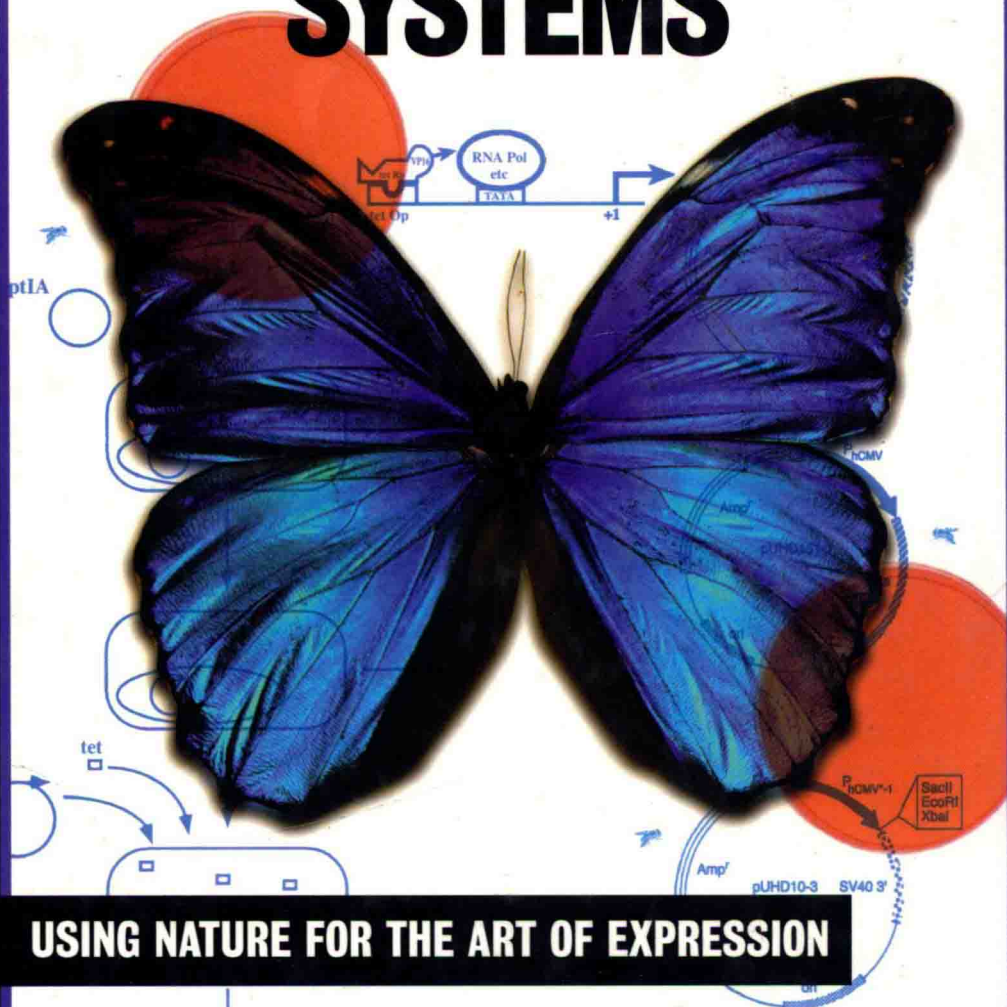


GENE EXPRESSION SYSTEMS



USING NATURE FOR THE ART OF EXPRESSION



EDITED BY
JOSEPH M. FERNANDEZ
JAMES P. HOFFLER



0055323

Gene Expression Systems

Using Nature for the Art of Expression

Edited by

Joseph M. Fernandez and James P. Hoeffler

Invitrogen Corporation

Carlsbad, California



ACADEMIC PRESS

San Diego London Boston New York Sydney Tokyo Toronto

This book is printed on acid-free paper. (∞)

Copyright © 1999 by ACADEMIC PRESS

All Rights Reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

Academic Press

a division of Harcourt Brace & Company

525 B Street, Suite 1900, San Diego, California 92101-4495, USA

<http://www.apnet.com>

Academic Press

24-28 Oval Road, London NW1 7DX, UK

<http://www.hbuk.co.uk/ap/>

Library of Congress Catalog Card Number: 98-85621

International Standard Book Number: 0-12-253840-4

PRINTED IN THE UNITED STATES OF AMERICA

98 99 00 01 02 03 MM 9 8 7 6 5 4 3 2 1

Gene Expression Systems

CONTRIBUTORS

Numbers in brackets indicate the pages on which the authors' contributions begin.

