

# Evolutionary Methods in Biotechnology

Clever Tricks for Directed Evolution

Edited by Susanne Brakmann  
and Andreas Schwienhorst



WILEY-  
VCH

WILEY-VCH Verlag GmbH & Co. KGaA



# Evolutionary Methods in Biotechnology

Edited by Susanne Brakmann and Andreas Schwienhorst

# Further Titles of Interest

G. Kahl

**The Dictionary of Gene Technology**

3rd edition 2004, ISBN 3-527-30765-6

A.S. Bommarius, B.R. Riebel

**Biocatalysis – Fundamentals and Applications**

2004, ISBN 3-527-30344-8

T. Dingermann, D. Steinhilber, G. Folkers (eds.)

**Molecular Biology in Medicinal Chemistry**

2003, ISBN 3-527-30431-2

S. Brakmann, K. Johnsson (eds.)

**Directed Molecular Evolution of Proteins**

2002, ISBN 3-527-30423-1

V. Braun, F. Götz (eds.)

**Microbial Fundamentals of Biotechnology**

2001, ISBN 3-527-30615-3

# Evolutionary Methods in Biotechnology

Clever Tricks for Directed Evolution

Edited by Susanne Brakmann  
and Andreas Schwienhorst



WILEY-  
VCH

WILEY-VCH Verlag GmbH & Co. KGaA

Dr. Susanne Brakmann  
Institut für Zoologie  
Angewandte Molekulare Evolution  
Universität Leipzig  
Liebigstr. 18  
04103 Leipzig, Germany  
sbrakma@uni-leipzig.de

Dr. Andreas Schwienhorst  
Institut für Mikrobiologie und Genetik  
Universität Göttingen  
Grisebachstr. 8  
37077 Göttingen, Germany  
aschwie1@gwdg.de

This book was carefully produced. Nevertheless, authors, editors and publisher do not warrant the information contained therein to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

Library of Congress Card No. applied for.

British Library Cataloguing-in-Publication Data: A catalogue record for this book is available from the British Library

Die Deutsche Bibliothek – CIP Cataloguing-in-Publication-Data: A catalogue record for this publication is available from Die Deutsche Bibliothek

ISBN 3-527-30799-0

© 2004 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Printed on acid-free paper

All rights reserved (including those of translation into other languages). No part of this book may be reproduced in any form – by photoprinting, microfilm, or any other means – nor transmitted or translated into a machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

Composition: EDV-Beratung Frank Herweg, Leutershausen  
Printing: betz-druck gmbh, Darmstadt  
Bookbinding: J. Schäffer GmbH & Co. KG, Grünstadt

Printed in the Federal Republic of Germany.

# Contents

## 1 Introduction

|  |   |
|--|---|
| <i>Susanne Brakmann and Andreas Schwienhorst</i> . . . . . | 1 |
| References . . . . .                                       | 3 |

## 2 Generation of Mutant Libraries Using Random Mutagenesis

|  |    |
|--|----|
| <i>Susanne Brakmann and Björn F. Lindemann</i> . . . . .               | 5  |
| 2.1 Introduction . . . . .   | 5  |
| 2.2 Materials . . . . .  | 6  |
| 2.2.1 Materials for Random PCR Mutagenesis . . . . .                   | 6  |
| 2.2.2 Materials for Mutator Strain Passage . . . . .                   | 6  |
| 2.3 Protocols . . . . .  | 7  |
| 2.3.1 Protocol for Random PCR Mutagenesis According to Joyce . . . . . | 7  |
| 2.3.2 Protocol for Mutator Strain Passage . . . . .                    | 8  |
| 2.4 Troubleshooting . . . . .  | 10 |
| References . . . . .   | 11 |

## 3 DNA Shuffling

|   |    |
|---|----|
| <i>Hikaru Suenaga, Masatoshi Goto, and Kensuke Furukawa</i> . . . . .                         | 13 |
| 3.1 Introduction . . . . .  | 13 |
| 3.2 Materials . . . . .   | 15 |
| 3.2.1 For Preparation of Parental Genes . . . . .   | 15 |
| 3.2.2 For Random Fragmentation by DNase I . . . . .   | 15 |
| 3.2.3 For Collection of DNA Fragments<br>in Specific Molecular Size Ranges . . . . .          | 16 |
| 3.2.4 For Reassembly of These Fragments by Primerless PCR . . . . .                           | 16 |
| 3.2.5 For Amplification of Reassembled Products<br>by Conventional PCR with Primers . . . . . | 16 |
| 3.3 Protocol . . . . .  | 17 |
| 3.3.1 Preparation of Parental Genes . . . . .   | 17 |
| 3.3.2 Random Fragmentation by DNase I . . . . .   | 17 |
| 3.3.3 Collection of DNA Fragments<br>in Specific Molecular Size Ranges . . . . .              | 18 |
| 3.3.4 Reassembly of These Fragments by Primerless PCR . . . . .                               | 19 |
| 3.3.5 Amplification of Reassembled Products<br>by Conventional PCR with Primers . . . . .     | 20 |

