing *flash*BAC DNA (100 ng per cotransfection) and positive control transfer plasmid (pAcRP23.lacZ; 500 ng per cotransfection).
6. Transfer plasmid DNA (500 ng per cotransfection) containing your gene of interest with suitable tag for expression testing (e.g., 6× histidine tag). Keep DNA sterile.
7. Sterile bijoux (Scientific Laboratory Supplies), or similar.
8. Sf9 cells from a culture in log phase growth (require 2×10^5 cells/well).
9. EX-CELL™ 420 serum-free insect cell culture medium (SAFC Biosciences).
10. Transfection reagent, e.g., Nanofectin (PAA Laboratories), Lipofectin, or Cellfectin® (Invitrogen).
11. Incubator set at 27–28°C.
12. 1% Virkon (Amtec) or other suitable disinfectant.

2.3. Recombinant Protein Expression in 24-Well Plates

1. CAS 1200 (Qiagen) or similar robot situated within a sterile environment, e.g., a Class II microbiological cabinet.
2. Carbon coated filtered tips for liquid level sensing (50 and 200 μL volume) (Alpha Laboratories).
3. 24-well tissue culture plates.
4. Inverted phase-contrast microscope.
5. Recombinant baculoviruses (Subheading 3.2).
6. Sf9 cells from a culture in log phase growth (require 4×10^5 cells/well).

7. EX-CELL™ 420 serum-free insect cell culture medium (SAFC Biosciences). Store in the dark at 4°C.
8. Incubator set at 27–28°C.
9. 1% Virkon (Amtec) or other suitable disinfectant.

2.4. SDS-Polyacrylamide Gel Electrophoresis of Recombinant Protein

1. Mini-PROTEAN gel system and gel assembly apparatus (BioRad), or similar.
2. Vacuum pump connected to side arm flask with bung.
3. Boiling water bath.
4. 10-well combs (1 mm width) or 15-well combs (0.75 mm width).
5. 1.5 M Tris–HCl pH 8.8: Mix Tris base 181.5 g with 100 mL distilled water. Adjust to pH 8.8 with 6N HCl. Make to 150 mL with distilled water and autoclave to sterilize.
6. 1 M Tris–HCl pH 6.8: Mix 121.1 g Tris base with 700 mL of distilled water and adjust the pH at 6.8 by adding 118 mL concentrated 6N HCl and then fine adjust if needed with acid. Add distilled water until final volume is 1 L and autoclave to sterilize.
7. 10% SDS: Add 10 g SDS in 90 mL water. Stir gently to dissolve. Make to 100 mL.
8. Ultra Pure Protogel 30% (w/v) acrylamide (National Diagnostics). This is a neurotoxin when unpolymerized and so care should be taken not to receive exposure.
9. *N*,*N*,*N*,*N*′-Tetramethyl-ethylenediamine (TEMED, Sigma). Store at 4°C.
10. Ammonium persulphate: Prepare a 10% (w/v) solution in water and immediately freeze in single use (200 μL) aliquots at −20°C.
11. 50% (v/v) Water saturated *n*-butanol: Mix 250 mL *n*-butanol with 250 mL distilled water in a glass bottle and allow to separate. The top layer contains water-saturated *n*-butanol. Store at room temperature (RT).
12. 10× SDS-polyacrylamide gel electrophoresis (SDS-PAGE) running buffer: Mix 900 mL of distilled water with 144 g glycine, 36.3 g Tris base, and 10 g SDS. Dilute 100 mL of this stock into 900 mL distilled water (1:10) to make 1× SDS-PAGE running buffer. Store at RT.
13. 4× SDS-PAGE loading buffer: Mix 12.5 mL 0.5 M Tris–HCl pH 6.8, 10 mL glycerol, 2 mL 2-mercaptoethanol, 0.8 g SDS, and 0.04 g bromophenol blue. Store frozen at −20°C.
14. PageRuler™ Prestained Protein Ladder Plus molecular weight markers (Fermentas Life Sciences), or similar.

2.5. Western Blotting of Recombinant Proteins

1. Phosphate-buffered saline with Tween (PBST): Prepare 10× stock with 80 g NaCl, 2 g KCl, 14.4 g Na_2HPO_4, 2.4 g KH_2PO_4, 10 mL Tween-20 in 800 mL distilled water. Adjust to pH 7.2 and add distilled water to 1,000 mL.
2. Blocking buffer: Mix 5% (w/v) nonfat dried milk in PBST.
3. Primary antibody: Mouse anti-His antibody (AbD Serotec).
4. Secondary antibody: Antimouse IgG conjugated to horse radish peroxidase (e.g., AbD Serotec).
5. Enhanced chemiluminescent (ECL) Western Blotting Substrate (Pierce).
6. Typhoon 8600 Variable Mode Imager with ImageQuant™ software (Molecular Dynamics) with X-ray film and cassettes (or similar phosphoimaging device).

2.6. Amplification of Recombinant Viruses in 24 Deep-Well Plates

1. Sf9 cells (0.9×10^6 cells/dish) from a culture in log phase growth.
2. EX-CELL™ 420 serum-free insect cell culture medium (SAFC Biosciences).
3. Incubator set at 27–28°C.
4. 1% Virkon (Amtec), or other suitable disinfectant.
5. Inverted microscope.
6. Plastic box to house dishes in the incubator.
7. Sterile bijoux or similar.
8. Selection of sterile pipettes (1–10 mL).
9. Virus from Subheading 2.3.

2.7. Quantification of Recombinant Baculovirus by Quantitative PCR

1. Stock of virus to be titrated (require 200 μL of virus).
2. *baculo*QUANT™ Kit (Oxford Expression Technologies) containing the following: Forward primer 2.5 μM (CGGCGTGAGTA TGATTCTCAAA), reverse primer 2.5 μM (ATGAGCAGAC ACGCAGCTTTT), dual labeled probe 2.5 μM (FAM-AAAAGTCTACGTTCACCACGCGCCAAA-TAMRA), DNA standard (25 ng/μL), and *baculo*QUANT™ Manual.
3. High Pure Viral Nucleic Acid Kit (Roche Molecular Diagnostics).
4. A QPCR mastermix, i.e., ABsolute Blue QPCR Low ROX Mix (ABgene).
5. Applied Biosystems 7500 Real-Time PCR Sequence Detection System.
6. Optical 96-well plates and plate seals (Applied Biosystems).
7. Filter pipette tips (Alpha Laboratories).

3. Methods

3.1. Insect Cell Culture

When working with insect cells aseptic technique must be used at all times and manipulation should be carried out in a microbiological safety cabinet or around a flame. Carry out any cell culture work prior to handling virus. It is very important that the insect cells used for the production of recombinant viruses and proteins are of the highest quality, i.e., taken from log phase cultures, not growing in clumps and with >95% viability. This is achieved by subculturing cells before they become overgrown, e.g., before reaching stationary phase, and by using cells that have been subcultured no more than 30 times from frozen stocks. The insect cells used are Sf9 cells (9), a clonal isolate of IPLB-Sf21AE cells (10) originally derived from the pupal ovarian cells of *S. frugiperda* (fall army worm). The optimal temperature for cell growth and infection for Sf9 cells is considered to be 27–28°C. Insect cell culture medium utilizes a phosphate buffering system, as opposed to the carbonate-based buffers used for mammalian cells, meaning that CO_2 incubators are not required.

3.1.1. Subculturing Insect Cells in Adherent Cultures

Cells can be obtained from a number of sources including The Health Protection Agency Culture Collections in the UK. Follow the manufacturer's protocol to raise the cells from liquid nitrogen. However, when reviving cells from liquid nitrogen, we often thaw the cells directly into a shake flask (following the supplier's protocol) and then remove 5 mL into a T-flask as a backup. The maintenance of cells in monolayer culture does not normally require counting of cell densities. Cells are observed using an inverted microscope until they become confluent when they are subcultured by diluting with fresh medium.

1. Observe the health and confluency of the cells using an inverted microscope (×10, ×20 and ×40 objectives) (see Note 1).
2. In a microbiological cabinet, when the cells are ~80% confluent, remove and discard the medium and floating cells from the T25 flask into 5% SDS or an autoclavable discard container. Replace an equal volume (5 mL) of fresh medium into the T25 flask.
3. Dislodge the cells from the surface of the flask by banging down firmly on a solid surface (see Note 2). Do not hold the neck of the flask when carrying out this operation and try to minimize foaming.
4. Using a graduated pipette, gently pipette the culture up and down to break up clumps of cells into a single cell suspension. Avoid causing bubbles and frothing in the culture.
5. Add 0.5 mL of this culture into a new T-flask containing 4.5 mL of fresh medium (1:10 dilution of cells to fresh medium) (see Note 3).

6. The following information should be recorded on each T-flask: (a) cell line, (b) passage number, (c) date of each passage, (d) density the culture has been split to (i.e., 1:5), and (e) medium used.
7. Keep the old flask on the bench as a backup flask and return the new flask to the 28°C incubator.
8. Routinely observe the cell growth and repeat subculturing procedure when cells reach ~80% confluency.

3.1.2. Cell Counting and Viability

When passaging cells in suspension culture, it is important to know the correct cell density to seed subsequent shake flasks and to calculate working cell numbers.

1. In a microbiological cabinet, remove a sample of cells from a culture using a sterile pipette and place into a bijou, avoiding rapid pipetting, which can cause foaming.
2. Place a glass cover slip over the Neubauer counting chambers and using a Pasteur pipette (or other pipette) load approximately 10 μL of this sample onto each side of the slide and allow it to fill the chamber by capillary action.
3. Count all of the cells within the 5 × 5 square grid on the counting chambers, using an inverted microscope (×10 objective). Count cells touching the triple line on the top and left of the squares. Do not count cells touching the triple lines on the bottom or right side of the squares (see Note 4).
4. If necessary, repeat the count to give a more statistically correct estimate of cell density.
5. Calculate the number of cells in 1 mL of culture; by multiplying the average number of cells from the 5 × 5 square by 10^4 (see Note 5). If the cells were diluted before counting, also multiply by the dilution factor.
6. To assess cell viability, mix 0.5 mL volumes of cell suspension and 0.5 mL 0.4% Trypan blue (see Note 6) (2% w/v) (1 in 2 dilution of cells) and incubate for 1–2 min, but no longer than 5 min.
7. Load the cells into a counting chamber and count the unstained (viable) and stained (nonviable) cells separately in the hemocytometer.
8. To obtain the total number of viable cells per milliliter, multiply the total number of viable cells by 2 (the dilution factor for trypan blue). To obtain the total number of cells per milliliter, add up the total number of viable and nonviable cells and multiply by 2.

3.1.3. Subculturing Insect Cells in Suspension Cultures

1. Determine cell number and viability of the cells grown in a T-flask (or a shake culture) as described above in Subheading 3.1.2. Ideally Sf9 cells grown in shake cultures should not be grown beyond 6×10^6 cells/mL.

2. Calculate how much of the existing culture needs to be diluted in fresh medium to obtain the required cell density and volume for the new culture. We generally maintain a 50 mL stock culture of Sf9 cells and these are split to a density of 0.3×10^6 cells/mL when they reach a density of ~3×10^6 cells/mL, generally twice a week (e.g., on a Monday and a Thursday). To do this pipette 45 mL prewarmed media into a fresh flask and then add 5 mL of cells (see Note 3).
3. Record the following information on the flask that the cell culture maintained in: (a) Cell line, (b) passage number, (c) date of each passage, (d) density the culture has been split to, and (e) medium used.
4. Also keep a separate record of liquid nitrogen batch and the date and passage when cells were first raised from liquid nitrogen.
5. Discard any unused cells into SDS or an autoclavable discard container.

3.2. Recombinant Baculovirus Production in 24-Well Plates

Prior to making recombinant viruses the gene of interest need to be transferred into a suitable transfer vector. There are many ways to do this from traditional cut and paste ligation to recombination and ligation independent cloning (LIC) methods. In our hands, the StarGate™ (IBA) system has proved to one of the most powerful in terms of ease of cloning and range of available combinations of plasmids and sequences (tags, signals, etc.). The method described below assumes that the user has already produced a library of recombinant transfer plasmids that are compatible with the *flash*BAC system, e.g., which have sequences for homologous recombination with *flash*BAC in insect cells.

