CONTENTS

Preface ***1***

Introduction ***3***

FIRST PART
AUTOMATED PHARMACOLOGICAL SCREENING

Chapter 1 - The pharmacological screening process: the small molecule, the biological target, the robot, the signal and the information ***7***

Eric MARÉCHAL - Sylvaine ROY - Laurence LAFANECHÈRE

1.1. Introduction 7

1.2. The screening process: technological outline 8
- 1.2.1. Multi-well plates, robots and detectors 8
- 1.2.2. Consumables, copies of chemical libraries and storage 10
- 1.2.3. Test design, primary screening, hit-picking, secondary screening 10

1.3. The small molecule: overview of the different types of chemical library 12
- 1.3.1. The small molecule 12
- 1.3.2. DMSO, the solvent for chemical libraries 12
- 1.3.3. Collections of natural substances 12
- 1.3.4. Commercial and academic chemical libraries 14

1.4. The target, an ontology to be constructed 14
- 1.4.1. The definition of a target depends on that of a bioactivity 14
- 1.4.2. Duality of the target: molecular entity and biological function 15
- 1.4.3. An ontology to be constructed 16

1.5. Controls 17

1.6. A new discipline at the interface of biology, chemistry and informatics: chemogenomics 17

1.7. Conclusion 18

1.8. References 19

Chapter 2 - Collections of molecules for screening: example of the French National Chemical Library ***23***

Marcel HIBERT

2.1. Introduction 23

2.2. Where are the molecules to be found? 25

2.3. State of progress with the European Chemical Library 27

2.4. Perspectives 28

2.5. References 28

Chapter 3 - The miniaturised biological assay: constraints and limitations 29
Martine KNIBIEHLER
3.1. Introduction 29
3.2. General procedure for the design and validation of an assay 30
3.2.1. Choice of assay 31
3.2.2. Setting up the assay 33
3.2.3. Validation of the assay and automation 35
3.3. The classic detection methods 36
3.4. The results 36
3.4.1. The signal measured: increase or decrease? 36
3.4.2. The information from screening is managed on three levels 37
3.4.3. Pharmacological validation 40
3.5. Discussion and conclusion 40
3.6. References 41

Chapter 4 - The signal: statistical aspects, standardisation, elementary analysis 43
Samuel WIECZOREK
4.1. Introduction 43
4.2. Normalisation of the signals based on controls 44
4.2.1. Normalisation by the percentage inhibition 44
4.2.2. Normalisation resolution 44
4.2.3. Aberrant values 46
4.3. Detection and correction of measurement errors 48
4.4. Automatic identification of potential artefacts 49
4.4.1. Singularities 49
4.4.2. Automatic detection of potential artefacts 50
4.5. Conclusion 52
4.6. References 52

Chapter 5 - Measuring bioactivity: Ki, IC50 and EC50 55
Eric MARÉCHAL
5.1. Introduction 55
5.2. Prerequisite for assaying the possible bioactivity of a molecule: the target must be a limiting factor 55
5.3. Assaying the action of an inhibitor on an enzyme under Michaelian conditions: Ki 56
5.3.1. An enzyme is a biological catalyst 57
5.3.2. Enzymatic catalysis is reversible 57
5.3.3. The initial rate, a means to characterise a reaction 59
5.3.4. Michaelian conditions 59
5.3.5. The significance of Km and Vmax in qualifying the function of an enzyme . 60
5.3.6. The inhibited enzyme: Ki 60
5.4. Assaying the action of a competitive inhibitor upon a receptor: IC50 62
5.5. Relationship between Ki and IC50: the CHENG-PRUSOFF equation 63
5.6. EC50: a generalisation for all molecules generating a biological effect (bioactivity) 64
5.7. Conclusion 64
5.8. References 65

Chapter 6 - Modelling the pharmacological screening: controlling the processes and the chemical, biological and experimental information..... 67
Sylvaine ROY
6.1. Introduction 67
6.2. Needs analysis by modelling 68
6.3. Capture of the needs 69
6.4. Definition of the needs and necessity of a vocabulary common to biologists, chemists and informaticians 69
6.5. Specification of the needs 69
6.5.1. Use cases and their diagrams 70
6.5.2. Activity diagrams 72
6.5.3. Class diagrams and the domain model 73
6.6. Conclusion 78
6.7. References 78

