High-Throughput Screening in Drug Discovery

Edited by
Jörg Hüser

WILEY-VCH Verlag GmbH & Co. KGaA
List of Contents

Preface XIV

List of Contributors XVI

Part I Concept of Screening

1 Chemical Genetics: Use of High-throughput Screening to Identify Small-molecule Modulators of Proteins Involved in Cellular Pathways with the Aim of Uncovering Protein Function
   Sa L Chiang
   1.1 Introduction 1
   1.2 Classical and Chemical Genetics 1
   1.2.1 Forward and Reverse Screens 3
   1.3 Identifying Bioactive Molecules 4
   1.4 Target Identification 5
   1.4.1 Hypothesis-driven Target Identification 5
   1.4.2 Affinity-based Target Identification 6
   1.4.3 Genomic Methods of Target Identification 7
   1.4.4 Proteomic Methods 9
   1.5 Discovery for Basic Research Versus Pharmacotherapy Goals 10
   1.6 Chemical Genetic Screens in the Academic Setting 11
   1.7 Conclusions 12

2 High-throughput Screening for Targeted Lead Discovery
   Jörg Hüser, Emanuel Lohmann, Bernd Kalthof, Nik Burkhardt, Ulf Brüggemeier, and Martin Bechem
   2.1 Chemical Libraries for High-throughput Screening 15
   2.2 Properties of Lead Structures 17
   2.3 Challenges to High-throughput Screening 19
   2.4 Assay Technologies for High-throughput Screening 21
   2.5 Laboratory Automation 24
   2.6 From Target Selection to Confirmed Hits - the HTS Workflow and its Vocabulary 25
List of Contents

2.7 Separating Specific Modulators from Off-Target Effects 29
2.8 Data Analysis and Screening Results 32
2.9 Conclusions 34

Part II Automation Technologies

3 Tools and Technologies that Facilitate Automated Screening
John Comley
3.1 Introduction - the Necessity to Automate 37
3.1.1 Compound Libraries 37
3.1.2 Targets and Data Points 38
3.1.3 Main Issues Facing HTS Groups Today 38
3.1.4 Benefits of Miniaturization 39
3.1.5 Benefits of Automated HTS 39
3.1.6 Screening Strategies 40
3.1.7 Ultra HTS (UHTS) 40
3.2 Sample Carriers 41
3.2.1 A Brief History of the Microplate 41
3.2.2 Microplate Usage Today 41
3.2.3 Microplate Arrays 42
3.2.4 Non-microplate Alternatives 43
3.2.4.1 Labchips 43
3.2.4.2 LabCDs 43
3.2.4.3 LabBrick 44
3.2.4.4 Arrayed Compound Screening 44
3.3 Liquid Handling Tools 45
3.3.1 Main Microplate Dispense Mechanisms 45
3.3.1.1 Pin Tools 45
3.3.1.2 Air and Positive Displacement 45
3.3.1.3 Peristaltic 46
3.3.1.4 Solenoid-syringe 47
3.3.1.5 Solenoid-pressure bottle 47
3.3.1.6 Capillary Sipper 48
3.3.1.7 Piezoelectric 48
3.3.1.8 Acoustic Transducer 48
3.3.2 HTS Liquid Handling Applications and Dispensing Technologies Used 49
3.3.2.1 Bulk Reagent and Cell Addition 49
3.3.2.2 Compound Reformatting and Nanoliter Dispensing 50
3.3.2.3 Cherry Picking and Serial Dilution 51
3.3.2.4 Microplate Washing 52
3.4 Detection Technologies 53
3.4.1 Main Detection Modalities Used in HTS 53
3.4.2 Plate Readers 54
3.4.3 Plate Imagers 55
### List of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4.3.1 Macro-imaging</td>
<td>56</td>
</tr>
<tr>
<td>3.4.3.2 Micro-imaging</td>
<td>57</td>
</tr>
<tr>
<td>3.4.4 Dispense and Read Devices</td>
<td>60</td>
</tr>
<tr>
<td>3.4.5 Other Detection Technologies</td>
<td>60</td>
</tr>
<tr>
<td>3.4.6 Automation of Detection Technologies</td>
<td>61</td>
</tr>
<tr>
<td>3.4.7 Potential Sources of Reading Error</td>
<td>61</td>
</tr>
<tr>
<td>3.5 Laboratory Robotics</td>
<td>62</td>
</tr>
<tr>
<td>3.5.1 Traditional Workstations</td>
<td>64</td>
</tr>
<tr>
<td>3.5.2 Robotic Sample Processors</td>
<td>64</td>
</tr>
<tr>
<td>3.5.3 Plate Storage Devices</td>
<td>64</td>
</tr>
<tr>
<td>3.5.4 Plate Moving Devices</td>
<td>65</td>
</tr>
<tr>
<td>3.5.5 Fully Integrated Robotic Systems</td>
<td>65</td>
</tr>
<tr>
<td>3.5.6 Turnkey Workstations</td>
<td>66</td>
</tr>
<tr>
<td>3.5.7 Automated Cell Culture Systems</td>
<td>66</td>
</tr>
<tr>
<td>3.5.8 Compound Management Systems</td>
<td>67</td>
</tr>
<tr>
<td>3.5.8.1 Current Practice in Compound Management</td>
<td>67</td>
</tr>
<tr>
<td>3.5.8.2 Plate-based versus Tube-based Liquid Compound Storage</td>
<td>68</td>
</tr>
<tr>
<td>3.5.8.3 Associated Automated Instrumentation</td>
<td>70</td>
</tr>
<tr>
<td>3.5.8.4 Sample Integrity and QC Testing</td>
<td>70</td>
</tr>
</tbody>
</table>