Rula Abbud [367], Department of Pharmacology, Case Western Reserve University, Cleveland, Ohio 44106

Roger N. Beachy [429], Division of Plant Biology, The Scripps Research Institute, La Jolla, California 92037

Robert P. Bennett [259], Invitrogen Corporation, Carlsbad, California 92008

Marc Better [95], XOMA Corporation, Santa Monica, California 90404

Jürgen Brosius [45], Institute for Experimental Pathology, Center for Molecular Biology of Inflammation (ZMBE), University of Münster, D-48149 Münster, Germany

E. S. Cole [399], Genzyme Corporation, Framingham, Massachusetts 01701

James M. Cregg [157], Department of Biochemistry and Molecular Biology, Oregon Graduate Institute of Science and Technology, Portland, Oregon 97291

John M. Curling [399], John Curling Consulting AB, S-75329 Uppsala, Sweden

Russell Durbin [9], Department of Hematology and Oncology, Wexner Institute for Pediatric Research, Children's Hospital, Columbus, Ohio 43205

Y. Echelard [399], Genzyme Transgenics Corporation, Framingham, Massachusetts 01701

Joseph M. Fernandez [1], Invitrogen Corporation, Carlsbad, California 92008

Eugenio Ferrari [65], Genencor International, Inc., Palo Alto, California 94304

Michael Galleno [331], Invitrogen Corporation, Carlsbad, California 92008

- S. Groet** [399], Genzyme Transgenics Corporation, Framingham, Massachusetts 01701
- M. Harvey** [399], Genzyme Transgenics Corporation, Framingham, Massachusetts 01701
- James P. Hoeffler** [1], Invitrogen Corporation, Carlsbad, California 92008
- Robert B. Kirkpatrick** [289], Department of Gene Expression Sciences, SmithKline Beecham Pharmaceuticals, King of Prussia, Pennsylvania 19406
- H. M. Meade** [399], Genzyme Transgenics Corporation, Framingham, Massachusetts 01701
- Brian Miller** [65], Genencor International, Inc., Palo Alto, California 94304
- Glen R. Nemerow** [111], Department of Immunology, The Scripps Research Institute, La Jolla, California 92037
- John H. Nilson** [367], Department of Pharmacology, Case Western Reserve University, Cleveland, Ohio 44106
- Christopher K. Raymond** [193], Department of Protein Expression Technology, ZymoGenetics, Inc., Seattle, Washington 98102
- Marijane Russell** [235], Invitrogen Corporation, Carlsbad, California 92008
- Allan Shatzman** [289], Gene Expression Sciences Department, SmithKline Beecham Pharmaceuticals, King of Prussia, Pennsylvania 19406
- August J. Sick** [331], Invitrogen Corporation, Carlsbad, California 92008
- T. E. Smith** [399], Genzyme Transgenics Corporation, Framingham, Massachusetts 01701
- Mark F. Stinski** [211], Department of Microbiology, University of Iowa, Iowa City, Iowa 52242
- Andreas Voloudakis** [429], Department of Cell Biology, The Scripps Research Institute, La Jolla, California 92037
- Daniel J. Von Seggern** [111], Department of Immunology, The Scripps Research Institute, La Jolla, California 92037
- Yanhai Yin** [429], Department of Medicine and School of Medicine, University of California, San Diego, La Jolla, California 92037
- M. W. Young** [399], Genzyme Transgenics Corporation, Framingham, Massachusetts 01701
- C. A. Ziomek** [399], Genzyme Transgenics Corporation, Framingham, Massachusetts 01701