|          |  |    |
|----------|--|----|
| 3.4      | Troubleshooting . . . . .  | 21 |
| 3.4.1    | Insufficient DNase I Fragmentation . . . . .   | 21 |
| 3.4.2    | Little or No Product of Primerless PCR . . . . .   | 21 |
| 3.4.3    | Little or No Product of PCR with Primers . . . . .   | 21 |
| 3.4.4    | The Product of PCR with Primers is Multi-banded . . . . .  | 22 |
| 3.5      | Amplification Examples . . . . .   | 22 |
|          | References . . . . .   | 23 |
| <b>4</b> | <b>DNA Recombination Using StEP</b>  |    |
|          | <i>Milena Ninkovic</i> . . . . .   | 25 |
| 4.1      | Introduction . . . . .   | 25 |
| 4.2      | Materials . . . . .  | 26 |
| 4.2.1    | StEP PCR . . . . .   | 26 |
| 4.2.2    | Purification of an Appropriate DNA Fragment . . . . .  | 27 |
| 4.2.3    | Equipment . . . . .  | 27 |
| 4.3      | Protocol . . . . .   | 27 |
| 4.4      | Technical Tips . . . . .   | 28 |
| 4.4.1    | Problem: Little or No PCR Product<br>(Full-length Product) after PCR . . . . .   | 28 |
| 4.4.2    | Problem: High Background Levels of DNA after PCR . . . . .   | 28 |
| 4.5      | StEP in Directed Evolution . . . . .   | 29 |
|          | References . . . . .   | 30 |
| <b>5</b> | <b>FACS Screening of Combinatorial Peptide and Protein Libraries<br/>Displayed on the Surface of <i>Escherichia coli</i> Cells</b> |    |
|          | <i>Thorsten M. Adams, Hans-Ulrich Schmoldt, and Harald Kolmar</i> . . . . .  | 31 |
| 5.1      | Introduction . . . . .   | 31 |
| 5.2      | Materials . . . . .  | 35 |
| 5.2.1    | <i>Escherichia coli</i> Strains and Plasmids . . . . .   | 35 |
| 5.2.2    | Liquid Media and Agar Plates . . . . .   | 35 |
| 5.2.3    | Biological and Chemical Materials . . . . .  | 36 |
| 5.2.4    | Equipment . . . . .  | 36 |
| 5.3      | Protocols . . . . .  | 36 |
| 5.3.1    | Verification of Cell Surface Exposure<br>of the Passenger Protein . . . . .  | 36 |
| 5.3.2    | Labeling of the Target Protein . . . . .   | 37 |
| 5.3.3    | Library Construction . . . . .   | 37 |
| 5.3.4    | Combinatorial Library Screening by FACS and MACS . . . . .   | 40 |
| 5.4      | Troubleshooting . . . . .  | 42 |
| 5.5      | Major Applications . . . . .   | 44 |
|          | References . . . . .   | 44 |
| <b>6</b> | <b>Selection of Phage-displayed Enzymes</b>  |    |
|          | <i>Patrice Soumillion</i> . . . . .  | 47 |
| 6.1      | Introduction . . . . .   | 47 |
| 6.2      | Materials . . . . .  | 50 |



|       |  |    |
|-------|--|----|
| 6.2.1 | Buffers, Reagents and Consumables . . . . .                          | 50 |
| 6.2.2 | Strains and Vectors . . . . .  | 50 |
| 6.3   | Protocols . . . . .  | 51 |
| 6.3.1 | The Phage-enzyme . . . . .   | 51 |
| 6.3.2 | Library Construction . . . . .                                       | 55 |
| 6.3.3 | Selection . . . . .  | 59 |
| 6.4   | Troubleshooting . . . . .  | 62 |
| 6.4.1 | Phage Titers Are Not Reproducible . . . . .                          | 62 |
| 6.4.2 | Phage-enzymes Degrade with Time . . . . .                            | 63 |
| 6.4.3 | Phages Are Not Genetically Stable . . . . .                          | 63 |
| 6.4.4 | The ‘out/in’ Ratio Does Not Increase with Selection Rounds . . . . . | 63 |
| 6.5   | Major Applications . . . . .   | 63 |
|       | References . . . . .   | 64 |