1. Confirm the health and viability of the insect cells to be used under an inverted microscope and prepare a stock of Sf9 cells in EX-CELL 420 medium at a cell density of 5×10^5 cells/mL. Design a simple liquid handling program for the robot (see Note 7) with the following steps: (a) Cell seeding, (b) cotransfection mixture preparation, and (c) aliquoting the cotransfection mixture into the 24 wells containing the seeded cells (see Note 8).
2. Use the robot to aliquot 400 μL (2×10^5 cells/well) into each well of a 24-well tissue culture plate. Allow the cells to settle and attach for 1 h before use at RT.
3. During the 1 h incubation period, program the robot to prepare 24 cotransfection mixes in the wells of a 96-well plate (made from polystyrene) and with U- or V-shaped well (see Note 9).
4. Dispense the following to each of 24 wells of a 96-well plate in the following order to give a final volume of 20 μL:

   8 μL EX-CELL 420 medium

   2 μL transfection reagent (e.g., Lipofectin™ or Nanofectin™)

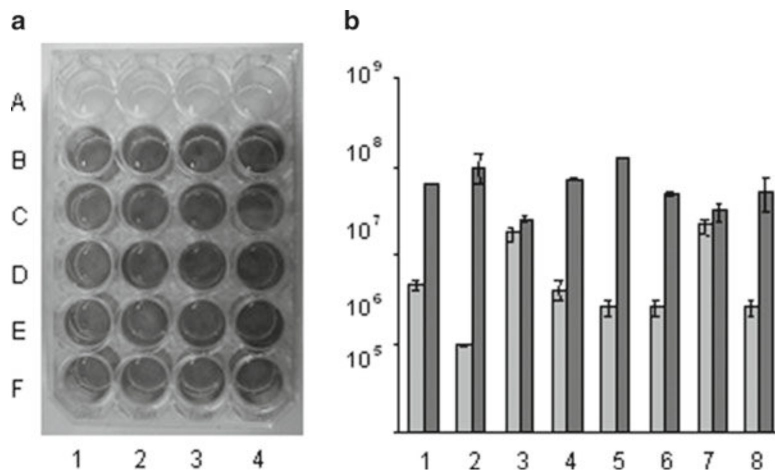


Fig. 2. Production of recombinant viruses in a 24-well plate. (**a**) 24-well plate containing 20 recombinant baculoviruses expressing β-galactosidase. Four mock infected controls are also shown in *lanes A1–A4* corresponding to mock-infected cells, bacmid DNA only, transfer vector DNA only and transfection reagent only, respectively. (**b**) Plaque assay titers (pfu/mL) are shown for eight of the cotransfections carried out in the 24-well plate shown in (**a**) (*light gray bars*), and plaque assay titers (pfu/mL) for the same eight recombinant viruses after amplification in shake flasks (*dark gray bars*). *Error bars*, s.d. $n=3$. Reproduced from (8) with permission from Wiley.

5 μL *flashBA*C DNA (100 ng)

5 μL transfer plasmid DNA (500 ng) containing the gene of interest (include one reaction containing the pAcRP23.lacZ positive control transfer plasmid containing *lacZ*, a mock-transfected reaction, e.g., minus transfer plasmid DNA, and a nontransfected control).

5. Mix the reagents by gentle pipetting (dispense and aspirate three times at slow speed).
6. Dispense the 20 μL cotransfection mix to the appropriate wells of the 24-well plate containing the cell monolayers (see Note 10).
7. Replace the lid and seal with parafilm to prevent evaporation. Incubate at 28°C for 5 days.
8. After 5 days, check the cells have grown and appear healthy in the mock and nontransfected control wells using an inverted microscope. Program the robot to harvest the culture medium containing recombinant virus from each well (or do this manually) and store at 4°C in the dark.
9. Add 100 μL media containing 5 μL X-gal (20% w/v) to the cells cotransfected with the pAcRP23.lacZ positive control transfer plasmid. *LacZ* expression should become apparent by the blue coloration produced by the β-galactosidase protein, confirming the presence of recombinant virus (Fig. 2).

3.3. Recombinant Protein Expression in 24-Well Plates

When making large numbers of recombinant viruses it is useful to screen at an early stage for the presence of the protein being expressed. Inoculating a second 24-well plate provides a rapid test of expression.

1. Prepare 10 mL of Sf9 cells in serum-free medium at a cell density of 5×10^5 cell/mL (see Note 11).
2. Use the robot to aliquot 400 μL (4×10^5 cells/well) into each well of a 24-well tissue culture plate. Allow the cells to settle and attach for 1 h before use.
3. During this incubation period, prepare the robotic system to dispense an aliquot of the recombinant viruses (generated in Subheading 3.2) directly onto these cells using standard liquid handling protocols (see Note 12).
4. Add 100 μL of each virus into each of the wells of the 24-well plate, allowing a well for a noninfected control and a well for a positive control virus containing *lac*Z.
5. Replace the lid and seal with parafilm to prevent evaporation. Incubate at 28°C for 48–72 h.
6. After 48–72 h, check the cells have grown and appear healthy in the noninfected control well using an inverted microscope. Harvest the culture medium and cells from each well into individual sterile containers.
7. For secreted proteins, the culture medium can be removed using the robot (or manually) and used directly for testing, for example, by SDS-PAGE and Western blot (Subheading 3.4); however, you may need to remove any floating cells by centrifugation prior to analysis.
8. The remaining cell monolayer can be lysed directly in the wells by the addition of 100 μL lysis buffer and medium speed repetitive pipetting. This cell lysate can then be removed and processed for SDS-PAGE and Western blot analysis.
9. Add 100 μL media containing 5 μL X-gal (20% w/v) to the cells cotransfected with the pAcRP23.lacZ positive control transfer plasmid and incubate at RT. *Lac*Z expression should become apparent by the blue coloration produced by the β-galactosidase protein. Alternatively, this sample can be processed for Western blot analysis of lacZ protein (Fig. 3).

3.4. SDS-Polyacrylamide Gel Electrophoresis of Recombinant Protein

Polyacrylamide gel electrophoresis (PAGE) is the most widely used method for analyzing protein mixtures qualitatively. The technique is based on determining the molecular mass of a protein. In order to do this, it is necessary to carry out the electrophoresis in the presence of the anionic detergent SDS. This detergent unfolds the protein and interacts with the unfolded chain such as to give a constant charge density. This means that

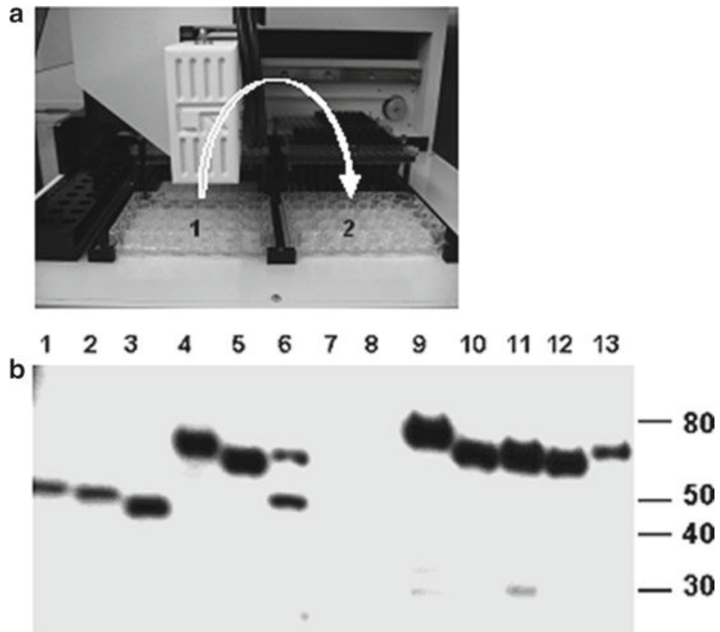


Fig. 3. Protein expression screening in 24-well plate. (**a**) A CAS-1200™ was used for plate to plate transfer of virus inoculum (1) onto insect cells (2) for protein expression. (**b**) Corresponding Western blot of supernatants from (**a**): Thy-1 (*lane 1*), rat Thy-1 (*lane 2*), CD7 (*lane 3*), CD200 receptor (*lane 4*), CD200 (*lane 5*), endothelial cell adhesion molecule (*lane 6*), CADM3 (*lane 7*), CD96 (*lane 8*), CEACAM8 (*lane 9*), PDL2 (*lane 10*), CRACC (*lane 11*), CD58 (*lane 12*), SLAM (*lane 13*); molecular weight markers (kDa). Reproduced from (8) with permission from Wiley.

separation is only based upon molecular mass. Calibration of the gel with marker proteins of known molecular mass allows us to estimate the molecular mass of unknown proteins. The method described below is a discontinuous buffer system, where the buffer in the reservoirs has a different pH and ionic strength from the buffer that casts the gel. Proteins are loaded on a stacking gel, where they can be concentrated into a sharp band before they enter the main separating gel.

These instructions assume the use of the mini-PROTEAN gel system (Bio-Rad) but they are easily adaptable to other minigel systems. It is critical that the glass plates for the gels are cleaned and rinsed extensively with distilled water. The casting stand gaskets must not be soaked for extended periods of time as they will swell up.

1. Assemble the unit according to the manufacturer's instruction, ensuring that the glass plates are flush and spacer plates are correctly orientated to avoid leaks.
2. Prepare a 0.75-mm thick, 12% resolving gel monomer by mixing 3.8 mL of 1.5 M Tris–HCl pH 8.8, 6 mL 30% acrylamide/bis solution, 4.9 mL water, and 150 μL 10% SDS.

3. Degas the solution under vacuum. Fill the vacuum pump cold trap with fresh water and then add some ice until the level reaches just below the overflow hole. Pour the mixture from step 2 into a side arm flask, replace the bung and attach via the side arm to the vacuum pump. Turn on the vacuum and degas the solution for 10–15 min, until no more bubbles are formed. Turn off the vacuum and empty the water from the cold trap.
4. Add 150 μL ammonium persulfate solution (APS) (see Note 13) and 6 μL TEMED. Mix by gently swirling.
5. Pour the gel smoothly to prevent it from mixing with air, leaving space for a stacking gel (1 cm below the comb teeth), and overlay with water-saturated isobutanol or water (see Note 14). The gel should polymerize in about 35–45 min.
6. Pour off the isobutanol and rinse the top of the gel twice with water or running buffer. At this point, the resolving gel can be stored at RT overnight; however, overlay with 5 mL of a 1:4 dilution of 1.5 M Tris–HCl pH 8.8 buffer to keep it hydrated.
7. Prepare the stacking gel by mixing the following: 0.38 mL of 1 M Tris–HCl pH 6.8, 0.5 mL 30% acrylamide/bis solution, 2.1 mL water, 30 μL 10% SDS. Then, degas the solution under vacuum for 15 min before adding 30 μL APS and 3 μL TEMED. Use about 0.5 mL of this to quickly rinse the top of the gel and dry with filter paper. Then, pour the stacker until it reaches the top of the short plate and insert the comb. The stacking gel should polymerize within 30 min.
8. Prepare the running buffer by diluting 100 mL of the 10× running buffer with 900 mL of water in a measuring cylinder and then add it to the upper and lower chambers of the gel unit.
9. Once the stacking gel has set, carefully remove the comb and use a syringe fitted with a 22-gauge needle to wash the wells with running buffer or water.
10. Prior to electrophoresis add 4× loading dye to each sample and boil at 100°C for 3–5 min.
11. Load 15 μL of each sample in a well using a pipette or a Hamilton syringe. Include one well for prestained molecular weight markers. If samples overspill into another well, rinse out carefully with running buffer before loading the next sample; however, this can be avoided by using a Hamilton syringe to load the wells. Rinse the syringe between samples.
12. Complete the assembly of the gel unit and connect to a power supply. The gel can be run at 200 V for approximately 35–45 min. Continue until the dye fronts are just running off the end of the gel.

3.5. Western Blotting of Recombinant Proteins

This protocol briefly describes the use of the iBlot™ Dry Blotting System (Invitrogen) to rapidly transfer the recombinant protein from the SDS-PAGE to a nitrocellulose membrane and is taken from the manufacturer's instructions.

1. Remove the gel from the cassette after completion of electrophoresis. If required wash the gel briefly in deionized water to remove any small gel pieces attached to the gel.
2. Use the appropriate gel transfer stacks, e.g., regular for blotting several mini gels and Mini for a single gel (see Note 15).
3. Open the iBlot system lid and remove the iBlot anode stack, Bottom (see iBlot manual) from the package. Keep the stack in the clear tray and place on the blotting surface of the iBlot and align it to the right hand side gel barriers.
4. Place the prerun gel on the transfer membrane of the anode stack and then overlay with iBlot Filter paper presoaked in distilled water and remove air bubbles using a roller (see Note 16).
5. Remove the cathode stack, Top (see iBlot manual) from its package and place over the filter paper with the electrode side facing up and aligned to the right hand edge. Roll again to remove air bubbles.
6. Place a disposable sponge with the metal contact in the upper right hand corner of the lid and close.
7. Select iBlot program 2 (P2), 23 V for 6 min.
8. After the program is complete, remove the nitrocellulose membrane from the stacks using forceps. Discard the stacks and the sponge.
9. The colored molecular weight markers should be clearly visible on the membrane.
10. Block membrane for 60 min with 10 mL blocking solution (PBST, 0.1% Tween® 20, and 5% skimmed milk) at RT on a rocking platform (transferred proteins-side upward).
11. Wash the membrane in 10 mL PBST for 5 min at RT on a rocking platform.
12. Dilute the mouse anti-His antibody 1:5,000 (v/v) in 10 mL blocking solution and incubate at RT on a rocking platform for 60 min.
13. Remove the primary antibody solution and wash the membrane in 10 mL PBST for 10 min. Repeat four times in total.
14. Incubate the membrane for 60 min in 10 mL horseradish peroxidase conjugated secondary antibodies diluted in blocking solution, according to manufacturer's instructions.
15. Remove the secondary antibody solution and wash the membrane in 10 mL PBST for 10 min. Repeat four times in total.