Chapter 7 - Quality procedures in automated screening 79
Caroline BARETTE
7.1. Introduction 79
7.2. The challenges of quality procedures 79
7.3. A reference guide: the ISO 9001 Standard 80
7.4. Quality procedures in five steps 82
7.4.1. Assessment 82
7.4.2. Action plan - planning 83
7.4.3. Preparation 83
7.4.4. Implementation 83
7.4.5. Monitoring 83
7.5. Conclusion 84
7.6. References 84

SECOND PART
HIGH-CONTENT SCREENING AND THE STRATEGIES FOR CHEMICAL GENETICS

Chapter 8 - Phenotypic screening of cells and the strategies for direct chemical genetics.. 87
Laurence LAFANECHÈRE
8.1. Introduction 87
8.2. The traditional genetics approach: from phenotype to gene and from gene to phenotype 88
8.2.1. Phenotype 88
8.2.2. Forward and reverse genetics 89
8.3. Chemical genetics 89
8.4. Chemical libraries for chemical genetics 90
8.4.1. Chemical library size 91
8.4.2. Concentration of molecules 91
8.4.3. Chemical structure diversity 91
8.4.4. Complexity of molecules 93

8.4.5. Accessibility of molecules to cellular compartments 93
8.4.6. The abundance of molecules 94
8.4.7. The possibility of functionalizing the molecules 94
8.5. Phenotypic tests with cells 94
8.6. Methods to identify the target 96
8.7. Conclusions 99
8.8. References 99

Chapter 9 - High information content screens for forward (phenotypic screening of organisms) and reverse (structural screening by NMR) chemical genetics 103
Benoît DÉPREZ
9.1. Introduction 103
9.2. Benefits of high-content screening 104
9.2.1. Summarised comparison of high-throughput screening and high-content screening 104
9.2.2. Advantages of high-content screening for the discovery of novel therapeutic targets 104
9.2.3. The nematode *Caenorhabditis elegans*: a model organism for high-content screening 105
9.2.4. Advantages of high-content screening for reverse chemical genetics and the discovery of novel bioactive molecules 108
9.3. Constraints linked to throughput and to the large numbers 110
9.3.1. Know-how 110
9.3.2. Miniaturisation, rate and robustness of the Assays 110
9.3.3. Number, concentration and physicochemical properties of small molecules . 111
9.4. Types of measurement for high-content screening 111
9.4.1. The critical information needed for screening 111
9.4.2. Raw, numerical results 111
9.4.3. Results arising from expert analyses 112
9.5. Conclusion 112
9.6. References 112

Chapter 10 - Some principles of Diversity-Oriented Synthesis 113
Yung-Sing WONG
10.1. Introduction 113
10.2. Portrait of the small molecule in DOS 114
10.3. Definition of the degree of diversity (DD) 116
10.3.1. Degree of diversity of the building block 116
10.3.2. Degree of stereochemical diversity 118
10.3.3. Degree of regiochemical diversity 119
10.3.4. Degree of skeletal diversity 121
10.4. Divergent multi-step DOS by combining elements of diversity 124
10.5. Convergent DOS: condensation between distinct small molecules 127
10.6. Conclusion 130
10.7. References 130

THIRD PART
TOWARDS AN IN-SILICO EXPLORATION OF CHEMICAL AND BIOLOGICAL SPACE

Chapter 11 - Molecular descriptors and similarity indices 143
Samia ACI
11.1. Introduction 143
11.2. Chemical formulae and computational representation 144
11.2.1. The chemical formula: a representation in several dimensions 145
11.2.2. Molecular information content 145
11.2.3. Molecular graph and connectivity matrix 147
11.3. Molecular descriptors 148
11.3.1. 1D descriptors 149
11.3.2. 2D descriptors 149
11.3.3. 3D descriptors 152
11.3.4. 3D *versus* 2D descriptors? 154
11.4. Molecular similarity 155
11.4.1. A brief history 155
11.4.2. Properties of similarity coefficients and distance indices 155
11.4.3. A few similarity coefficients 156
11.5. Conclusion 156
11.6. References 158