**7 Selection of Aptamers**

|       |  |    |
|-------|--|----|
|       | <i>Heiko Fickert, Heike Betat, and Ulrich Hahn</i> . . . . .   | 65 |
| 7.1   | Introduction . . . . .   | 65 |
| 7.2   | Materials . . . . .  | 66 |
| 7.2.1 | Immobilization of Target Molecules . . . . .                   | 66 |
| 7.2.2 | PCR . . . . .  | 67 |
| 7.2.3 | <i>In vitro</i> Transcription . . . . .                        | 67 |
| 7.2.4 | RNA Purification . . . . .                                     | 67 |
| 7.2.5 | Selection of Aptamers . . . . .                                | 67 |
| 7.2.6 | Reverse Transcription . . . . .                                | 68 |
| 7.3   | Protocols . . . . .  | 68 |
| 7.3.1 | Selection of RNA Aptamers . . . . .                            | 68 |
| 7.3.2 | Selection of 2'-Modified RNA Aptamers . . . . .                | 75 |
| 7.3.3 | Selection of ssDNA Aptamers . . . . .                          | 76 |
| 7.3.4 | Cloning and Sequencing . . . . .                               | 77 |
| 7.3.5 | Characterization of Aptamers . . . . .                         | 77 |
| 7.3.6 | Example: Isolation of Moenomycin A-specific Aptamers . . . . . | 79 |
| 7.4   | Troubleshooting . . . . .                                      | 82 |
| 7.5   | Major Applications . . . . .                                   | 83 |
|       | References . . . . .   | 83 |

**8 Methods for Selecting Catalytic Nucleic Acids**

|       |   |     |
|-------|---|-----|
|       | <i>Benjamin L. Holley, and Bruce E. Eaton</i> . . . . . | 87  |
| 8.1   | Introduction . . . . .                                  | 87  |
| 8.2   | Materials and Equipment . . . . .                       | 88  |
| 8.3   | Protocols . . . . .                                     | 91  |
| 8.3.1 | Generating the Starting Library . . . . .               | 91  |
| 8.3.2 | Transcription . . . . .                                 | 97  |
| 8.3.3 | Ligation . . . . .                                      | 99  |
| 8.3.4 | Nucleic Acid-catalyzed Reactions . . . . .              | 102 |
| 8.3.5 | Reverse Transcription . . . . .                         | 104 |

|           |  |     |
|-----------|--|-----|
| 8.3.6     | Partitioning . . . . .   | 105 |
| 8.4       | Troubleshooting . . . . .  | 108 |
| 8.5       | Major Applications . . . . .   | 109 |
|           | References . . . . .   | 109 |
| <b>9</b>  | <b>High-throughput Screening<br/>of Enantioselective Industrial Biocatalysts</b>       |     |
|           | <i>Manfred T. Reetz</i> . . . . .  | 113 |
| 9.1       | Introduction . . . . .   | 113 |
| 9.2       | Materials and Equipment . . . . .  | 115 |
| 9.2.1     | Assays Based on Mass Spectrometry . . . . .  | 115 |
| 9.2.2     | Assays Based on NMR Spectrometry . . . . .   | 116 |
| 9.2.3     | Assay Based on FTIR Spectroscopy . . . . .   | 116 |
| 9.2.4     | Assays Based on UV/Visible Spectroscopy . . . . .                                      | 116 |
| 9.2.5     | Enzyme-coupled UV/Visible-based Assay for Hydrolases . . . . .                         | 117 |
| 9.3       | Protocols . . . . .  | 117 |
| 9.3.1     | Assays Based on Mass Spectrometry . . . . .  | 117 |
| 9.3.2     | Assays Based on NMR Spectroscopy . . . . .   | 121 |
| 9.3.3     | Assay Based on FTIR Spectroscopy . . . . .   | 125 |
| 9.3.4     | Assays Based on UV/Visible Spectroscopy . . . . .                                      | 129 |
| 9.3.5     | Enzyme-coupled UV/Visible-based Assay for Hydrolases . . . . .                         | 132 |
| 9.3.6     | Further Assays . . . . .   | 133 |
| 9.4       | Troubleshooting . . . . .  | 138 |
| 9.4.1     | Comments on the Kazlauskas Test . . . . .  | 138 |
| 9.4.2     | Potential Problems when Performing Kinetic Resolution . . . . .                        | 138 |
| 9.5       | Conclusions . . . . .  | 139 |
|           | References . . . . .   | 139 |
| <b>10</b> | <b>Computer-assisted Design of Doped Libraries</b>                                     |     |
|           | <i>Dirk Tomandl and Andreas Schwienhorst</i> . . . . .                                 | 143 |
| 10.1      | Introduction . . . . .   | 143 |
| 10.2      | Materials . . . . .  | 146 |
| 10.3      | Protocol . . . . .   | 147 |
| 10.4      | Troubleshooting . . . . .  | 150 |
| 10.5      | Major Applications . . . . .   | 150 |
|           | References . . . . .   | 151 |
| <b>11</b> | <b>Directed <i>in silico</i> Mutagenesis</b>   |     |
|           | <i>Markus Wiederstein, Peter Lackner, Ferry Kienberger, Manfred J. Sippl</i> . . . . . | 153 |
| 11.1      | Introduction . . . . .   | 153 |
| 11.2      | Materials . . . . .  | 155 |
| 11.2.1    | PDB Files . . . . .  | 155 |
| 11.2.2    | Knowledge-based Potentials . . . . .   | 156 |
| 11.2.3    | Polyprotein, Z-scores . . . . .  | 160 |
| 11.2.4    | <i>In silico</i> Mutagenesis . . . . .   | 162 |
| 11.2.5    | Summary . . . . .  | 163 |