### Part III Assay Technologies

#### 4 Functional Cell-based Assays for Targeted Lead Discovery in High-throughput Screening

**Jörg Hüser, Bernd Kalthof, and Jochen Strayle**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Introduction</td>
<td>75</td>
</tr>
<tr>
<td>4.2 Reporter Gene Technologies</td>
<td>78</td>
</tr>
<tr>
<td>4.3 Membrane Potential Indicators</td>
<td>82</td>
</tr>
<tr>
<td>4.4 Ca(^{2+}) Indicators</td>
<td>88</td>
</tr>
<tr>
<td>4.5 Conclusions</td>
<td>90</td>
</tr>
</tbody>
</table>

#### 5 Biochemical Assays for High-throughput Screening

**William D. Mallender, Michael Bembenek, Lawrence R. Dick, Michael Kuranda, Ping Li, Saurabh Menon, Eneida Pardo, and Tom Parsons**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 General Considerations for Biochemical High-throughput Screening</td>
<td>93</td>
</tr>
<tr>
<td>5.2 Expression and Purification of Recombinant Enzymes</td>
<td>95</td>
</tr>
<tr>
<td>5.2.1 Design of Expression Constructs</td>
<td>98</td>
</tr>
<tr>
<td>5.2.2 Expression Assessment and Optimization</td>
<td>99</td>
</tr>
<tr>
<td>5.2.3 Purification</td>
<td>99</td>
</tr>
<tr>
<td>5.3 Peptidases</td>
<td>102</td>
</tr>
<tr>
<td>5.3.1 Application of Fluorogenic Substrates to Configure Peptidase Screens</td>
<td>102</td>
</tr>
<tr>
<td>5.3.2 The Value of Continuous Assays</td>
<td>107</td>
</tr>
<tr>
<td>5.4 Oxidoreductases</td>
<td>107</td>
</tr>
<tr>
<td>5.4.1 NAD(P)-dependent Oxidoreductases</td>
<td>108</td>
</tr>
</tbody>
</table>
List of Contents