16. Finally, rinse the membrane in 20 mL water for 2 min and repeat this wash.
17. In a dark room, add 1 mL ECL Reagent 1 to 1 mL ECL Reagent 2 (both from Amersham ECL™ kit) and then immediately add to the blot for 1 min on a rocking platform to ensure even coverage. Use 0.125 mL Working Solution per cm^2 of membrane.
18. After 1 min remove the reagent mix, blot with Kim-Wipes, and then place between the leaves of an acetate sheet protector that has been cut to the size of an X-ray film cassette. Press out any bubbles from between the sheets.
19. The acetate containing the membrane is then placed in an X-ray film cassette with film for a suitable exposure time, typically 5–45 min.
20. Scan the film using the Typhoon 8600 Variable Mode Imager and analyze with ImageQuant™ software (or with other suitable device), according to the manufacturer's instructions. If the signal is too intense, try reducing the exposure time (see Note 17).

3.6. Amplification of Recombinant Viruses in 24 Deep-Well Plates

1. Growth of insect cells and viruses in 24 deep-well blocks has been described previously (11, 12) and is a rapid method for amplifying multiple viruses for future work. This method may also be used for protein expression and is especially useful for optimization of expression parameters where multiple viruses are being tested in different cell lines, at different moi, over time and can also be used to produce a high titer stock of virus which can be used to directly inoculate cells for protein production, for example 1 L of High-Five™ cells at m.o.i. 1. Alternatively, it can be used as a master stock of virus to produce subsequent batches for use in protein production (Fig. 4).

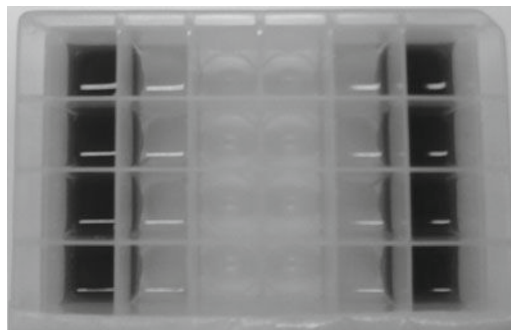


Fig. 4. 24 deep-well block. *Lanes 1* and *6* show wells containing Sf 9 cells infected with a recombinant virus expressing β-galactosidase. *Lanes 2* and *5* show wells containing mock-infected Sf 9 cells. *Lanes 3* and *4* are empty wells. Image supplied by E. Locanto (Oxford Expression Technologies Ltd).

2. Observe health and viability of cells under an inverted microscope (see Note 18).
3. Set up 50 mL cultures at low density, e.g., 0.5×10^6 Sf9 cells/mL and allow the cells to grow to the required density for approximately 48 h. Monitor cell growth using an inverted microscope and hemocytometer as described in Subheading 3.1.2 or using a cell counter.
4. When the cells reach the correct density (2×10^6 Sf9 cells/mL) add 4.5 mL to each well of a 24 deep-well plate.
5. Add 0.05 mL of the recombinant virus seed stock (from Subheading 3.2) to 1 well, repeat up to 24 wells and seal the plate with a breathable membrane. Incubate with shaking at 350 rpm on a shaking incubator until the cells appear well infected (normally 4–5 days) (see Note 19).
6. After 4–5 days, remove the cells by centrifugation at low speed (e.g., $3{,}000 \times g$), at 4°C for 15 min.
7. Decant the clarified culture medium aseptically using the robot either into individual bejoux or a fresh 24 deep-well block and store the recombinant viruses in the dark at 4°C (see Note 20).

3.7. Titration of Recombinant Baculovirus by Quantitative PCR

Prior to using the viruses produced in Subheading 3.5, it is important to determine their titers. A rapid and highly reproducible titration method (*baculo*QUANT) based on real-time quantitative PCR (qPCR) has been developed in our laboratory (13), which offers the versatility to titrate not only single viruses but also multiple viruses in a semiautomated high throughput format. qPCR uses a target-specific probe tagged with a reporter fluorescent dye (FAM) at its 5′ end and a quencher dye (TAMRA) at the 3′ end which anneal between two primer sites. Fluorescent resonance energy transfer (FRET) inhibits the reporter dye until it is cleaved by the 5′–3′ exonuclease activity of *Taq* polymerase during PCR amplification. This produces a fluorescence signal which increases in intensity proportional to the amount of PCR product produced. The increase in fluorescence can be quantified in real-time using the Cycle threshold (C_t) of the reaction: The point at which the fluorescence intensity rises above the background-noise baseline at the most exponential phase of the PCR reaction. By comparing this Ct value to a standard curve of virus titers, it is possible to establish an accurate titer, which can be converted to equivalent plaque forming units per milliliter (qpfu/mL), using an equation derived from linear regression analysis. The *baculo*QUANT titration method is described below.

1. Remove 200 μL of culture media from your recombinant baculovirus stock (from Subheading 3.5) and isolate the viral DNA using the High Pure Viral Nucleic Acid Kit, following the manufacturer's instructions. Elute the viral DNA from the column in a final volume of 50 μL.

2. Prepare a log series of tenfold dilutions of the DNA standard supplied in the *baculo*QUANT kit in final volumes of 10 μL (25, 2.5, 0.25, 0.025, 0.0025 ng/μL). 2 μL of each of the diluted standards will then be used in each DNA standard reaction (see step 5).
3. Prepare QPCR reactions as a master mix, according to how many viruses you need to titrate. Each virus sample, along with the control sample and each of the five standard DNA samples (step 2), should be analyzed in triplicate. Control reactions contain water in the place of DNA.
4. Aliquot 23 μL of master mix into the appropriate number of wells in a 96-well plate. Each reaction will be carried out in triplicate, i.e., 3×5 wells for standards, 3 wells for water control, and 3 wells per unknown virus.
5. Add 2 μL of your purified viral DNA from step 1 or 2 μL DNA standards/water, giving a final volume of 25 μL per reaction.
6. Dispense each 25 μL reaction into a 96-well plate, making a note of their locations on the plate.
7. Enter the position of each reaction into the ABI Sequence Detection System (ABI-SDS) software along with the fluorescent dyes used and the standard DNA dilutions.
8. Carry out DNA amplification following the manufacturer's instructions, e.g., ABI Quantitative Analysis using the ABI-SDS default cycling conditions (see Note 21).
9. At the end of the qPCR cycle, export the Cycle threshold (Ct) values into an Excel spreadsheet and calculate the mean Ct value for each virus. Using the login and password information provided with the kit enter the Ct values in to the conversion software on the OET Web site (http://www.oxfordexpressiontechnologies.com/). This will convert the qPCR derived titer to its plaque assay equivalent (Qpfu/mL).

4. Notes

1. Healthy cells attach well to the bottom of a T-flask forming a monolayer and double every 18–24 h. Loosely attached cells or cells floating in the medium are frequent in cultures that are overgrown, whereas in healthy cultures only a few will be visible.
2. Trypsin and other enzymes are not recommended to dislodge insect cells as they may damage the cells. Cell scrapers can also cause cell damage and should be used only if absolutely necessary.
3. Each time a culture is split, its passage number increases by 1.

4. Healthy cells look bright, round and refractile and many should be in the process of dividing into daughter cells.
5. The 5×5 square gives the number of cells present in 0.1 μL of culture.
6. Nonviable cells adsorb the dye and appear blue, live and healthy cells exclude the blue-colored dye. Healthy, log-phase cultures should contain more than 95% unstained, viable cells. However, viable cells will adsorb Trypan blue over time, and this can affect counting and viability results.
7. The CAS-1200 robot software is simple to use and once a program is designed it can be stored for future use. It is also useful to carry out a virtual run prior to starting the experiment to test the program. Most robotic software will operate along similar principles but some useful guides include setting the pipette mode to reuse where possible to avoid pipette tip wastage and reducing the pipetting speed to avoid frothing of the media and cell damage.
8. Plasticware used to prepare the transfection mixture must be made from polystyrene and not from polypropylene, which can bind the DNA–lipofectin complexes.
9. There is no need to change the medium, simply add the cotransfection mix to the medium and mix by gentle pipetting.
10. It is recommended to prepare 10% more than required to reduce the likelihood of the robot failing to detect the cells in the dispense tubes when the volume becomes low.
11. Each manufacturer's robot will have different operating protocols, but these generally follow a similar format.
12. Ensure the TEMED and APS are fresh or gels may take longer to polymerize.
13. Do not leave isobutanol for longer than 1 h or the alcohol will dehydrate the top of the gel.
14. It is important to wear gloves when handling the gel, 3 MM papers and nitrocellulose filter, as oils and secretions from the skin will prevent transfer of proteins from the gel to the filter. It is also important that the blotting paper is not larger than the gel, as the overhanging edges of the paper and the filter will contact around the gel, and provide the current with an alternate route, thereby making transfer inefficient.
15. Take great care to remove all air bubbles between the sandwich layers by rolling them out with a 10 mL pipette.
16. Quantification of data can be done by scanning densitometry of the films, provided that care is taken to ensure that the signal has not saturated. It is important to wear gloves when handling the gel, 3 MM papers and nitrocellulose filter, as oils and

secretions from the skin will prevent transfer of proteins from the gel to the filter. It is also important that the blotting paper is not larger than the gel, as the overhanging edges of the paper and the filter will contact around the gel, and provide the current with an alternate route, thereby making transfer inefficient.

17. Take great care to remove all air bubbles between the sandwich layers by rolling them out with a 10-mL pipette.
18. Using a low ROX master mix will allow the system to automatically set the Ct threshold which helps ensure consistency between runs. The use of a high ROX master mix will result in the system being unable to detect the Ct threshold and the operator will be required to set it manually.
19. It is important that the cells are healthy and have good viability to ensure that virus replication occurs efficiently to amplify high titer stocks of virus for subsequent use in expression studies.
20. Virus-infected cells will appear uniformly rounded and grainy with distinct enlarged nuclei.
21. The virus inoculum may be stored for 6–12 months or longer in the dark at 4°C. The titer of the virus will start to decrease over time, but the addition of serum (2–5%) will slow this process down considerably.

Acknowledgments

The authors would like to thank Elisabetta Locanto and Adam Chambers for technical assistance in developing the 24 deep-well block method described in this chapter.

References

1. McCall, E. J., Danielsson, A., Hardern, I. M., et al., (2005) Improvements to the throughput of recombinant protein expression in the baculovirus/insect cell system. *Protein Expr.Purif.* 42, 29–36.
2. Sarafanov, A., and Saenko, E. (2004) High-throughput optimization of protein expression in the baculovirus system based on determination of relative expression efficiency of viral stocks. *Anal. Biochem.* 328, 98–100.
3. Redaelli, L., Zolezzi, F., Nardese, V., et al., (2005) A platform for high-throughput expression of recombinant human enzymes secreted by insect cells. *J. Biotechnol.* 120, 59–71.
4. Schlaeppi, J. M., Henke, M., Mahnke, M., et al., (2006) A semi-automated large-scale process for the production of recombinant tagged proteins in the baculovirus expression system. *Protein. Expr. Purif.* 50, 185–195.
5. Shrestha, B., Smee, C., and Gileadi, O. (2008) Baculovirus expression vector system: an emerging host for high-throughput eukaryotic protein expression. *Methods Mol. Biol.* 439, 269–289.
6. Karkkainen, H. R., Lesch, H. P., Maatta, A. I., et al., (2009) A 96-well format for a high-throughput baculovirus generation, fast titering and recombinant protein production in insect and mammalian cells. *BMC. Res. Notes.* 2, 63.
7. Yao, L., Sun, J., Xu, H., et al., (2010) A novel economic method for high throughput production

of recombinant baculovirus by infecting insect cells with Bacmid-containing diminopimelate-auxotrophic *Escherichia coli*. *J. Biotechnol.* 145, 23–29.