|           |   |     |
|-----------|---|-----|
| 11.3      | Protocol . . . . .  | 163 |
| 11.3.1    | ProSa Setup and Interaction . . . . .   | 163 |
| 11.3.2    | ProSa Objects . . . . .   | 164 |
| 11.3.3    | Session 1 (mut_script1.cmd) . . . . .   | 164 |
| 11.3.4    | Session 2 (mut_script2.cmd) . . . . .   | 166 |
| 11.3.5    | Session 3 (mut_script3.cmd) . . . . .   | 168 |
| 11.3.6    | Session 4 (mut_script4.cmd) . . . . .   | 170 |
| 11.3.7    | Tips & Tricks . . . . .   | 171 |
| 11.4      | Troubleshooting . . . . .   | 173 |
| 11.5      | Major Applications . . . . .  | 174 |
|           | References . . . . .  | 175 |
| <b>12</b> | <b>RNA Folding <i>in silico</i></b>   |     |
|           | <i>Christoph Flamm, Ivo L. Hofacker, Peter F. Stadler</i> . . . . .   | 177 |
| 12.1      | Introduction . . . . .  | 177 |
| 12.2      | Materials . . . . .   | 178 |
| 12.2.1    | Typographical Conventions . . . . .   | 179 |
| 12.2.2    | RNA Web Services . . . . .  | 180 |
| 12.3      | Protocols . . . . .   | 181 |
| 12.3.1    | Secondary Structures for Individual Sequences . . . . .   | 181 |
| 12.3.2    | Consensus Structures of a Sample of Sequences . . . . .   | 183 |
| 12.3.3    | Sequence Design . . . . .   | 184 |
| 12.3.4    | Analysis of SELEX Experiments . . . . .   | 186 |
| 12.3.5    | A Note for the Experts: Write your Own RNA Programs . . . . .   | 187 |
| 12.4      | Troubleshooting . . . . .   | 187 |
| 12.5      | Caveats . . . . .   | 188 |
|           | References . . . . .  | 189 |
| <b>13</b> | <b>Patenting in Evolutionary Biotechnology</b>  |     |
|           | <i>Martina Leimkühler and Hans-Wilhelm Meyers</i> . . . . .   | 191 |
| 13.1      | Introduction . . . . .  | 191 |
| 13.2      | The Nature of Patents . . . . .   | 191 |
| 13.3      | What Can Be Patented . . . . .  | 192 |
| 13.4      | The Requirement of Novelty . . . . .  | 193 |
| 13.5      | The Requirement of Inventiveness . . . . .  | 196 |
| 13.6      | The Requirement of Utility . . . . .  | 197 |
| 13.7      | The Requirements of Enablement and Written Description . . . . .  | 197 |
| 13.8      | Patent Prosecution . . . . .  | 199 |
| 13.9      | Search Tools . . . . .  | 204 |
| 13.10     | The First-to-invent Principle of the United States<br>and Its Consequences on Laboratory Notebook Keeping . . . . . | 206 |
| 13.11     | Summary . . . . .   | 209 |
|           | References . . . . .  | 209 |
|           | <b>Subject Index</b> . . . . .  | 211 |



## List of Contributors

Thorsten M. Adams  
Institut für Mikrobiologie und Genetik  
Universität Göttingen  
Grisebachstr. 8  
37077 Göttingen  
Germany

Heike Betat  
Max-Planck-Institut für  
Evolutionäre Anthropologie  
Deutscher Platz 6  
04103 Leipzig  
Germany

Susanne Brakmann  
Institut für Zoologie  
Angewandte Molekulare Evolution  
Universität Leipzig  
Liebigstr. 18  
04103 Leipzig  
Germany

Bruce E. Eaton  
Department of Chemistry  
North Carolina State University  
Raleigh, NC 27695  
USA