5.4.2 Non-NAD(P) Cofactor-dependent Oxidoreductases 110
5.4.3 Oxidases or Oxygen-utilizing Oxidoreductases 111
5.4.4 General Considerations 113
5.5 Transferases, Synthetases and Lipid-modifying Enzymes 114
5.5.1 Streptavidin-Biotin Capture 114
5.5.2 Ionic Capture 117
5.5.3 Hydrophobic Capture 117
5.6 Kinases 120
5.6.1 Streptavidin-Biotin Capture 120
5.6.2 Homogeneous Time-resolved Fluorescence (HTRF) 122
5.6.3 Pyruvate Kinase-Lactate Dehydrogenase Assay System 122
5.7 Pitfalls and Reasons for Assay Development Failures 125

6 Image-based High-content Screening - A View from Basic Sciences
Peter Lipp and Lars Kaestner
6.1 Introduction 129
6.2 HCS Systems Employing Confocal Optical Technologies 132
6.3 Single-point Scanning Technology 134
6.4 Line Scanning Technology 136
6.5 Multi-beam Technology 138
6.6 Structured Illumination 145
6.7 Summary and Perspectives 147

Part IV Data Analysis

7 Methods for Statistical Analysis, Quality Assurance and Management of Primary High-throughput Screening Data
Hanspeter Cubler
7.1 Introduction 151
7.1.1 Overview 151
7.1.2 Problems during the Analysis of Primary HTS Data 151
7.2 Statistical Considerations in Assay Development 155
7.3 Data Acquisition, Data Preprocessing, and HTS Data Analysis Environment 158
7.4 Data Normalization 160
7.5 Robust Statistics in HTS Data Analysis 163
7.5.1 The General Problem 163
7.5.2 Threshold Setting - a Simple Model 164
7.5.3 Threshold Setting - a Complex Model 165
7.5.4 The Most Important Robust Estimation Methods for Data Summaries 165
7.5.5 An Illustrative Example: Performance of Location and Scale Estimators on Typical HTS Data Activity Distributions 167
7.5.6 A Robust Outlier Detection Method 168
7.5.7 Outlier-resistant Versions of Simple HTS Data Quality Indicators 170
List of Contents