8. Possee, R. D., Hitchman, R. B., Richards, K. S., et al., (2008) Generation of baculovirus vectors for the high-throughput production of proteins in insect cells. *Biotechnol. Bioeng.* 101, 1115–1122.
9. Summers, M. D., and Smith, G. E. (1987) A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures. *Texas Agric. Exp. Stn. Bull. 1555.*
10. Vaughn, J. L., Goodwin, R. H., Tompkins, G. J., et al.,(1977) The establishment of two cell lines from the insect *Spodoptera frugiperda* (Lepidoptera; Noctuidae). *In Vitro* 13, 213–217.
11. Bahia, D., Cheung, R., Buchs, M., et al., (2005) Optimisation of insect cell growth in deep-well blocks: development of a high-throughput insect cell expression screen. *Protein. Expr. Purif.* 39, 61–70.
12. Chambers, S. P., Austen, D. A., Fulghum, J. R., et al., (2004) High-throughput screening for soluble recombinant expressed kinases in *Escherichia coli* and insect cells. *Protein. Expr. Purif. 36*, 40–47.
13. Hitchman, R. B., Siaterli, E. A., Nixon, C. P., et al., (2007) Quantitative real-time PCR for rapid and accurate titration of recombinant baculovirus particles. *Biotechnol. Bioeng.* 96, 810–814.

Chapter 34

Recombinant Protein Expression in Milk of Livestock Species

Zsuzsanna Bösze and László Hiripi

Abstract

Producing complex recombinant proteins in the milk of transgenic animals offers several advantages: large amounts of proteins can be obtained, and in most cases, these proteins are properly folded, assembled, cleaved, and glycosylated. The level of expression of foreign genes in the mammalian gland cannot be predicted in all cases, and appropriate vectors must be used. The main elements of these vectors are as follows: a well-characterized specific promoter, the coding region of the gene of interest, preferably with a homologous or heterologous intron, to improve transcription efficiency, and an insulator or boundary element to counteract the chromosomal position effects at the integration site. Once high expression levels are achieved, and the recombinant protein is purified, an essential step in the analysis of the final product is determining its degree of glycosylation. This is an important readout because it can affect among other parameters the stability and immunogenicity of the recombinant protein.

Key words: Recombinant proteins, Milk, Expression vectors, Transgenic animals, Posttranslational modifications, Purification

1. Introduction

Recombinant protein expression in milk has been extensively studied in the last 25 years. During the past decade, this approach has experienced significant improvement both from the methodological point of view and in terms of products reaching the market. The US Food and Drug Administration (FDA) issued its first approval for ATryn, a biological product produced by genetically engineered animals on February 6, 2009. ATryn (human antithrombin III) is a protein useful to treat a rare clotting disorder. In addition to FDA approval, this product has been also approved by the European Medicines Agency for use in the prevention of clotting conditions during surgical procedures in patients with

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_34, © Springer Science+Business Media, LLC 2012

hereditary antithrombin deficiency. These patients are at high risk of blood clots during medical interventions, such as surgery, and before, during, and after childbirth. ATryn is a therapeutic protein derived from the milk of goats and is manufactured by GTC Biotherapeutics Inc. The amount of ATyrn obtained per year from transgenic goats is equivalent to that obtained from 90,000 human blood samples. The product Ruconest has been approved in the EU for the treatment of hereditary angioedema. This protein is produced by the Pharming Group NV in the milk of transgenic rabbits and is the second transgenic animal derived drug on the market. Other pharmaceutical proteins produced in bovine, goat or rabbit milk are currently at different phases of clinical trials (1). This success story underlines the importance and safety of engineered mammary-gland-derived products and makes milk from transgenic animals a viable alternative production platform for recombinant proteins.

The methods for developing genetically modified livestock animals are numerous, their efficiency varies from one species to the other and they are continuously evolving. Therefore, for details in pronuclear microinjection and for the classic somatic cell cloning, we refer the reader to previous reviews and laboratory manuals (2–5).

In this chapter, we first review the available vectors for mammary-gland-specific transgene expression, and then we focus on the special requirements and methods in transgenic vector design for recombinant protein expression in milk. Last, we present an overview of the methods to isolate and characterize the milk-born recombinant proteins.

2. Main Advantages and Challenges of Recombinant Protein Production in Milk

The advantage of transgenic livestock animals to produce recombinant proteins lies in the high level and quality of the product obtained at a relatively low cost compared to other production systems. The specific features and limitations of the different production systems, including their capacity in proper posttranslational modifications have been summarized in a recent review (1).

The gross composition of the transgenic milk including fat, total protein, lactose, and dry matter should be determined and potential side effects of the transgene expression considered for each established transgenic line. Casein micelle size determination and measurement of viscosity could be also informative parameters to consider (6). A general purification protocol for recombinant proteins expressed in milk involves removal of lipids and caseins, followed by different chromatographic methods, depending on the molecular mass, the level of expression, and the specific properties of the expressed protein. The characterization of the

recombinant proteins should include determination of the molecular mass, the glycosylation pattern, and other posttranslational modifications, if adequate. Determining the degree of glycosylation of recombinant proteins is an important issue because it can affect among other parameters their stability and immunogenicity. Glycoproteins exist as a population of glycoforms due to macro and micro heterogeneity. Species specificity in the glycosylation of proteins such as the case of IgG underscores the importance of choosing the appropriate species for producing fully functional recombinant proteins (7). One of the main challenges for the production of recombinant therapeutic glycoprotein is ensuring the consistency in the glycoforms. To date, only restricted number of publications exists with detailed structural analysis of the N- and O-glycans produced in transgenic animal bioreactors: the recombinant human antitrombin III and erythropoietin produced in transgenic goats (8, 9), the recombinant human lactoferrin produced in transgenic cows (10), the recombinant human protein C and acid α-glucosidase produced in the milk of transgenic rabbits (11, 12), and the recombinant human Factor IX purified from transgenic pig milk (13). Those analyses were performed with different combination of capillary electrophoresis, anion exchange, normal phase and high performance liquid chromatography, mass spectrometry methods. The biological activity of the recombinant protein should be individually evaluated on a case-by-case basis.

3. Vectors for Mammary-Gland-Specific Transgene Expression

To achieve significant accumulation of recombinant proteins highly efficient, mammary-gland-specific expression vectors are preferable. General rules for vector construction to generate animals in which there is mammary-gland-specific expression have been established, yet the optimal vector should be designed on a case-by-case basis. For biotechnological applications the main elements of a mammary-gland-specific vector are the following: (a) a well-characterized mammary-gland-specific promoter, (b) the coding region of the gene of interest, preferably with a homologous or heterologous intron, to improve transcription efficiency, and (c) an insulator or boundary element to counteract the chromosomal position effects at the integration site (Fig. 1).

3.1. BACs Versus YACs Driven Expression?

Transgenes expressed in bacterial artificial chromosomes (BACs) achieve expression at physiological levels with the same developmental timing and expression patterns as endogenous genes. BAC-driven transgene expression tends to be copy number dependent and integration site independent. High expression level of heterologous milk proteins, namely, the human and goat alpha-lactalbumin (14, 15), and the porcine whey acidic protein (16) in

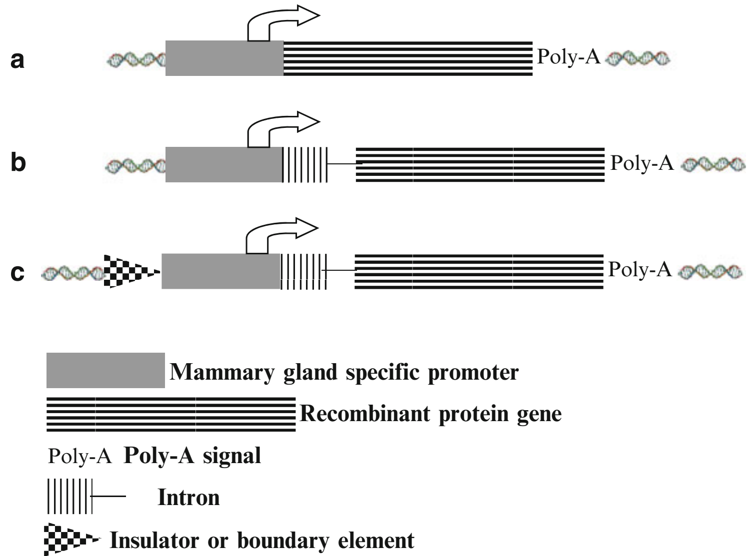


Fig. 1. Design of basic type mammary-gland-specific vectors. (**a**) The coding region of the recombinant protein gene (preferably a genomic clone) placed under a mammary-gland-specific promoter and supplemented with a homologous or heterologous poly-A tail. (**b**) The minigene consists of the cDNA of the recombinant protein, plus a homologous or heterologous intron to increase transcription efficiency, placed under a mammary-gland-specific promoter. (**c**) The constructs described under (**a**) and (**b**) flanked by a boundary element, which shields the construct from undesired chromosomal position effects.

transgenic mice and recently the human lactoferrin in transgenic cattle (17) has been achieved by either YAC or BAC transgenesis.

Yeast artificial chromosomes (YACs), such as BACs, belong to the class of large genomic constructs and ensure optimal transgene expression. However, the growing number of easily accessible BAC libraries and the availability of recombineering methods, makes BAC transgenesis more attractive compared to YACs. BAC clones can be grown and modified in *Escherichia coli*, and this technology offers a fast alternative to the knockout technology in embryonic stem cells. The coding sequence of a milk protein can be replaced through homologous recombination by the coding sequence of the recombinant protein of interest within a couple of days. Homologous recombination could be also used to inactivate genes, which are also present on the selected BAC clone, but their expression in the transgenic animal is not desirable.

3.2. Mammary-Gland-Specific Promoters

Efficient promoters of the following milk protein genes are well characterized: Bovine αs1-, and goat β-casein, rabbit and mouse whey acidic protein, and the ovine β-lactoglobulin (18). All of them are patent-protected and exclusive rights to use them for commercial aims have been granted to various biotechnology companies. As an example, transgene constructs under the rabbit

whey acidic protein promoter for gene expression in the milk of transgenic rabbits is protected by worldwide patents and has been exclusively licensed to BioProtein Technologies.

The commercially available pBC1 vector (pBC1 Milk Expression Vector Kit, Invitrogen) is based on the β-casein promoter. Its efficiency albeit at relative low levels was proven among others in recombinant hirudin expressing transgenic mice (19) and in a patent application for expressing human erythropoietin in milk.

Alternative choices of promoters are the equine β and κ-casein (20) or ovine αS1-casein (21), which have not been used yet to create transgenic animals. Those and other milk protein promoters have been characterized in vitro in mammary gland cell lines.

The combination of lentiviral vectors and indicator genes provides an efficient tool to perform the quick in vivo analysis of the tissue specificity and activity of the selected milk protein gene promoters in transgenic mice (22). It is important to note that the milk protein promoter of choice should not necessarily arise from the same species in which the recombinant protein is going to be produced. Examples that illustrate this concept are the expression of the acid α-glucosidase under the bovine αs1-casein promoter in transgenic rabbits and the expression of the rabbit kappa casein under the rabbit whey acidic promoter in transgenic mice (23, 24).

3.3. Boundary Elements to Improve Expression of Transgenic Constructs

Insulators or boundary elements are genomic regulatory elements that serve to organize the expression of the loci they protect from neighboring DNA sequences. Boundaries are defined functionally in tests that measure their capacities to enhance, block and counteract chromosomal position effects. Therefore, our knowledge of how they work at the mechanistic level is still limited. The first identified and best characterized insulator is the 5′HS4 site at the 5′end of chicken β-globin locus (25). Its functional parts have been already dissected: the nuclear factor CTCF binding site is associated with enhancer-blocking capacity and its role in several insulator elements have been underlined (reviewed in ref. 26). The “barrier activity”, the protection from chromosomal position effects was correlated with the USF1 binding sites in the 5′HS4 insulator element and its associated chromatin remodeling activities, which contribute to the establishment of active chromatin and prevention of adjacent condensed and inactive chromatin (27, 28). The histone hyperacetylation status of chromatin around transgenes flanked with 5′HS4 element is thought to counteract the DNA hypermethylation of promoter elements seen in unprotected transgenes (29). The chicken 5′HS4 element has been successfully used to obtain homogeneous gene expression in transgenic mice (30, 31).

Scaffold or matrix attachment regions (S/MARs) show insulator properties, and their utility to improve gene expression in transgenic mice has been also demonstrated. The S/MARs are A+T

rich sequences defined based on their capacity to interact with the nuclear matrix (32). The best example is the chicken lysozyme MAR element, which has been successfully used to improve gene expression in transgenic mice (33). However, not all S/MAR sequences are able to improve transgene expression. It is thought that there are tens of thousands S/MARs in the genome of higher eukaryotes. There is a category of S/MARs that function as boundary element, when they separate a gene from other gene's regulatory modules (34). Computer prediction of S/MARs is a highly significant task; however, while a number of S/MAR predictors have been developed, their accuracy still needs to be improved (35).