Heiko Fickert  
Universität Hamburg  
Institut für Biochemie  
und Lebensmittelchemie  
Martin-Luther-King Platz 6  
20146 Hamburg  
Germany

Christoph Flamm  
Institut für Theoretische Chemie  
und Strukturbioogie  
Universität Wien  
Währingerstr. 17  
1090 Wien  
Austria

Kensuke Furukawa  
Department of Bioscience  
and Biotechnology  
Kyushu University  
Fukuoka 812-8581  
Japan

Masatoshi Goto  
Department of Bioscience  
and Biotechnology  
Kyushu University  
Fukuoka 812-8581  
Japan

Ulrich Hahn  
Universität Hamburg  
Institut für Biochemie  
und Lebensmittelchemie  
Martin-Luther-King Platz 6  
20146 Hamburg  
Germany

Ivo L. Hofacker  
Institut für Theoretische Chemie  
und Strukturbioogie  
Universität Wien  
Währingerstr. 17  
1090 Wien  
Austria

Benjamin L. Holley  
Department of Chemistry  
North Carolina State University  
Raleigh, NC 27695  
USA

Ferry Kienberger  
Center of Applied Molecular  
Engineering  
Institute of Chemistry  
and Biochemistry  
University of Salzburg  
Jakob Haringerstr. 5  
5020 Salzburg  
Austria

Harald Kolmar  
Institut für Mikrobiologie und Genetik  
Universität Göttingen  
Grisebachstr. 8  
37077 Göttingen  
Germany

Peter Lackner  
Center of Applied Molecular  
Engineering  
Institute of Chemistry  
and Biochemistry  
University of Salzburg  
Jakob Haringerstr. 5  
5020 Salzburg  
Austria

Martina Leimkühler  
Evotec OAI AG  
Schnackenburgallee 114  
22525 Hamburg  
Germany

Björn Lindemann  
Bioagency AG  
Schnackenburgallee 116a  
22525 Hamburg  
Germany

Hans-Wilhelm Meyers  
Deichmannhaus am Dom  
Bahnhofsvorplatz 1  
50667 Köln  
Germany

Milena Ninkovic  
Max-Planck-Institut  
für Experimentelle Medizin  
Hermann-Rein-Str. 3  
37075 Göttingen  
Germany

Manfred T. Reetz  
Max-Planck-Institut  
für Kohlenforschung  
Kaiser-Wilhelm-Platz 1  
45470 Mülheim/Ruhr  
Germany

Hans-Ulrich Schmoltd  
Institut für Mikrobiologie und Genetik  
Universität Göttingen  
Grisebachstr. 8  
37077 Göttingen  
Germany

Andreas Schwienhorst  
Institut für Mikrobiologie und Genetik  
Universität Göttingen  
Grisebachstr. 8  
37077 Göttingen  
Germany

Manfred J. Sippl  
Center of Applied Molecular  
Engineering  
Institute of Chemistry  
and Biochemistry  
University of Salzburg  
Jakob Haringerstr. 5  
5020 Salzburg  
Austria

Patrice Soumillion  
Institut des Sciences de la Vie  
Université Catholique de Louvain  
Place Pasteur 1/1b  
1348 Louvain-la-Neuve  
Belgium

Peter F. Stadler  
Institut für Informatik  
Universität Leipzig  
Kreuzstrasse 7b  
04103 Leipzig  
Germany

Hikaru Suenaga  
Institute for Biological Resources  
and Functions  
National Institute of Advanced  
Industrial Science and Technology  
Tsukuba, Ibaraki 305-8566  
Japan

Dirk Tomandl  
Graffinity Pharmaceuticals AG  
Im Neuenheimer Feld 518 - 519  
69120 Heidelberg  
Germany

Markus Wiederstein  
Center of Applied Molecular  
Engineering  
Institute of Chemistry  
and Biochemistry  
University of Salzburg  
Jakob Haringerstr. 5  
5020 Salzburg  
Austria