Chapter 12 - Lipophilicity of molecules: a predominant descriptor for QSAR 161
Gérard GRASSY - Alain CHAVANIEU
12.1. Introduction 161
12.2. History 161
12.3. Theoretical foundations and principles of the relationship between the structure of a small molecule and its bioactivity 162
12.3.1. QSAR, QPAR and QSPR 162
12.3.2. Basic equation of a QSAR study 163
12.4. Generalities about lipophilicity descriptors 163
12.4.1. Solubility in water and in lipid phases: conditions for bioavailability 163
12.4.2. Partition coefficients 164
12.4.3. The partition coefficient is linked to the chemical potential 164
12.4.4. Thermodynamic aspects of lipophilicity 165
12.5. Measurement and estimation of the octanol/water partition coefficient 166
12.5.1. Measurement methods 166
12.5.2. Prediction methods 168
12.5.3. Relationship between lipophilicity and solvation energy: LSER 171
12.5.4. Indirect estimation of partition coefficients from values correlated with molecular lipophilicity 171
12.5.5. Three-dimensional approach to lipophilicity 173
12.6. Solvent systems other than octanol/water 174
12.7. Electronic parameters 175
12.7.1. The HAMMETT parameter, σ 175
12.7.2. SWAIN and LUPTON parameters 176

12.8. Steric descriptors 177
12.9. Conclusion 177
12.10. References 177

Chapter 13 - The annotation and classification of chemical space in chemogenomics 179
Dragos HORVATH
13.1. Introduction 179
13.2. From the medicinal chemist's intuition to a formal treatment of structural information 179
13.3. Mapping structural space: predictive models 182
13.3.1. Mapping structural space 182
13.3.2. Neighbourhood (similarity) models 183
13.3.3. Linear and non-linear empirical models 186
13.4. Empirical filtering of drug candidates 188
13.5. Conclusion 189
13.6. References 189

Chapter 14 - The annotation and classification of biological space in chemogenomics 193
Jordi MESTRES
14.1. Introduction 193
14.2. Receptors 194
14.2.1. Definitions 194
14.2.2. Establishing the 'RC' nomenclature 195
14.2.3. Ion-channel receptors 196
14.2.4. G protein-coupled receptors 197
14.2.5. Enzyme receptors 198
14.2.6. Nuclear receptors 198
14.3. Enzymes 199
14.3.1. Definitions 199
14.3.2. The 'EC' nomenclature 200
14.3.3. Specialised nomenclature 201
14.4. Conclusion 201
14.5. References 202

Chapter 15 - Machine learning and screening data 205
Gilles BISSON
15.1. Introduction 205
15.2. Machine learning and screening 207
15.3. Steps in the machine-learning process 210
15.3.1. Representation languages 211
15.3.2. Developing a training set 213
15.3.3. Model building 214
15.3.4. Validation and revision 215
15.4. Conclusion 217
15.5. References and internet sites 217

Chapter 16 - Virtual screening by molecular docking 221
Didier ROGNAN
16.1. Introduction 221
16.2. The 3 steps in virtual screening 221
16.2.1. Preparation of a chemical library 221
16.2.2. Screening by high-throughput docking 224
16.2.3. Post-processing of the data 226
16.3. Some successes with virtual screening by docking 228
16.4. Conclusion 229
16.5. References 230

APPENDIX
BRIDGING PAST AND FUTURE

Chapter 17 - Biodiversity as a source of small molecules for pharmacological screening: libraries of plant extracts 235
Françoise GUERITTE, Thierry SEVENET, Marc LITAUDON, Vincent DUMONTET
17.1. Introduction 235
17.2. Plant biodiversity and North-South co-development 237
17.3. Plant collection: guidelines 238
17.4. Development of a natural-extract library 239
17.4.1. From the plant to the plate 239
17.4.2. Management of the extract library 239
17.5. Strategy for fractionation, evaluation and dereplication 241
17.5.1. Fractionation and dereplication process 241
17.5.2. Screening for bioactivities 243
17.5.3. Some results obtained with specific targets 244
17.5.4. Potential and limitations 247
17.6. Conclusion 248
17.7. References 248

Glossary 249

Authors 261