Members of families of DNA retrotransposon-like short repetitive elements (SINE, Alu) have also been found to participate in the establishment of functionally distinct domains within the mammalian genome. A new insulating mechanism has been described in detail in the murine growth hormone locus (36). There a SINE2b element, with a 190 bp long consensus sequence contains two promoters transcribing in opposite directions. Upon activation of the RNA polymerase II promoter, chromatin modification is triggered, and results in decondensation of the locus, its transfer to euchromatin, and this allows interaction with nuclear factors. Another retrotransposon-like element, a 323 bp long Alu sequence identified in the 5′ region of the human keratin-18 gene conferred position-independent and copy number dependent expression of two heterologous transgenes (37).

Locus control regions (LCRs) contrary to insulators are tissue specific regulatory elements. They have strong and specific enhancer activity and contribute to open chromatin structure (38). The principal restriction for the use of LCRs in recombinant gene expression is their tissue specificity. The best well-characterized human β-globin locus LCR provides strong erythroid specific gene expression (39).

Recently, high throughput approaches have allowed the identification and characterization of novel insulator sequences and authors have demonstrated that insulator sequences from mouse and chicken maintain their insulator capacity in the remote vertebrate zebrafish (40).

4. From the BAC Clone to Preparing DNA for Microinjection

4.1. Selection and Characterization of a BAC Clone

For mammary-gland-specific expression, a BAC clone that harbors a milk protein gene should be selected from a BAC library. The ideal BAC clone harbors the milk protein gene of interest including most if not all its validated regulatory elements. BAC libraries from livestock animals are publicly available from different sources and available for the academic scientific community on a

cost-recovery basis (see Note 1). For human and mouse the two major commercially available BAC libraries are the CITB and the RPCI. Four distinct human BAC libraries have been developed by the California Institute of Technology. Research Genetics offers three of these libraries. The RPCI libraries also commercially available were originally constructed by the Roswell Park Cancer Institute.

The number of BACs validated for use to direct mammary-gland-specific expression of a transgene is limited. So far, the whey acidic protein BAC transgenes are the best characterized (41). In case of not characterized BAC clones, the use of comparative genomic approaches is highly recommended in the genomic environment of the gene of interest to identify the conserved regulatory elements among homologous loci in different species (41). The basis of using and alignment analysis of the genomic information from different species is the observation that evolutionarily conserved DNA sequences retained high percentage of homology in noncoding regions (42, 43). Regulatory regions of the milk specific kappa casein gene were compared from six different species and highly conserved mammary-gland-specific transcription factor sites were identified (44).

4.2. Homologous Recombination Based on BAC Modification in E. coli

This technique relies on homologous recombination between cloned or synthesized genomic fragments and the cognate BAC locus, resulting in the insertion of the heterologous sequences or the deletion of endogenous ones. The first method originally developed by Heintz et al. is based on bacterial recombinase A for the recombination of homologous fragments of approximately 1 kb (45, 46), whereas the method developed by Stewart et al. is based on λ phage DNA repair enzymes to induce recombination of homologous fragments as short as 50 nucleotides (47). There are also commercially available kits based on λ-mediated Red/ET recombination (see Note 2).

4.3. Purification Methods for BAC DNA

The most common methods to purify high quality BAC DNA are the alkaline lysis method, a combination of the alkaline lysis method plus a CsCl gradient, and size exclusion chromatography. The efficiency to generate transgenic animals using BAC DNA purified by various methods was recently compared by Van Kuren and collaborators (48). Examination of transgenic efficiency (the number of transgenic founders produced per microinjected eggs) showed no difference between the different purification methods.

Several manufacturers offer commercially available kits for BAC purification based on the alkaline lysis method. In our experience the NucleoBond BAC 100 Kit (Clontech) gives the best compromise of volume management and size (capacity) (see Note 3).

The reader should take into account that measuring DNA concentration by UV spectrometry is not very reliable and may not be

meaningful, since there can be absorbing contaminants as well as significant amounts of *E. coli* DNA. The same is true for flurometry due to the presence of *E. coli* DNA. In our experience a good choice is to evaluate the integrity, quality, and quantity of the BAC DNA by pulsed-field gel electrophoresis using linearized DNA (see Note 4).

A combination of the alkaline lysis and CsCl gradient can be also used for BAC DNA isolation; however, the protocol is more time-consuming and requires an ultracentrifuge.

Although circular BAC DNA can be used for microinjection, linearized DNA is preferable. Separation of the BAC backbone is not necessary but can be beneficial. Separation and purification of the microinjection fragment can be achieved by sepharose 4B-CL chromatography or pulsed field gel electrophoresis followed by gel isolation of the fragment (see Note 5).

5. Purification and Characterization of the Recombinant Protein

Milk collection methods vary depending on the species. Milk of transgenic cows is collected and tested during hormonal induction of lactation at first, while in other livestock species in which the generation time is much shorter milk is collected during normal lactation. Milk is a complex biocolloid system composed of three phases: dispersed lipid phase, aqueous serum phase, and casein micellar phase, which contains 70–80% of milk proteins. The majority of recombinant proteins are naturally secreted and found in the aqueous phase of the milk, which is the milk fraction remaining after the eliminations of lipids and caseins. However, in special cases when the recombinant protein is hydrophobic, it will be found concentrated in the milk fat globules (49, 50). In case of the recombinant human alkaline phosphatase, which was secreted with the milk fat globules, a single phase extraction procedure of whole milk samples was performed with 8% (v/v) butanol at room temperature, with constant shaking for 1 h at 100 rpm to extract the recombinant alkaline phosphatase (51).

To induce lactation in transgenic cows, we recommend that at 6–8 months of age an intramuscular injection of medroxyprogesterone acetate (25 mg/kg/day) and estradiol benzoate (7.5 mg/kg/day) for 7 days is administered (17). After milking, the whey could be obtained by ultracentrifugation at $100{,}000 \times g$ 1 h and 20°C to precipitate casein and discard the fat.

In the case of transgenic goats, the first milking can take place 2 h after kidding. After that, animals can be milked twice daily either by hand or using a portable milk collector. Milking could continue until the animals decrease their milk production to less than 100 g/day for 2 consecutive days (52).

In the case of transgenic pigs, lactating sows should be injected intramuscularly with 30–60 IU of oxytocin and milking should

take place after 2–5 min either by hand or machine. Collection of minimum five daily samples from each sow is recommended. Immediately after collection, the milk should be diluted 1:1 with 200 mM EDTA pH 7.0 to solubilize caseins and then frozen. Small aliquots of the milk/EDTA mixture should be taken and centrifuged for 30 min 3,000 × *g* at 4°C. The fat layer should be separated and discarded (53).

For transgenic rabbits, mothers should be separated from their pups overnight, injected with oxytocin (0.5–2 IU), and then hand milked for 5 min (54). Milk samples could be collected once a day during the 4 weeks long lactation period. Milk samples should be diluted with 3 volumes of distilled water, skimmed by centrifugation at 2,500 × *g* at 4°C for 20 min, freeze-dried and stored at −20°C.

Primarily characterization of recombinant proteins could be performed by Western analysis, with species-specific antibodies and by determining the specific activity of the recombinant protein in the crude milk extract (blood clotting, enzymatic, antibacterial etc.). For further purification steps, different chromatographic methods are available to separate the recombinant protein from all other milk components. Milk with its complex nature is challenging for chromatographic processing and several approaches have been reported to reduce the complexity of milk prior processing. Human myelin basic protein was found associated with the casein micellar phase in the milk of transgenic cows. A new approach was developed for the deconstruction of the calcium core of the casein micelles, which increased the recovery of the recombinant protein by more than 25% (55).

Further purification can be achieved by anion exchange chromatography (56) and high-pressure liquid chromatography (10). The molecular mass of the purified recombinant protein could be determined by matrix-assisted laser desorption-ionization time-of-flight (MALDI-TOF) mass spectrometry.

Structural analysis of the specific glycosylated residues of the recombinant proteins produced in transgenic animal's milk has been performed in a handful number of cases as outlined in the above sections. To perform that kind of detailed examinations consultancy with an expert glycobiologist is recommended. For evaluation the basic glycosylation pattern of a milk born recombinant protein, deglycosylation with N-glycosidase F and Endo H is suggested as per the manufacturer's instructions.

6. Conclusions

Multicellular organisms, plants, and animals are the most suitable for producing large amount of recombinant proteins, which are biologically active, properly modified posttranslationally and their production cost is low. Production of recombinant proteins in the

milk of transgenic livestock animals is in the most advanced stage. Mammary-gland-specific vectors and efficient methods to create transgenic animals have been developed. Methods to separate and characterize the recombinant proteins in milk are also available, but should be adapted on case-by-case bases to the expressed proteins. Emerging data reveal species-specific differences in the glycosylation pattern of the milk-born recombinant proteins, which might have important consequences in the choice of livestock species for biopharming (1). Pharmaceutical proteins produced for replacement therapy should consistently recapitulate the complex posttranslational modifications of the endogenous human protein. However, variations in the percentage of sialylated glycans and core fucosylation were reported to depend on the level of transgene expression and the stage of lactation beyond the species specific variations (13). Those interanimal and daily variations should be further explored and the availability of production herds of different species facilitates the progress.

7. Notes

1. BAC libraries available on cost recovery bases can be found at http://bacpac.chori.org/.
2. A kit for homologous recombination can be obtained at http://www.genebridges.com/.
3. The Transgenic Core facility of the University of Michigan has developed a robust protocol for purification of BAC DNA using the NucleoBond BAC 100 Kit (Clontech). This protocol can be found at http://www.med.umich.edu/tamc/BACnucleo.html.
4. An alternative option to eliminate contaminating *E. coli* DNA is using the Qiagen Large-Construct Kit. This product is designed for the preparation of up to 50 μg of ultrapure BAC DNA.
5. A detailed description of a robust protocol to purify BAC DNA via gel exclusion chromatography can we found at the University of Michigan Transgenic Core Web page: http://www.med.umich.edu/tamc/BACcol.html.

Acknowledgments

This work was supported by the grant OM-00118/2008.

References

1. Houdebine, L. M. (2009) Production of pharmaceutical proteins by transgenic animals. *Compar Immunol. Microbiol Infect. Dis.* **32**, 107–121.
2. Fan, J., and Watanabe, T. (2003) Transgenic rabbits as therapeutic protein bioreactors and human disease models. *Pharmacol. Therap.* **99**, 261–282.
3. Robl, J. M., Wang, Z., Kasinathan, P., et al., (2007) Transgenic animal production and animal biotechnology. *Theriogenology* **67**, 127–133.
4. Sumer, H., Liu, J., Tat, P., et al., (2009) Somatic cell nuclear transfer: pros and cons, *J. Stem Cells* **4**, 85–93.
5. Wall, R. J. (2001) Pronuclear microinjection, *Cloning Stem Cells* **3**, 209–220.
6. van Veen, H. A., Geerts, M. E., van Berkel, P. H., et al., (2002) Analytical cation-exchange chromatography to assess the identity, purity, and N-terminal integrity of human lactoferrin, *Anal. Biochem.* **309**, 60–66.
7. Koles, K., van Berkel, P. H., Pieper, F. R., et al.,(2004) N- and O-glycans of recombinant human C1 inhibitor expressed in the milk of transgenic rabbits. *Glycobiology* **14**, 51–64.
8. Jongen, S. P., Gerwig, G. J., Leeflang, B. R., et al., (2007) N-glycans of recombinant human acid alpha-glucosidase expressed in the milk of transgenic rabbits. *Glycobiology* **17**, 600–619.
9. Gil, G. C., Velander, W. H., and Van Cott, K. E. (2008) Analysis of the N-glycans of recombinant human Factor IX purified from transgenic pig milk. *Glycobiology* **18**, 526–539.
10. Fujiwara, Y., Miwa, M., Takahashi, R., et al., (1999) High-level expressing YAC vector for transgenic animal bioreactors. *Molec. Reprod. Develop.* **52**, 414–420.
11. Fujiwara, Y., Takahashi, R. I., Miwa, M., et al., (1999) Analysis of control elements for position-independent expression of human alpha-lactalbumin YAC. *Molec. Reprod. Develop.* **54**, 17–23.
12. Rival-Gervier, S., Viglietta, C., Maeder, C., et al., (2002) Position-independent and tissue-specific expression of porcine whey acidic protein gene from a bacterial artificial chromosome in transgenic mice. *Molec. Reprod. Develop.* **63**, 161–167.
13. Yang, P., Wang, J., Gong, G., et al., (2008) Cattle mammary bioreactor generated by a novel procedure of transgenic cloning for large-scale production of functional human lactoferrin, *PloS One* **3**, e3453.
14. Baranyi, M., Hiripi, L., Szabo, L., et al.,(2007) Isolation and some effects of functional, low-phenylalanine kappa-casein expressed in the milk of transgenic rabbits. *J. Biotechnol.* **128**, 383–392.
15. Raju, T. S., Briggs, J. B., Borge, S. M., et al., (2000) Species-specific variation in glycosylation of IgG: evidence for the species-specific sialylation and branch-specific galactosylation and importance for engineering recombinant glycoprotein therapeutics. *Glycobiology* **10**, 477–486.
16. Edmunds, T., Van Patten, S. M., Pollock, J., et al., (1998) Transgenically produced human antithrombin: structural and functional comparison to human plasma-derived antithrombin, *Blood* **91**, 4561–4571.
17. Montesino, R., Toledo, J. R., Sanchez, O., et al., (2008) Monosialylated biantennary N-glycoforms containing GalNAc-GlcNAc antennae predominate when human EPO is expressed in goat milk. *Arch. Biochem. Biophys.* **470**, 163–175.
18. Maga, E. A., and Murray, J. D. (1995) Mammary gland expression of transgenes and the potential for altering the properties of milk, *Biotechnology* **13**, 1452–1457.
19. Yen, C. H., Yang, C. K., Chen, I. C., et al., (2008) Expression of recombinant Hirudin in transgenic mice milk driven by the goat beta-casein promoter. *Biotechnol. J.* **3**, 1067–1077.
20. Lenasi, T., Kokalj-Vokac, N., Narat, M., et al., (2005) Functional study of the equine beta-casein and kappa-casein gene promoters, *J. Dairy Res.* 72 *Spec No*, 34–43.
21. Bhure, S., and Sharma, B. (2007) Cloning and characterization of ovine alphaS1-casein gene promoter: a transfection study in rat mammary gland cell line. *DNA Seq.* **18**, 39–46.
22. Lois, C., Hong, E. J., Pease, S., et al., (2002) Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science* **295**, 868–872.
23. Bijvoet, A. G., Van Hirtum, H., Kroos, M. A., et al., (1999) Human acid alpha-glucosidase from rabbit milk has therapeutic effect in mice with glycogen storage disease type II. *Human Molec. Gen.* **8,** 2145–2153.
24. Hiripi, L., Baranyi, M., Szabo, L., et al., (2000) Effect of rabbit kappa-casein expression on the properties of milk from transgenic mice. *J. Dairy Res.* **67**, 541–550.
25. Chung, J. H., Whiteley, M., and Felsenfeld, G. (1993) A 5′ element of the chicken beta-globin

domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila. *Cell* **74**, 505–514.

26. Filippova, G. N. (2008) Genetics and epigenetics of the multifunctional protein CTCF. *Curr. Top. Develop. Biol.* **80**, 337–360.
27. West, A. G., Huang, S., Gaszner, M., et al., (2004) Recruitment of histone modifications by USF proteins at a vertebrate barrier element. *Molec. Cell* **16**, 453–463.
28. Huang, S., Li, X., Yusufzai, T. M., et al., (2007) USF1 recruits histone modification complexes and is critical for maintenance of a chromatin barrier, *Molec. Cell. Biol.*27, 7991–8002.
29. Mutskov, V. J., Farrell, C. M., Wade, P. A., et al., (2002) The barrier function of an insulator couples high histone acetylation levels with specific protection of promoter DNA from methylation. *Genes Develop.* **16,** 1540–1554.
30. Wang, Y., DeMayo, F. J., Tsai, S. Y., et al., (1997) Ligand-inducible and liver-specific target gene expression in transgenic mice. *Nature Biotechnol.* **15**, 239–243.
31. Potts, W., Tucker, D., Wood, H., et al., (2000) Chicken beta-globin 5′HS4 insulators function to reduce variability in transgenic founder mice. *Biochem. Biophys. Res. Comm.* **273**, 1015–1018.
32. Cremer, T., Kreth, G., Koester, H., et al., (2000) Chromosome territories, interchromatin domain compartment, and nuclear matrix: an integrated view of the functional nuclear architecture. *Crit. Rev. Eukaryiot. Gene Exp.* **10**, 179–212.
33. Phi-Van, L., and Stratling, W. H. (1996) Dissection of the ability of the chicken lysozyme gene 5′ matrix attachment region to stimulate transgene expression and to dampen position effects. *Biochemistry* **35**, 10735–10742.
34. Goetze, S., Baer, A., Winkelmann, S., et al., (2005) Performance of genomic bordering elements at predefined genomic loci. *Molec. Cell. Biol.* **25**, 2260–2272.
35. Evans, K., Ott, S., Hansen, A., et al., (2007) A comparative study of S/MAR prediction tools, *BMC Bioinformatics* **8**, 71.
36. Lunyak, V. V., Prefontaine, G. G., Nunez, E., et al., (2007) Developmentally regulated activation of a SINE B2 repeat as a domain boundary in organogenesis, *Science* **317**, 248–251.
37. Willoughby, D. A., Vilalta, A., and Oshima, R. G. (2000) An Alu element from the K18 gene confers position-independent expression in transgenic mice, *The J. Biol.Chem.* **275**, 759–768.
38. Li, Q., Peterson, K. R., Fang, X., et al., (2002) Locus control regions, *Blood* **100**, 3077–3086.
39. Grosveld, F., van Assendelft, G. B., Greaves, D. R., et al., (1987) Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell* **51**, 975–985.
40. Bessa, J., Tena, J. J., de la Calle-Mustienes, E., et al., (2009) Zebrafish enhancer detection (ZED) vector: a new tool to facilitate transgenesis and the functional analysis of cis-regulatory regions in zebrafish, *Dev. Dyn.* **238**, 2409–2417.
41. Montoliu, L., Roy, R., Regales, L., et al., (2009) Design of vectors for transgene expression: The use of genomic comparative approaches, *Comp. Immunol. Microbiol. Infect. Dis.* **32**, 81–90.
42. Woolfe, A., Goodson, M., Goode, D. K., et al., (2005) Highly conserved non-coding sequences are associated with vertebrate development, *PLoS Biology* **3**, e7.
43. de la Calle-Mustienes, E., Feijoo, C. G., Manzanares, M., et al., (2005) A functional survey of the enhancer activity of conserved non-coding sequences from vertebrate Iroquois cluster gene deserts. *Genome Res.* **15**, 1061–1072.
44. Gerencser, A., Barta, E., Boa, S., et al., (2002) Comparative analysis on the structural features of the 5′ flanking region of kappa-casein genes from six different species. *Genet. Sel. Evol.* **34,** 117–128.
45. Yang, X. W., Model, P., and Heintz, N. (1997) Homologous recombination based modification in Escherichia coli and germline transmission in transgenic mice of a bacterial artificial chromosome, *Nature Biotechnol.* **15**, 859–865.
46. Sparwasser, T., Gong, S., Li, J. Y., et al., (2004) General method for the modification of different BAC types and the rapid generation of BAC transgenic mice, *Genesis* **38**, 39–50.
47. Muyrers, J. P., Zhang, Y., Testa, G., et al., (1999) Rapid modification of bacterial artificial chromosomes by ET-recombination. *Nucl. Acids Res.* **27**, 1555–1557.
48. Van Keuren, M. L., Gavrilina, G. B., Filipiak, W. E., et al., (2009) Generating transgenic mice from bacterial artificial chromosomes: transgenesis efficiency, integration and expression outcomes. *Transgenic Res.* **18**, 769–785.
49. DiTullio, P., Cheng, S. H., Marshall, J., et al., (1992) Production of cystic fibrosis transmembrane conductance regulator in the milk of transgenic mice. *Bio/technology* **10**, 74–77.
50. Bodrogi, L., Brands, R., Raaben, W., et al., (2006) High level expression of tissue-nonspecific alkaline phosphatase in the milk of transgenic rabbits. *Transgenic Res.* **15**, 627–636.

51. Ahn, Y. S., and Snow, L. D. (1993) Selective extraction of alkaline phosphatase and 5′-nucleotidase from milk fat globule membranes by a single phase n-butanol procedure. *Prep. Biochem.* **23**, 409–419.
52. Baldassarre, H., Hockley, D. K., Dore, M., et al., (2008) Lactation performance of transgenic goats expressing recombinant human butyryl-cholinesterase in the milk. *Transgenic Res.* **17**, 73–84.
53. Van Cott, K. E., Lubon, H., Russell, C. G., et al., (1997) Phenotypic and genotypic stability of multiple lines of transgenic pigs expressing recombinant human protein C. *Transgenic Res.* **6**, 203–212.
54. Hiripi, L., Makovics, F., Halter, R., et al., (2003) Expression of active human blood clotting factor VIII in mammary gland of transgenic rabbits. *DNA Cell Biol.* **22**, 41–45.
55. Al-Ghobashy, M. A., Williams, M. A., Brophy, B., et al., (2009) On-line casein micelle disruption for downstream purification of recombinant human myelin basic protein produced in the milk of transgenic cows. *J. Chromat.* **877**, 1667–1677.
56. Wang, J., Yang, P., Tang, B., et al., (2008) Expression and characterization of bioactive recombinant human alpha-lactalbumin in the milk of transgenic cloned cows. *J. Dairy Sci.* **91**, 4466–4476.

Index

A

Adaptation....16, 19, 23, 220, 292, 390, 453, 454, 457, 459
Adjuvant(s)....75–77, 274, 536
Affinity chromatography....116
Aggregation index (AI)....146, 148, 150, 152, 153
Agrobacterium
  A. rhizogenes....502, 504, 509
  A. tumefaciens....472, 485, 487, 490–492, 498, 502, 504, 507, 509, 542, 549
Agrobacterium-mediated transformation....471, 488, 490, 498
Allelic exchange....220, 221
Amino acid modification....185
Ammonium sulfate protein precipitation....409, 549
Amplification....192
Animal cells....4, 11, 272, 274
Antibody
  fragment....100, 345, 349
  monoclonal....67, 70, 71, 74–76, 78–82, 84, 86, 93, 95–98, 100, 112, 116, 121, 125, 126, 330, 364, 368, 384, 388, 420, 423, 428, 430, 545, 553, 556, 598, 599, 603
  polyclonal....67–71, 73–85, 87, 95–98, 100, 101, 116, 126, 159, 506, 515, 517, 545, 599, 603
Antigen....12, 69, 71–77, 80, 82, 90, 94–97, 100, 101, 157, 158, 271, 278, 360, 398
Antioxidant(s)....7, 11, 39, 40, 43, 398, 560
Antisense
  expression....284
  RNA....178, 179, 196, 284
Aptamer....381–390
Aqueous two phase system (ATPS)....528, 529, 531–533
Arabidopsis (Arabidopsis thaliana)....596
Aspergillus nidulans....46, 393–415

B

Bacillus....5, 169, 186, 284, 330, 345
Bacterial artificial chromosomes (BACs)....55, 581–593, 610, 611, 616, 631–632, 634–636, 638
Bacterial host....13, 169, 194, 331, 412
Baculovirus....56, 151, 609–626
Basement membrane....318
Bioactivity....68, 495, 496, 542, 554, 557, 558
Biolistics....471, 474
Bioreactor....26, 263, 283, 288, 291, 292, 295, 309, 347, 523, 527, 602, 631
Biotechnology
  industrial....282, 345
  plant....470, 471, 477, 478

C

Carrier protein....73
Cassette exchange....582, 583, 585–588, 591
Cell
  counting....365, 606, 611, 614, 615
  death....11, 14, 267, 348, 365
  lines....12, 80, 158, 324, 345, 490, 492, 493, 498, 581–593, 595, 597, 598, 600–603, 615, 616, 622, 633
  lysate....120, 122, 124, 139, 213–216, 247, 264, 277, 322, 448, 565, 604, 618
  lysis....17, 112–115, 122, 124, 150, 187, 188, 238, 244, 286, 290, 435, 446
  viability....17, 365, 498, 606, 615, 616, 623, 626
Cell-free protein synthesis (CFPS)....24, 565–578
Cellular location....157–158
Chaperone(s)
  chemical....7–9, 12–15, 22, 24, 150
  molecular....5, 9, 11, 12, 21–22, 25, 27, 596, 597, 600, 601, 603, 606
Chinese Hamster Ovary (CHO) cells....345, 593, 595–606
Chloroplast....474, 475
Chromatin
  elements....56, 58, 582, 633
  positional effects....582
  reassembly....60, 61
Chromohalobacter....168–176, 179, 181, 184, 187, 188, 192

Argelia Lorence (ed.), *Recombinant Gene Expression: Reviews and Protocols, Third Edition*,
Methods in Molecular Biology, vol. 824, DOI 10.1007/978-1-61779-433-9_7, © Springer Science+Business Media, LLC 2012