7.6 Measures of HTS Data Quality, Signaling of Possible QC Problems, Visualizations 170

7.6.1 Trends and Change Points 171

7.6.2 Positional Effects - Summary Statistics and Views 175

7.6.3 Positional Effects - Heat Maps, Trellis Plots, and Assay Maps 177

7.6.4 Distribution Densities - Histograms, Smoothed Distributions 180

7.6.5 Numerical Diagnostics and Fully Automated QC Assessment 181

7.6.6 Possible Sources of Systematic Errors and Trends 182

7.7 Correction of Position-dependent Response Effects 182

7.7.1 The General Problem 182

7.7.2 Plate Averaging (Multiple Plates) 184

7.7.3 Median Polish Smoothing (Single Plates, Multiple Plates) 185

7.7.4 Change Point Detection (Multiple Plates, Plate Sequences) 186

7.7.5 Parametric (Polynomial) Surface Fitting (Single Plates, Multiple Plates) 187

7.7.6 Nonparametric Surface Fitting, Local Regression, and Smoothing (Single Plates, Multiple Plates) 189

7.7.7 Expansion into Orthogonal Basis Functions (Single Plates) 190

7.7.8 Empirical Orthogonal Function (EOF) Analysis, Singular Value Decomposition (SVD) (Multiple Plates) 191

7.7.9 Some Remarks on the Correction of Individual Plates versus Complete Plate Sets 194

7.7.10 Position-dependent Correction of Background Response Surface 195

7.8 Hit Identification and Hit Scoring 199

7.9 Conclusion 201

8 Chemoinformatic Tools for High-throughput Screening Data Analysis

Peter C. Nell and Stefan M. Mundt

8.1 Introduction 207

8.1.1 Definition of Chemoinformatics 208

8.1.2 High-throughput Screening 208

8.1.2.1 Random Screening 209

8.1.2.2 Sequential Screening 209

8.2 Workflow of High-throughput Screening and Use of Chemoinformatics 211

8.3 Chemoinformatic Methods Used in HTS Workflow 214

8.3.1 Substructure Search/Similarity Search 214

8.3.1.1 Structural Descriptors/Fingerprints 215

8.3.1.2 Measures of Similarity 217

8.3.2 Clustering 218

8.3.2.1 Hierarchical Clustering 219

8.3.2.2 Nonhierarchical Clustering 219

8.3.2.3 Partitioning Methods 220

8.3.4 Principal Components Analysis 221

8.3.3 Maximum Common Substructure - Distill 222
8.3.4 Mode of Action-Profiling 223
8.3.5 Artificial Neural Networks (ANNs) 224
8.3.6 Decision Trees/Recursive Partitioning 226
8.3.7 Reduced Graph-based Methods 226
8.3.7.1 Reduced Graph Theory Analysis: Baytree 227
8.3.8 Fragment-based Methods - Structural Units Analysis 230
8.4 Chemoinformatic Methods in the Design of a Screening Library 231
8.4.1 Drug and Lead Likeness 231
8.4.2 ADME Parameters 233
8.4.2.1 Absorption 233
8.4.2.2 Distribution 234
8.4.2.3 Metabolism 234
8.4.2.4 Excretion 235
8.4.2.5 Toxicity 235
8.4.3 Diversity 236
8.5 Integrated Software Packages 238
8.5.1 Commercially Available Packages 239
8.5.1.1 Accelrys: DIVA® 239
8.5.1.2 BioSolvIT: HTSview 239
8.5.1.3 SciTegic: Pipeline Pilot™ 240
8.5.1.4 Bioreason: ClassPharmer™ Suite 241
8.5.1.5 Spotfire Lead Discovery 242
8.5.1.6 LeadScope 243
8.5.1.7 OmniViz 244
8.5.1.8 SARNavigator 245
8.5.2 In-house Packages 246
8.6 Conclusions 247

9 Combinatorial Chemistry and High-throughput Screening
Roger A. Smith and Nils Criebenow
9.1 Introduction 259
9.2 Categories of Compound Libraries for High-throughput Screening 260
9.3 Synthesis Techniques and Library Formats 261
9.3.1 Solid-phase Synthesis 262
9.3.1.1 Parallel Solid-phase Synthesis Techniques and Tools 265
9.3.1.2 Pool/Split Techniques with Encoding 265
9.3.2 Solution-phase Synthesis 269
9.3.2.1 Polymer-supported Reagents and Scavengers 269
9.3.2.2 Extraction Techniques for Purification 271
9.3.2.3 Purification by Chromatography 273
9.3.3 Library Formats 273
9.3.3.1 One-bead One-compound Libraries 274
9.3.3.2 Pre-encoded Libraries 275
9.3.3.3 Spatially Addressable Libraries 276
9.4 Library Design and Profiling Approaches 277
List of Contents

9.5 Impact of Combinatorial Libraries on Drug Discovery 277
9.5.1 Lead Identification 278
9.5.2 Lead Optimization 283
9.5.3 Clinical Drug Candidates 286
9.6 Conclusion 290

10 High-throughput Screening and Data Analysis
Jeremy S. Caldwell and Jeff Janes

10.1 Introduction 297
10.2 Analysis of Cellular Screening Data 298
10.2.1 Quality Control and Analysis 298
10.2.2 Enrichment for Hits 301
10.2.3 Meta-data Analysis 302
10.3 Massively Parallel Cellular Screens 305
10.3.1 Data Analysis of Multidimensional Datasets 307
10.3.2 Multidimensional Cellular Profiling for MOA prediction 311
10.3.3 Cellular Profiling in Lead Exploration 314
10.4 Systematic Serendipity 317
10.5 Conclusion 320

Appendix 323

Index 333