# 1 Introduction

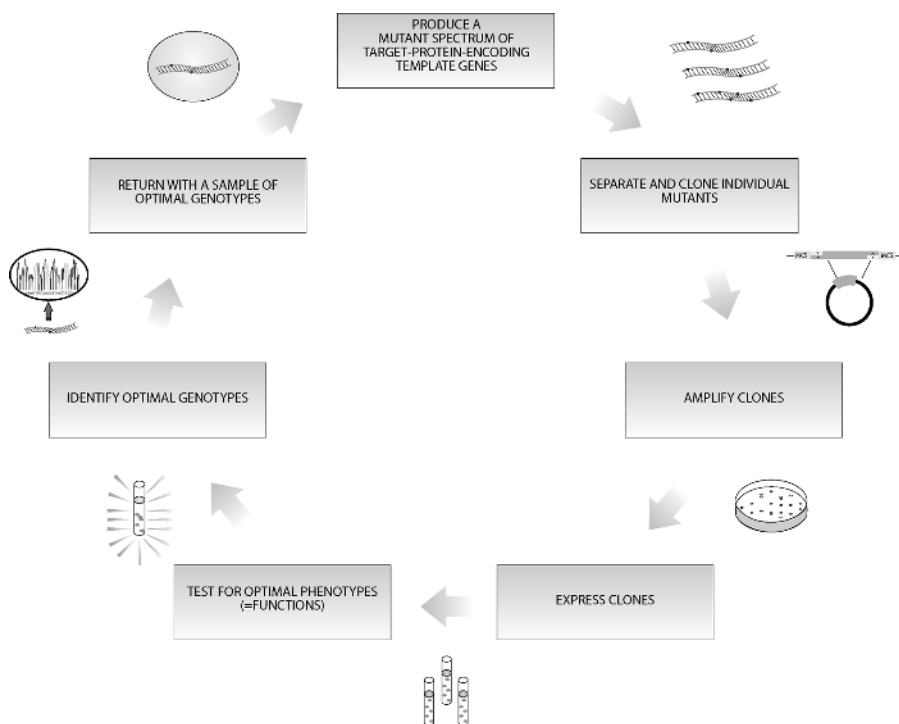
*Susanne Brakmann and Andreas Schwienhorst*

Since the landmark papers of Manfred Eigen [1, 2] and Sol Spiegelman [3, 4], the concept of Darwinian evolution has had a major impact on the design of biomolecules with tailored properties. ‘Directed evolution’, ‘applied evolution’, and ‘evolutionary biotechnology’ are different expressions that all describe an ‘evolutionary’ type of optimization strategy that comprises several cycles each consisting of (1) molecular library preparation to create the desired molecular diversity, (2) functional selection or screening, and (3) error-prone amplification or chemical modification of selected species to generate a new library of molecules (Fig.1.1). The ultimate goal is to identify molecular species that are well-adapted to a given profile of defined demands. Biocatalysts, for example, may be generated to exhibit high processivity, enantioselectivity, or tolerance to high temperatures or organic solvents.

The book presented here is intended as a practical state-of-the-art compilation of methods related to the topic of directed evolution and hence is complementary to the recent successful book *Directed Molecular Evolution of Proteins* [5]. The methods are described in sufficient detail to serve as ‘recipes’ in a ‘cookbook’. They are easy to follow by laboratory staff, from the technical assistant to the postdoctoral academic or industrial specialist.

The sequence of chapters mirrors the steps in a standard directed-evolution experiment. In the beginning, various methods for the creation of molecular diversity are considered. S. Brakmann and B.F. Lindemann (Chapter 2) present protocols for the generation of mutant libraries by random mutagenesis. Two chapters deal with the particularly powerful approach of *in-vitro* recombination. H. Suenaga, M. Goto, and K. Furukawa (Chapter 3) describe the application of DNA shuffling, and M. Ninkovic (Chapter 4) presents DNA recombination by the StEP method.

Next, several chapters are concerned with techniques of selection and/or mass screening technologies. T. Adams, H.-U. Schmoldt, and H. Kolmar (Chapter 5) describe the FACS-based screening of combinatorial peptide and protein libraries. P. Soumillon (Chapter 6) presents some of the latest developments in the selection of phage-displayed enzymes. In Chapter 7, H. Fickert, H. Betat, and U. Hahn provide methods for the selection of specific target-binding nucleic acids, i. e., aptamers. Related methods for the generation of catalytic nucleic acids are described by B.L. Holley and B.E. Eaton (Chapter 8). The part on functional selection and screening closes with a description of high-throughput screening approaches, in



**Fig. 1.1.** Scheme of directed evolution. Starting from a pool of mutant genes, single clones are expressed and their phenotype is evaluated in a selection or screening step. Clones with desired phenotypes provide genes that are the basis for the subsequent cycle of selection.

particular, to produce enantioselective industrial biocatalysts, provided by M.T. Reetz (Chapter 9).

Combinatorial mutagenesis easily produces a degree of molecular diversity that far exceeds the number of different proteins or functional nucleic acids that can be produced in a single experiment. As the number  $n$  of randomized amino acid positions in a protein grows, the number of possible combinations increases as  $20^n$ . Hence, complete coverage of a library with 9 randomized positions requires a library size well above  $10^{11}$  molecules. Since in a standard random library, functional molecules are usually highly diluted in a large background of nonfunctional, e. g., misfolded, molecules, it may be meaningful to restrict variations to a certain subset of promising molecules. Three chapters deal with theoretical computer-based methods to predict these promising molecular species. D. Tomandl and A. Schwienhorst (Chapter 10) report a ‘doping’ algorithm that helps to design random codons for only subsets of amino acids, at the same time minimizing stop codons. M. Wiederstein, P. Lackner, F. Kienberger, and M.J. Sippl (Chapter 11) provide algorithms to predict (mutant) protein structures as a means of *in silico* mutagenesis, e. g., to enhance the probability of generating properly folded mutant proteins. C. Flamm, I.L.