Chromosomal integration........ 156–157, 382, 383, 385, 433–450, 458, 460, 461
Cloning........ 111–113, 117–118, 126, 127, 136–137, 156, 169, 170, 174–182, 205, 208, 209, 217, 222, 225, 226, 232, 236, 240–241, 252–254, 257–258, 267, 309, 320, 321, 330, 335, 336, 393–415, 437, 449, 485, 488–490, 504, 507, 509, 536, 570, 583, 616, 630
Codon
  bias........ 339, 497
  optimization 474
  usage........ 16, 22, 339, 462, 474
Co-expression........ 4, 9, 25, 110, 122–123, 146, 204, 235–248, 475, 527, 597, 600
Cofactors........ 7, 8, 14, 189, 252, 457, 557
Cold-adapted bacteria 204
Competent cells........ 135, 142, 161, 236, 241, 258, 366, 405–406, 425, 434, 490, 504, 507
Complementation........ 179, 183, 283, 335, 412, 420, 423, 424, 458
Conditional gene expression........ 381, 382
Conjugation........ 73, 170, 181–182, 184, 186, 189, 196, 204–206, 212, 220, 222, 226, 252, 259–261, 271
Constitutive
  expression 157, 252, 253, 309, 310, 320
  promoter(s)........ 156, 161, 205, 254, 334, 339, 372, 373, 383, 472, 496
Continuous-exchange cell-free (CECF) system........ 570–573, 575, 576
Copy number........ 16, 55, 56, 156, 158, 181, 189, 276, 283, 288–290, 295, 336, 339, 367, 368, 434, 455, 474, 476, 498, 510, 583, 590, 631, 634
Corn (maize, Zea mays) 472, 473, 475, 502
Counter-selectable marker........ 221, 227
Cre/lox........ 371–379
Cryptic transcription 59, 60
Culture
  batch........ 287
  continuous 211, 309, 324, 348, 351, 365
  fed-batch 287, 348
  media........ 38, 41, 138, 181, 253, 399, 410, 493–496, 518, 519, 623
Cytoplasm........ 5, 7, 13, 21, 26, 131, 168, 187, 341, 475

D

Deglycosylation 103, 537, 541, 556–559, 637
Deletion mutant 186, 227–230
Dialysis........ 59, 114, 124, 134, 140, 147, 150, 151
DNA
  amplification........ 191, 196, 209, 210, 212, 225, 257, 320, 321, 386, 402, 412, 436, 624
  extraction........ 135, 229, 318, 366–368, 374, 376, 404, 485, 586
  genomic 58, 139, 196, 205, 208–210, 225, 226, 229, 230, 257, 282, 366–367, 394, 399, 401, 402, 449, 472, 510–512
  linear........ 278, 280, 284, 321, 366, 444, 446, 448, 590, 592, 602, 636
  methylation........ 54
  plasmid (pDNA) 118, 127, 137, 156, 158, 159, 161, 162, 173, 175, 193, 194, 241, 242, 256, 258–260, 271–296, 310, 318, 320, 321, 367, 374, 375, 403, 434, 444, 446, 611, 617
  purification 293
  quantification........ 190
  synthetic 58, 133
  vaccines........ 273, 274, 283, 284, 289, 292, 294, 296
DnaK........ 9, 10, 21, 22, 24, 110, 112, 113, 116, 118, 122–123, 125, 127, 199
DNase........ 56, 146, 148, 151, 194, 365, 396, 400
Dot-blot........ 394, 406–407, 413, 490, 492, 498
Downstream purification........ 493

E

E2 epitope tag 420
Electrocompetent cells........ 142, 162–163, 585–587, 589, 591
Electroporation........ 134, 137–138, 161, 163, 171, 182, 194, 274, 310, 320, 321, 323, 335, 337, 338, 366, 397, 405–406, 584–586, 588, 589, 593
Electrotransfer (electroblotting) 38, 84, 89, 538, 545, 546
Embryo........ 475, 567, 569, 577
Endomembrane system 542, 557
Endophyte........ 371–379
Endoplasmic reticulum (ER)........ 5, 14, 15, 342, 343, 475, 488, 489, 528, 530, 542, 596, 597
Endotoxin
  contamination........ 150
  removal 153, 547, 548, 550, 561, 563
  test........ 539, 549–550
Enhancer........ 52–56, 58, 168, 496, 502, 503, 506–508, 515, 530, 539, 546, 633, 634
Enterokinase (EK)........ 132, 135, 136, 139–142
Enzymatic cleavage 132, 135, 136, 140–141, 341
Enzyme assay 554, 557
Enzyme-linked immunosorbent assay (ELISA)
  direct........ 77, 92–95, 103, 555–557
  sandwich........ 69, 80, 90, 95–97, 103, 494, 506, 517–520, 524, 555
Epichloë........ 374, 377
Epitope tagging........ 420
Escherichia coli (E. coli)........ 4, 8, 71, 109–128, 204, 222, 235, 252, 276, 286, 291, 310, 330, 374, 383, 397, 504, 507, 566, 583, 596, 632
Essential genes........ 156, 220, 284, 381–390

Ethanol production 333, 457, 460, 462, 463
Eukaryotes.............................. 4, 59, 60, 309, 311, 343, 557, 585, 591, 592, 634
Eukaryotic
proteins.. 343, 394, 566
systems..14, 151, 331, 360, 566
Export signal 131–142
Expression
heterologous genes............................... 51–61, 156, 235, 254, 268, 308, 317, 333, 359–369, 433–450
heterologous proteins....................... 155–165, 169, 205, 251, 309, 334
recombinant proteins........................... 27, 56, 109, 119, 124, 126, 312, 313, 317, 363–365, 537, 542–547, 583, 611–612, 618, 629–638
vector....................................16, 56, 112, 114, 116–118, 126, 127, 132, 135–137, 169, 170, 175–177, 179–181, 204, 205, 208–209, 212, 217, 253, 254, 309–310, 318, 321, 324, 336, 344, 383, 396, 411, 420, 489, 566–567, 570, 581–589, 597, 631, 633
Extremophiles 168

F

Fermentation 155, 330–334, 339, 344–349, 451–455, 457, 460, 462
Filamentous fungi..330, 342
Flagellum-specific export 131–142
Foldases.. ..4, 146
Folding promoting agents...3–27
Freedom to operate (FTO) 477, 478
Freezing...... ....78, 126, 168, 188, 211, 247
Functional genomics..565
Fungal transformation 378, 420–426, 448
Fungi...................... . 313, 330, 333, 342, 374, 377, 448, 596
Fusion
expression 116, 117, 122
protein(s) 4, 10, 37, 109–128, 132, 133, 138–140, 142, 145, 420, 475, 487, 488, 528, 530–534

G

β-Galactosidase (LacZ)....................82, 187, 190, 205, 208, 212–217, 340, 617, 618, 622
β-Glucuronidase (GUS)..502, 557
Gene
inactivation 221
silencing.. 54, 473, 498, 527
targeting 278, 336
therapy..................................55, 56, 272, 273, 275–285
transfer.. 169, 310, 458
Generally recognized as safe (GRAS) 331, 332, 452
Genetic
exchange 171, 173, 182
transformation 473, 502
Genome sequencing 204, 205, 220, 313, 335, 337, 338, 342–344, 394
Genotype... ....286, 287, 296, 373, 376, 420, 435, 517, 565
Gland specific transgene expression 631–634
Glutathione (GSH)............................5, 11, 13, 18, 20, 236, 238, 247
Glycoproteins. 12, 98, 308, 311, 318, 334, 343, 484, 487, 488, 492–496, 498, 499, 536, 541, 542, 557, 558, 561, 563, 597, 631
Glycosylation
N-glycosilation 308, 311, 318, 337, 343, 557
O-glycosylation 343, 344, 484, 497
Golgi..341–343
Gram-negative bacteria 174, 179, 181, 184, 186, 549
Gram-positive bacteria 221
Green fluorescent protein (GFP)
enhanced GFP... 485, 598, 603
GroEL.....................9, 21–24, 110, 112, 113, 116, 118, 119, 122–123, 127

H

Hairy roots 501–525
Halomonadaceae 167–197
Halomonas.................168–176, 179, 181, 182, 184, 186–188
Halophilic bacteria 167–169, 172, 173, 175–179, 188, 192
Heat-shock protein(s) (HSPs)
Hsp40.. .... 596, 597, 599
Hsp70.. ....110, 596, 597, 599, 600
Hsp104 .... 596, 597, 599, 600
Herbicide resistance...168, 483
Heterologous proteins 110, 155–165, 169, 203, 205, 251, 265–266, 268, 272, 292, 308–310, 330–334, 336, 339, 341, 342, 344–346, 349, 351, 360, 394, 417, 475, 529
Hexa-histidine tag (HIS-tag)........................ 12, 13, 69, 71, 92–94, 96, 97, 115–116, 123–125, 132, 236, 238, 239, 243, 244, 246, 394, 409, 412, 414, 536, 537, 539, 541, 542, 545, 547, 549, 555, 556, 561, 598
High-copy number..................158, 181, 283, 288, 290, 336
High-throughput (HT)
analysis...4
expression 609–626
Homologous recombination...................185, 221, 278, 309, 310, 385, 423, 438, 440, 444, 445, 458, 582, 583, 585, 588–589, 592, 609, 610, 616, 632, 635, 638
Hormone...........................4, 15, 41, 55, 277, 330, 540, 634
Human antithrombin III..629
Hydrophobin 527–534
Hydroxyproline-O-glycosylation...................................484
Hygromycin..310, 373
Hyp-Glyco technology..484, 493

I

Immobilized-metal affinity chromatography (IMAC) 116, 127, 151
Immunoblotting 68, 69, 80, 83, 85, 161, 425, 542, 545, 553, 561
Immunodetection 38, 99, 100, 253, 265–266, 364, 394, 406–407, 413
Inclusion bodies 6, 9, 21, 71, 110, 116, 122, 123, 132, 142, 204, 247, 264, 282, 331, 341
Infection 182, 183, 277, 285, 307, 360, 473, 474, 509, 614
Insect cells 56, 151, 311–313, 558, 609–626
Insertion 183–185, 220, 226–228, 258, 260, 283, 285–287, 309, 378, 382, 383, 385–389, 411, 412, 438, 459, 509–510, 512, 584, 591, 635
Insertional mutagenesis 184, 185
Insulator(s) 53–56, 582, 631–634
Integration
  site 52, 53, 55, 386, 390, 434, 582, 583, 631
  vector 434, 437
Intellectual property (IP) 157, 159, 470, 471, 476–478
Interferon α2 485
Interleukin-12 .... 66, 536, 551
Interleukin(s) 66, 536, 551
In vitro transcription 196, 570–572
In vitro translation 27, 570, 572
Isopropyl-β-D-thiogalacto pyranoside (IPTG) 111, 119, 121, 124, 126, 142, 146, 147, 237, 238, 242, 243

K

Knockout mutant 227

L

Lactococcus lactis (L. lactis) 155–165
LacZ. *See* β-galactosidase (LacZ)
Laminin 311, 318, 321
Leishmania tarentolae (L. tarentolae) 307–313, 317–324
Ligase 112, 117, 118, 133, 135, 206, 210, 224, 225, 236, 241, 256–258, 397, 403, 462, 485, 567
Ligation 117, 118, 135, 185, 210, 241, 257, 259, 284, 397, 403, 583, 616
Light scattering 148, 152, 282
Lignocellulosic biomass 452, 460, 463
Linking number (Lk) 278, 279
Locus control regions (LCRs) 634
Loss-of-function mutation
Luciferase (Luc) 24, 26, 56, 283, 566, 568, 573, 598–605

M

Magnetic beads 99, 146, 151
Malate induction 217
Mammalian cell(s) 12, 52, 56, 98, 308, 309, 312, 313, 317, 318, 324, 330, 346, 536, 581–593, 595, 614
Mammalian-type post-translational modification 313
Marker excision 372, 374, 377, 461
Matrix-assisted laser-desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) 38
Matrix attachment regions (MARs) 55, 582, 633
Membrane 4, 38, 70, 112, 131, 150, 157, 168, 205, 235, 252, 293, 318, 362, 384, 397, 423, 462, 475, 505, 529, 538, 565, 599, 621
Metabolic engineering 294–296, 330, 352, 470, 501, 581
Metabolic burden 295, 339, 345
Metal 14, 37, 116, 150, 151, 175, 179, 252, 271, 275, 394, 408, 409, 414, 545, 621
Metal chelation affinity chromatography 37
Mice 55, 70, 79, 81, 157, 158, 285, 496, 582, 583, 632–634
Milk 84, 87, 103, 116, 164, 238, 239, 246, 265, 322, 362, 385, 388, 397, 406, 423, 428, 604, 613, 621, 629–638
Minimal medium 9, 134, 138, 140, 171, 187, 189, 211–215, 253, 261, 283, 289, 360, 361, 394, 395, 399, 422
Mitochondria 13, 24, 307, 475
Molecular farming 470, 483
Mutagenesis 170, 183–187, 194, 195, 283, 378, 452, 459, 485, 490

N

Neomycin phosphotransferase II (nptII) 489, 507, 508
Neotyphodium coenophialum 373, 378
N-glycan 311, 338, 343, 344, 596
Nicotiana benthamiana 77, 527, 530, 542
Northern blot 191, 192, 363, 364, 418, 505, 512–515, 520, 521
Nucleosome 52, 57–61
  +1 nucleosome 60, 61
  positioning 51–61