CONTENTS

Contributors xi

Introduction

Joseph M. Fernandez and James P. Hoeffler

The Art of Expression 1

Purpose of This Book 2

Selecting a Suitable Expression System: Considerations 3

Genomics and the Future of Protein Expression Systems 5

Section I

PROKARYOTIC EXPRESSION SYSTEMS

1. Gene Expression Systems Based on Bacteriophage T7 RNA Polymerase

Russell Durbin

Introduction 10

T7 RNAP in Its Natural Habitat 11

Biochemistry of T7 RNAP 14

T7-Based Gene Expression Systems 24

Conclusions 36

References 37

2. Expression Vectors Employing the *trc* Promoter

Jürgen Brosius

Introduction 46

Vectors for *trc*-Driven Expression 51

Additional Hosts for *trc*-Driven Expression 57

Outlook for Further Improvements of Expression Vectors

Employing the *trc* Promoter 58

References 59

3. *Bacillus* Expression: A Gram-Positive Model
Eugenio Ferrari and Brian Miller
Introduction 66
Bacilli Used as Industrial Production Organisms 67
Genetic Manipulation in Bacilli 71
Plasmid Vectors 75
Transcriptional Regulation 77
Conclusions 81
Appendix 82
References 89
4. *araB* Expression System in *Escherichia coli*
Marc Better
Introduction 95
The *araB* Promoter and How It Works 97
Use of the *ara* System for Expression of Recombinant
Products 99
Conclusions 104
References 106

Section II

EUKARYOTIC EXPRESSION SYSTEMS

5. Adenoviral Vectors for Protein Expression
Dan J. Von Seggern and Glen R. Nemerow
Adenovirus Biology 112
Adaptation of Adenovirus as a Gene Transfer
Vector 117
Construction of Adenovirus Expression Vectors 119
Applications of Adenoviral Vectors in Gene
Expression 128
Considerations in the Use of Adenovirus Expression
Vectors 138
Advantages and Disadvantages of Adenoviral
Vectors 141
Future Directions 143
Conclusions 144
References 145

6. Expression in the Methylophilic Yeast *Pichia pastoris*

James M. Cregg

Introduction 158

Background Information 161

Construction of Expression Strains 166

Important Considerations in the Use of the Expression
System 175

Conclusion 183

References 184

7. Recombinant Protein Expression in *Pichia methanolica*

Christopher K. Raymond

Background 193

DNA Transformation 194

Expression Vector and Identification of an Expression
Strain 196

Gene Disruptions and Generation of a Protease-Deficient
Strain 198

Fermentation 199

Evaluation of Protein Expression 199

Expression of the 65-kDa Isoform of Human Glutamate
Decarboxylase in *P. pastoris* and *P. methanolica* 202

Protein Secretion 202

Benefits and Liabilities of the *P. methanolica* Expression
System 205

Summary 206

Appendix 207

References 207

8. Cytomegalovirus Promoter for Expression in Mammalian Cells

Mark F. Stinski

Introduction 212

The Cytomegalovirus Enhancer-Containing Promoter 214

Positive Regulation 216

Experimental Procedures 219

Effects Upstream of the Cytomegalovirus Enhancer 221

Effects Downstream of the Cytomegalovirus Promoter 222

Negative Regulation 223

Conclusions	225
References	228

9. Inducible Mammalian Expression Systems

Marijane Russell

Background: Development of Inducible Expression Systems	235
Literature Review	245
Author's Experience	248
Comparing Inducible Mammalian Expression Systems Currently Available	250
Conclusions and Future Directions	250
References	253

10. Protein Expression in Mammalian Cells Using Sindbis Virus

Robert P. Bennett

Introduction	260
Applications	267
Protein Expression	272
Comparisons with Other Systems	279
New Directions and Conclusions	282
References	283

Section III

EXPRESSION IN INSECT SYSTEMS

11. *Drosophila* S2 System for Heterologous Gene Expression

Robert B. Kirkpatrick and Allan Shatzman

Introduction	290
Properties of S2 Line	290
Experimental Procedures	292
Literature Review	300
Choosing an Expression System	315
Future Directions	318
Conclusions	321
Appendix	321
References	325

12. Baculovirus Expression Vector System

Michael Galleno and August J. Sick

Introduction 332

Background 334

Basic Research of Baculoviruses as a Tool for Gene
Expression 347

Expression Examples 349

Guidelines to Optimize Heterologous Gene Expression in
Baculovirus 352

Summary and Future 354

Appendix 355

References 359

Section IV

TRANSGENIC EXPRESSION

13. Recombinant Protein Expression in Transgenic Mice

Rula Abbud and John H. Nilson

Overview and Background 368

Transgenic Mice: General Overview 368

Examples of Recombinant Protein Expression in Transgenic
Mice 371

Modeling Human Disease through Targeted
Overexpression 378

Biomass: Transgenic Mice as Model Bioreactors 381

Future of Transgenic Mice Approaches: Need for Inducible
Expression 383

Conclusions 390

References 391

14. Expression of Recombinant Proteins in the Milk of Transgenic Animals

*H. M. Meade, Y. Echelard, C. A. Ziomek, M. W. Young,
M. Harvey, E. S. Cole, S. Groet, T. E. Smith, and J. M. Curling*