Hofacker, and P.F. Stadler (Chapter 12) pursue a similar goal concerning functional nucleic acids and provide various *in silico* tools to predict RNA folding.

In the past 10 years, directed evolution has gained considerable attention as a commercially important strategy for rapidly designing molecules with properties tailored for the biotechnological and pharmaceutical market. Therefore, legal protection of methods and molecules has become an important issue. Hence, the book closes with Chapter 13, by M. Leimkühler and H.W. Meyers on patenting issues in evolutionary biotechnology.

Since the first evolution experiments by Sol Spiegelman, Manfred Eigen, and coworkers, the field of directed evolution itself has evolved into a plethora of different methodologies that can hardly be covered comprehensively in a standard textbook. We nevertheless tried to provide a collection of protocols useful to the novice as well as to the scientist experienced in the field. We hope to provide a practical starting point and at the same time inspire scientists to develop their own variations on the evolutionary theme.

We thank all the authors for their contributions, and Peter Gölitz and Frank Weinreich of Wiley-VCH for their help in publishing this book.

## References

1. Eigen, M. *Die Naturwissenschaften*, **1971**, 58, 465–523.
2. Eigen, M. and Gardiner, W. *Pure Appl. Chem.*, **1984**, 56, 967–976.
3. Spiegelman, S., Haruna, I., Holland, I.B., Beaudreau, G., and Mills, D. *Proc. Natl. Acad. Sci. USA*, 1965, 54, 919-927.
4. Mills, D.R., Peterson, R.L., and Spiegelman, S. *Proc. Natl. Acad. Sci. USA*, **1967**, 58, 217–224.
5. Brakmann, S., Johnsson, K., eds. *Directed Molecular Evolution of Proteins: or How to Improve Enzymes for Biocatalysis*, Wiley-VCH, Weinheim, **2002**.



## 2 Generation of Mutant Libraries Using Random Mutagenesis

*Susanne Brakmann and Björn F. Lindemann*

### 2.1 Introduction

Engineering enzymes by applying directed evolution strategies involves the generation of molecular libraries that are as large and as diverse as possible. However, mutant libraries of enzymes, which usually consist of more than 100 amino acids, are inaccessible by automatic chemical synthesis. These are better available by mutagenesis at the nucleotide level. During the past decade, a series of experimental strategies has been developed for generating DNA mutant libraries that differ in diversity, that is, in the extent of sequence space covered, and in their way to deal with complex libraries.

*Random mutagenesis* is a widespread strategy which targets whole genes. This may be achieved by passing cloned genes through mutator strains [1, 2], by treating DNA or whole bacteria with various chemical mutagens [3–6], or by “error-prone” [7, 8] or “hypermutagenic” PCR [9]. Due to its simplicity and versatility, random PCR mutagenesis has emerged as the most common technique and can result in mutation frequencies as high as 10% per nucleotide position. The incorporation of nucleotide analogs that promote base pair mismatching during PCR has even been found to cause overall mutation frequencies of up to 19% per position and PCR [10]. With alterations of some PCR conditions, the mutation rate may be adjusted to the appropriate level (see Table 2.2). Usually, a maximal number of mutants (and no wildtype) is required, of which as many variants as possible should be active. For example, catalytically active variants of enzymes like HIV reverse transcriptase, *Taq* polymerase, or HSV-1 thymidine kinase almost never contain more than five amino acid substitutions [11]. We should also mention that the number of amino acid substitutions accessible by error-prone PCR is limited, because on the one hand, the reaction may bias the distribution of mutation type (depending on the sequence), and on the other hand, multiple substitutions within a single codon are extremely rare.

Alternative random mutagenesis strategies have been developed for targeting single or a few amino acids or selected regions of a protein that might be important for a certain function. By focusing on only the positions of interest and their close environment or by reducing the set of amino acids per randomized position (see Chapter 10 by Tomandl), the library size can be drastically reduced. Typically,

randomization of small gene fragments is achieved by substituting a wildtype gene fragment with a synthetic oligonucleotide which contains random positions or regions (random cassettes [12, 13]) or semi-random ranges (spiked oligonucleotides [14]). Randomization of defined positions or regions is achieved with automatic solid-phase DNA synthesis, by programming the desired International Union of Biochemistry (IUB) mix codes. The introduction of stop codons can be reduced by allowing only G and C (IUB mix code: S) at the third position of each codon. Complete permutation of a single amino acid position may thus enable finding nonconservative replacements that are inaccessible by random point mutagenesis.