O

Ocs element 502
Open reading frame (ORF) 383, 385, 438
ORF. *See* Open reading frame (ORF)
Origin of replication 12, 116, 178, 204, 205, 208, 253, 254, 271, 276, 283–285, 339, 383, 429, 591, 592
Osmolytes 7–9, 15, 19, 22–25, 168
OTR. *See* Oxygen transfer rate (OTR)

Oxidoreductase......596, 597
Oxygen transfer rate (OTR)......292, 295

P

Packing......44, 528
PAGE. *See* Polyacrylamide Gel Electrophoresis (PAGE)
Patents......476–478, 633
Pathway engineering......344
PCR. *See* Polymerase chain reaction (PCR)
Pentose fermentation......459
Peptide mass fingerprint......45
Periplasm...... 5, 13–15, 17, 21, 26, 295, 342
Periplasmic space......5, 331, 342
Phage...... 6, 25, 116, 182, 183, 194, 280, 285, 292, 310, 635
Photosynthetic bacteria......251–268
Pichia pastoris......10, 330, 360, 361, 366, 393–415
Plant-based
  bioproduction......483–499
Plant cell
  culture......486, 493, 499
  wall......394
Plant variety protection (PVP)......476
Plasmid
  concatamers......278, 279
  high copy number...... 158, 283, 288, 290
  loss......278, 292, 293
  low copy number......189
  supercoiling......279, 287, 289
  Ti......502
Polyacrylamide Gel Electrophoresis (PAGE)......68–69, 71, 115, 239, 245, 264–265, 320, 362, 363, 422, 427–428, 430, 486–487, 537, 543–545, 573–574, 612, 618–620
Polyadenilation..277, 308, 411, 472, 508, 582, 585, 589, 592
Polyhistidine tag. *See* Hexa-histidine tag
Polylinker......208, 224, 225
Polymerase chain reaction (PCR)...... 318, 320, 402, 447
Positional effect(s)......582
Post-transcriptional gene silencing......473
Post-translational modifications......331, 343
Potato *(Solanum tuberosum)*......473
Prokaryotes......4, 167, 585, 591, 592
Promoter
  artificial......583
  chimeric......58–60, 502, 503
  constitutive......156, 161, 205, 254, 334, 339, 372, 373, 383, 472, 496
  heat-shock......418
  inducible......156, 161, 252, 339, 360, 361, 364, 381, 475, 476, 497
  osmoregulated......187
  35S...... 472, 485, 489, 496, 502, 508, 530
  super promoter......501–525
  synthetic......224, 225
  temperature induced......418
  yeast...... 61, 335, 339, 381
Protease(s)...... 14, 110, 112, 114, 126, 151, 168, 231, 255, 263, 321, 330, 340, 342, 344, 348, 475, 551, 604, 606
Protein
  accumulation...... 110, 263–266, 475–476
  aggregation......8, 10, 11, 595, 596
  bodies......528
  complexes...... 11, 127, 237–239, 242
  expression...... 27, 56, 109, 110, 117–122, 124, 126, 132–134, 138–140, 147, 148, 155–165, 217, 235, 275, 308–310, 312, 313, 317, 320, 322–323, 334, 345, 363–364, 394, 406–411, 413, 417–430, 470, 474, 517, 524, 537, 541–547, 570, 583, 611–612, 618, 619, 622, 629–638
  folding......3–5, 7, 12, 13, 15, 24, 26, 27, 110, 116, 252, 311, 312, 331, 341, 342, 597
  fusion...... 4, 10, 37, 109–127, 132, 133, 138–140, 142, 145, 420, 475, 487, 488, 528, 530–534
  identification......38, 49
  N-terminal sequencing...... 103, 537, 540, 542, 549–551, 553–554, 562, 563
  over-expression......21
  product validation......542–547
  protein interactions...... 237, 239, 242, 248
  purification......112–115
  quality......153, 341, 393
  quantity......78, 341
  secretion......231, 342
  solubility......112–115, 247
  stability...... 7, 9, 25, 26, 596
  yield...... 268, 394, 475, 484, 485, 496, 527, 554–557, 566, 575
Protoplast......368, 371, 373–376, 378, 502
Protozoan parasite......307–313, 318
Pseudoalteromonas haloplanktis...... 203–217, 219–232, 345
Pull-down assay...... 238, 239, 243–245
PVP. *See* Plant variety protection (PVP)

Q

qPCR. *See* Real time PCR (qPCR)

R

Real time PCR (qPCR)......192, 367, 613
Recombinant
  antibody......4, 25, 65–104, 330, 360, 637
  DNA...... 330, 352, 420, 472

gene expression..55, 205, 634
heterologous 55
homologous..............................185, 221, 278, 309, 310, 385, 423, 438, 440, 444, 445, 458, 582, 583, 585, 588–589, 592, 609, 610, 616, 632, 635
Recombinant (*Continued*)
protein(s) 3–27, 37–50, 55, 66, 109–127, 131–142, 169, 203–17, 220, 238, 265, 272, 307–313, 317, 330, 360, 407, 469–480, 483–499, 502, 527–563, 565–578, 581, 595–606, 611, 629–638
protein detection...121
Redox
agents..8
state...13
Refolding.......................4, 7, 10, 11, 14, 21, 22, 71, 110, 596
Regulated gene expression
system 208–212, 381–390, 394
Regulatory affairs...478
Renilla luciferase... 566, 568, 573
Replication 12, 116, 156, 158, 177–180, 183, 204, 205, 208, 253, 254, 271, 275–276, 278, 282–286, 289, 290, 336, 339, 383, 412, 429, 437, 473–474, 591, 592, 626
Reporter gene(s) 26, 170, 185, 187–189, 205, 283, 360, 502, 503, 520
Restriction enzyme........................112, 118, 126, 127, 141, 193, 206, 209, 224, 236, 240, 256, 257, 318, 321, 397, 401, 404, 412, 435, 440, 444–446, 485, 488, 504, 510, 567, 583
Retrovirus.. ....272, 273
Reverse transcription 192, 320
Rhodobacter capsulatus 251–268
Ribosome..267, 566, 591
Riboswitch...382
RNA
mRNA..........................11, 24, 178, 189–192, 196, 308, 309, 363, 382, 394, 399–401, 411, 418, 472–475, 513–515, 520–522, 525, 565, 572, 577, 597
purification 194
rRNA.................................168, 309, 320, 364, 387, 388
siRNA
tRNA...................................26, 54, 221, 224, 225, 227, 283, 566, 567, 572, 573
Rosa26 locus 55, 583–585
RT-PCR.................. 174, 189, 191, 192, 194, 368, 396, 400

S

Saccharomyces cerevisiae (*S. cerevisiae*) 7, 11, 15, 54, 59, 330–337, 340–343, 345, 351, 359–364, 384, 385, 389, 393, 394, 406, 411, 418, 429, 451–463, 596
Salmonella typhimurium (*S. typhimurium*) 132, 133, 135, 137
Schizosaccharomyces pombe (*S. pombe*)................ 15, 331, 434, 446–447
Secreted protein....... 132, 140, 141, 158–160, 331, 342, 345, 349, 350, 360, 484, 485, 493, 519, 522, 618
Secretion.................................6, 10, 15, 131–142, 157–159, 230, 231, 294, 309, 310, 330, 333, 341–345, 349, 351, 352, 359, 360, 394, 411, 483–499, 502, 522, 557, 625, 626
Selectable marker...................................277, 282, 283, 310, 371–379, 412, 429, 459, 472, 507
Signal sequence..................................... 139–141, 309–311, 319, 341, 401, 411, 475, 485, 489
Silencing (gene silencing)............................... 473, 498, 527
Silencing barriers 53, 54
Site-directed mutagenesis...............................185–187, 490
Size exclusion chromatography 146, 635
Sodium dodecyl polyacrylamide gel electrophoresis (SDS-PAGE) 37, 68, 112, 134, 149, 158, 207, 243, 264, 322, 363, 408, 425, 437, 493–494, 530, 538, 576, 598–599, 612
Solid medium 170, 172–174, 195, 223, 226–228, 256, 260–262, 365, 368, 429, 434, 436, 447, 449, 485, 486, 498, 567
Soluble aggregates 146, 152
Solutes.............................. 5, 7, 17, 168, 184, 193, 206, 223
Southern blot...................505, 510–512, 516, 588, 589, 591
Spodoptera frugiperda 610
Stable expression 530, 582
Stress response................. 4, 5, 8, 25, 26, 292, 419, 454, 597
Subcloning..................................... 136–137, 488–489, 571
Surface antigen of hepatitis B virus (HBsAg) 360
Surfactants..529
Systems biology...4, 345

T

Tagging...................................252, 383, 420, 424, 429, 433
T-DNA.. 498, 502, 508, 510
Terminal deoxynucleotidyl transferase mediated dUTP nick end labelling (TUNEL) assay..365–366
Tetrad analysis 386–387, 390
T7 expression
vectors...253
Therapeutic protein(s) 343, 360, 495, 501, 523, 630
Tissue specific expression 472, 476, 634
Tobacco *(Nicotiana tabacum)* 504, 509
Tobacco BY2 cells 488, 490–493
Transcription
factors... 52, 360, 635
initiation 189–191
start site 58, 191
termination 110, 189, 208, 210, 335, 411, 412, 474
Transcript stability..472–473, 476

Transduction.. 56, 97, 181–183, 194
Transfection......................................71, 272, 283, 284, 310, 320–321, 371–379, 581, 589, 593, 595, 597, 598, 600–603, 605, 611, 616, 617, 625
Transformation
  chloroplast 474
  DNA injection......................73, 274, 602, 630, 634–636
  electroporation..................................163, 171, 182, 274, 310, 321, 323, 335, 337, 338, 405–406, 585–586, 588, 593
  flower dip...................................471, 488, 490, 497, 498
  plastids...471
Transgene(s) 52, 53, 55, 56, 58, 59, 88, 94, 367, 472–476, 502, 503, 509, 510, 512–515, 542, 547, 582, 583, 630–635, 638
Transgenic
  animal(s)...............................56, 313, 630–633, 635–638
  plant(s)......................................313, 471, 473, 474, 476, 479, 480, 509, 513, 514, 516, 519
Transient
  expression 77, 371, 471, 502, 527, 530, 536, 542, 597, 599
  transfection 372–379, 389, 601, 603
Translation..................................24, 27, 225, 226, 247, 283, 309, 310, 336, 340, 346, 382, 389, 472, 474, 565–567, 570–573, 576, 577, 589, 597
Translational enhancer......................496, 503, 507, 508, 530
Transposon(s) 169, 172, 173, 183–185, 195, 196
Trehalose.................................5, 7, 8, 11, 12, 15, 17, 22, 150
Trichloroacetic acid (TCA) precipitation of protein 38, 41–42, 134, 163
Trypanosomatidae 307–310, 313
Trypsin digestion..............................41, 43–45, 48, 49, 415, 551, 552, 562
Tumor suppressor gene p53..360
Two-hybrid analysis 237, 241, 242
Type III secretion 132

U

Unfolded protein response (UPR)........... 342, 343, 346, 454
Untranslated region (UTR).....................308, 309, 382, 383, 389, 472, 473, 503, 514, 570

V

Vaccine.................. 4, 66, 155–158, 273, 274, 283–285, 289, 292, 294, 296, 330, 360, 470, 476, 536
Vector(s)
  adenovirus...272, 273
  binary................. 485, 488, 490–491, 496–498, 504, 507
  cloning...............................126, 174–182, 485, 488, 504
  expression 16, 56, 112, 114, 116–118, 126, 127, 132, 135–137, 169, 170, 175–177, 179–181, 204, 205, 208–212, 217, 253, 254, 309–310, 318, 321, 324, 336, 344, 383, 396, 411, 420, 489, 566–567, 570, 581–589, 597, 598, 631, 633
  shuttle.......................................178, 180–182, 189, 205, 394, 411, 429, 460
  suicide.........................178, 184–186, 195, 221, 224–231
  viral.......................................55, 272, 274, 277, 471–473
Viral promoter..277, 472

W

Western blot.................... 68–70, 80, 83, 101, 102, 537–539, 542, 543, 545–547, 550, 553–555, 561
Wheat germ extract.................566, 569–570, 572, 573, 577
Wound inducible promoter 520–523

X

Xylanase.. 42, 408, 411, 414
Xylosidase... 42, 46, 408, 410–411

Y

Yeast
  fission... 417–430, 433–450
  industrial strains 451–463
  laboratory strains 452, 453, 455, 458, 459
Yeast artificial chromosome (YAC) 55, 360, 582, 583, 631–632
Yeast promoter...................................61, 335, 339, 381, 418

Z

Zeocin...335, 338, 361, 366, 397, 403–406, 412, 413, 461
Zymobacter 168, 171, 176, 182