Introduction 400

Expression of Heterologous Proteins in Milk 401

Milk-Specific Transgenes 403

Insertion of the Transgene into the Germ Line 405

Transgenic Animal Production	408
Biosynthesis of Milk Proteins	409
Milk Secretion from the Mammary Gland	412
Lactation and Milk Output	412
Milk Composition and Purification of the Target Protein	413
Quality Issues in Transgenic Production	417
Regulatory Considerations	418
Current Status and Future Directions	420
References	421

15. Recombinant Protein Expression in Plants

Andreas E. Voloudakis, Yanhai Yin, and Roger N. Beachy

General Introduction	430
Transformation Methods	431
Promoters Used for Recombinant Protein Accumulation in Plants	435
Expression of Recombinant Proteins in Plants and Agricultural Biotechnology	438
Recombinant Protein Expression in Plants to Obtain New Products	442
Protein Targeting and Accumulation	447
Virus-Mediated Expression Systems	448
Summary/Discussion	450
References	452

Introduction

SO MANY POSSIBILITIES: HOW TO CHOOSE A SYSTEM TO ACHIEVE YOUR SPECIFIC GOAL

Joseph M. Fernandez and James P. Hoeffler

Invitrogen Corporation, Carlsbad, California 92008

The Art of Expression

Purpose of This Book

Selecting a Suitable Expression System: Considerations

Genomics and the Future of Protein Expression Systems

The Art of Expression

Our understanding of gene expression has increased greatly since the mid-1980s, resulting in new developments in protein expression system technology. A key element in creating efficient and economic expression systems has been the construction of vectors that include, along with the gene of interest, the appropriate promoter and other regulatory sequences. Recombinant DNA techniques have enabled unique pairings of promoters and structural genes in a wide variety of vectors for expression of desired recombinant proteins. These new gene combinations are currently being utilized in numerous prokaryotic and eukaryotic organisms to produce recombinant products of both academic and industrial importance. Notably, all the elements that constitute an expression system—structural genes, control sequences, markers, and induc-

ers—are present in nature; scientists have only had to mix and match them from a variety of organisms to create diversity and flexibility in the regulation of gene expression, much as a painter mixes colors on a palette to obtain the desired shades. The result is a wide range of effective expression systems available to researchers to achieve their specific objectives: from high-level expression of recombinant products for large-scale production, to subtle expression for studying protein function in the cell, from prokaryotes to transgenic animals. Selecting the most suitable system depends on a series of parameters, such as time, resources, and intended use. Often, after weighing advantages and disadvantages, there is not always a clear winner, but there may be some areas of overlap among alternatives. The choice is indeed so wide and complex that selecting the right system has become an art in itself, further emphasizing the fact that Nature is a palette for the art of expression.

Purpose of This Book

The aim of this book is to provide the latest information on state-of-the-art protein expression systems and to help the reader select one that will best suit individual goals and resources. Leading scientists in the field review the most popular prokaryotic and eukaryotic expression systems: from bacteria (Section I) to yeast (Section II) and to insects (Section III), mammalian cell cultures (Section II), and transgenic animals and plants (Section IV). Advantages and disadvantages of each system are surveyed and summarized for easy reference. Each chapter illustrates how a system works and what proteins are most likely to work well, addressing potential problems and suggesting solutions. Although some information is offered on methods for generating recombinants, detailed protocols are beyond the scope of this book and, for these, most authors provide references. In light of the increasingly important role of protein expression, most chapters also contribute insight into future developments. Finally, in an effort to provide further guidance to the reader, prokaryotic and eukaryotic systems are compared for commonly desirable characteristics in a convenient, easy-to-read chart.

Selecting a Suitable Expression System: Considerations

When selecting a protein expression system, a number of considerations must be made, including the intended use, time frame, availability of resources, and the characteristics of the recombinant product. These considerations will affect the choice of expression system and type of promoter to be used. As for the intended use, this may be the isolation of the recombinant protein or the study of protein function in a cell or organism. If the objective is large-scale production of the gene product, then a yeast such as *Pichia pastoris* (discussed in Chapter 6) or a transgenic animal system (see Chapter 14) may be suitable options. Both systems allow for high-level production of exogenous eukaryotic proteins, but differ in a number of characteristics related to additional considerations. If time and cost are of concern, the yeast offers the advantages of rapid growth in inexpensive medium and easy handling of microbes, which translates into economical high-level production of the gene product. A transgenic system, conversely, involves higher costs and time-consuming activities associated with transgenic animal development, dairying, and testing; this system, however, is able to produce correctly processed protein and would therefore be a better choice when the expressed protein requires posttranslational modifications not possible in the yeast.