In this chapter, two approaches are described for the introduction of random point mutations into whole genes: (1) PCR mutagenesis and (2) mutator strain passage. Both procedures involve the cloning of target genes into custom plasmid vectors ready for the functional expression of enzyme variants. Alternatively, mutant gene libraries may be expressed by using commercially available *in vitro* transcription/translation systems. However, this topic is not discussed here.

## 2.2 Materials

### 2.2.1 Materials for Random PCR Mutagenesis

1. Template DNA encoding the gene of interest.
2. Oligonucleotide primers containing the desired restriction sites for cloning.
3. Expression vector with suitable promoter, multiple cloning site, and fusion tag, where applicable (e.g., six-histidine tag).
4. *Taq* DNA polymerase and buffer.
5. Deoxynucleoside triphosphates (10 mM each).
6.  $\text{MnCl}_2$  (100 mM).
7.  $\text{MgCl}_2$  (100 mM).
8. PCR and gel purification (spin) kit.
9. Agarose gel electrophoresis equipment.
10. Restriction endonucleases, alkaline phosphatase, T4 DNA ligase.
11. Competent *E. coli* cells (high quality is required;  $\geq 10^9$  transformants/ $\mu\text{g}$  supercoiled DNA).
12. Luria Bertani (LB) media and appropriate antibiotic.

### 2.2.2 Materials for Mutator Strain Passage

1. Plasmid vector encoding the target gene in a genetic context ready for expression in *E. coli*.
2. Mutator strain: XL1-Red (*mutD*, *mutS*, *mutT*; Stratagene).

3. Plasmid preparation (spin) kit.
4. Amplification strain: XL1–Blue (Stratagene).
5. LB media and appropriate antibiotic.

## 2.3 Protocols

### 2.3.1 Protocol for Random PCR Mutagenesis According to Joyce

A series of parameters is used to substantially increase the overall error frequency of *Taq* DNA polymerase. This enzyme lacks 3'-5' exonuclease activity and exhibits an error rate of  $0.8\text{--}1.1 \times 10^{-4}$  base substitutions/bp of product under standard conditions [15, 16]. The mutagenic PCR conditions include (1) increased  $\text{Mg}^{2+}$  concentration for stabilizing noncomplementary base pairs [17], (2) the addition of  $\text{Mn}^{2+}$  for reducing the base pairing specificity [18], (3) unbalanced dNTP stoichiometry for forcing misincorporation [7], and (4) increased polymerase concentration for enhancing the probability of elongation of misprimed termini [19]. The protocol below largely follows the procedure originally conceived by G. Joyce [8, 20]:

1. Prepare a 10X dNTP mix consisting of 2 mM each of dATP and dGTP and 10 mM each of dCTP and dTTP.
2. Setup a PCR reaction starting with 0.05–0.2 pmol of template DNA, 50 pmol of each primer, 10  $\mu\text{L}$  10X PCR buffer, 10  $\mu\text{L}$  10X dNTP mix, 0.5 mM  $\text{MnCl}_2$ , 5 U *Taq* DNA polymerase, and water to a final volume of 100  $\mu\text{L}$ . The manganese solution should be added just prior to the polymerase (see section 2.4, note 1).
3. Perform PCR cycling following the standard conditions for this template/primer system.
4. Analyze 5  $\mu\text{L}$  of the reaction on an 0.8% agarose gel (see section 2.4, notes 2–4). Usually, the yield of an error-prone PCR reaction is lower than that of a standard PCR; however, one 100  $\mu\text{L}$  reaction will yield  $\approx 1\text{--}2 \mu\text{g}$  of crude PCR product ( $10^{10}\text{--}10^{11}$  molecules). For efficient cloning, 2–5 100- $\mu\text{L}$  reactions should be prepared.
5. Purify the product using a PCR purification (spin) kit.
6. Digest PCR product and vector according to standard protocols [21]. Dephosphorylate the vector using alkaline phosphatase and purify the DNA by agarose gel electrophoresis.
7. Ligate vector and insert, applying at least 3-fold molar excess of the insert (PCR product).
8. Transform competent *E. coli* according to the supplier's manual and cultivate in LB media (plates or liquid cultures, depending on the selection/screening approach; see section 2.4, note 5).

Using DNA fragments of various origins, nucleotide compositions, and lengths (maximum of  $\approx 3$  kb), we observed mutation frequencies of  $0.93 \pm 0.06\%$  per