When the user's objective is to study the function of a recombinant protein in a cell or organism, the ability to regulate expression may be important, leading to the choice of a system that allows for maximum optimization. There are numerous stable and transient expression systems, for regulated or constitutive production, that can be used for functional studies. The choice will depend on factors such as the need for posttranslational processing to obtain a biologically active product, toxicity of the recombinant product on the host, the amount of protein needed, and the time frame of the study. Viruses, for instance, make powerful tools for the expression of heterologous gene products in higher eukaryotes because of their high transfection efficiencies and the high level of recombinant protein expression. The Sindbis virus (reviewed in Chapter 10) is an example of an effective vector system particularly suitable for transient expression, as infection with this virus results in eventual death of the host cells. In contrast, retroviruses integrate into the host genome

and therefore represent an attractive system if stable protein expression is desired, allowing for optimization of a cell line.

One important consideration is the effect of the recombinant protein on the host. If the protein is toxic to or if it inhibits growth of the host cells, a transient system or a system that uses an inducible promoter may be desirable. In Chapter 4, the reader will find information on the arabinose expression system, which allows for the tightly regulated expression of recombinant prokaryotic and eukaryotic proteins in bacteria, and Chapter 9 discusses inducible mammalian expression systems that enable regulated inducible expression, including prokaryotic control elements in eukaryotic cells, as well as the novel insect system based on regulatory elements from *Drosophila*.

Finally, to help the reader in the identification of a suitable choice, we have provided a quick comparison chart (Table 1) in

Table 1
Comparison of Expression Systems

Desired characteristics	Expression system			
	Bacteria	Yeast	Insect	Mammalian cell culture
Cell growth	Rapid	Rapid	Slow	Slow
Complexity of growth medium	Minimum	Minimum	Complex	Complex
Cost of growth medium	Low	Low	High	High
Expression level	High	Low to high	Low to high	Low to moderate
Extracellular expression	Secretion to periplasm	Secretion to medium	Secretion to medium	Secretion to medium
Posttranslational modifications				
Protein folding	Refolding usually required	Refolding may be required	Proper folding	Proper folding
N-linked glycosylation	None	High mannose	Simple, no sialic acid	Complex
O-linked glycosylation	No	Yes	Yes	Yes
Phosphorylation	No	Yes	Yes	Yes
Acetylation	No	Yes	Yes	Yes
Acylation	No	Yes	Yes	Yes
γ -Carboxylation	No	No	No	Yes

which the expression systems featured in this book are compared for commonly desired characteristics. Each system category (bacteria, yeast, insect, mammalian cell cultures) is reviewed, respectively, in Section I, II, III and IV of this book. Whatever the challenge, selecting the right system is an important key to success.

Genomics and the Future of Protein Expression Systems

Developments in DNA array technology have made it possible to define gene expression profiles as never before, enabling the simultaneous analysis of thousands of genes, large-scale gene discovery, and mapping of genomic DNA clones. DNA arrays are used to measure the expression levels of prokaryotic and eukaryotic genes and can quickly elucidate the correlation between gene expression and biochemical pathways. In addition, they have provided information on the expression patterns of many previously unknown genes. DNA chip technology is advancing rapidly, with applications in diagnostics (mutation detection), gene discovery, gene expression, and mapping, as well as pharmaceutical development. Because of the abundance of data generated by this technology, protein expression systems will be increasingly more important, as researchers study the structure and function of gene products identified by DNA arrays. However, with the plethora of expression systems available, one thing to bear in mind is that there is always an element of unpredictability in the behavior of any given protein in a system. No matter how advanced the state of technology and how carefully all parameters are evaluated when selecting a system, no system is totally predictable and ultimately, the only sure way to find out if a system will work is to try and experiment with it. As Mark Ratner pointed out in *Bio/Technology* 1989, "Expression systems are protein specific. You must be able to play around with each one, insert your gene of choice, tweak it, and then see what you've got."