

Contents

1	Introduction	1
	Jaroslav Koča, Radka Svobodová Vařeková, Lukáš Pravda, Karel Berka, Stanislav Geidl, David Sehnal and Michal Otyepka	
	References.	4
 Part I Patterns, Fragments and Data Sources		
2	Biomacromolecular Fragments and Patterns	7
	Lukáš Pravda	
2.1	Pattern Examples	8
2.1.1	Active Site and Their Inhibition – Cyclooxygenase Inhibitors.	8
2.1.2	Allosteric Site – Structural Flexibility of HIV Protease.	9
2.1.3	Transcription Factor – Zinc Finger Motif	9
2.2	Pattern Prediction	10
2.2.1	Ubiquitin-Binding Domain Prediction	11
2.2.2	Pattern Detection	12
2.2.3	Phosphorylation of Drug Binding Pockets.	12
	References.	13
3	Structural Bioinformatics Databases of General Use	17
	Karel Berka	
3.1	How a Biomacromolecule Looks Codes What It Does.	17
3.2	Worldwide Protein Data Bank (PDB) – Essential Structure Repository	19
3.2.1	Protein Data Bank in Europe (PDBe)	20
3.2.2	RCSB PDB	22
3.3	Other Notable Databases.	23
3.3.1	PDBsum – Pictorial View on PDB Database.	23
3.3.2	PDB_REDO and WHY_NOT Databases for Curated Structures	23

3.3.3	CATH and Pfam Databases for Classification of Protein Folds and Sequences.	23
3.3.4	PDB Flex, Pocketome and PED3 Databases to Analyze Protein Flexibility and Disorder.	24
3.3.5	OPM and MemProtMD Databases for Membrane Protein.	25
3.3.6	NDB and GFDB Databases for Other Macromolecules.	25
3.3.7	UniProt and ChEMBL Databases – Power of Connection	26
3.4	Conclusion	27
3.5	Exercises.	27
3.5.1	Use of PDBe.	27
3.5.2	Use of RCSB and ChEMBL	28
3.5.3	Use of PDBsum	28
3.5.4	Use of CATH	28
	References.	29
4	Validation	31
	Radka Svobodová Vařeková, David Sehnal, Lukáš Pravda, Stanislav Geidl and Jaroslav Koča	
4.1	Introduction and Motivation	31
4.2	Nipah G Attachment Glycoprotein Validation Example	32
4.3	Objects of Validation	33
4.4	Source Data for Validation	34
4.5	Validation Approaches	34
4.6	Evolution of Validation Tools.	35
4.7	How to Handle Structures with Errors	35
4.8	Exercises.	36
	References.	38
 Part II Detection and Extraction		
5	Detection and Extraction of Fragments	43
	Lukáš Pravda, David Sehnal, Radka Svobodová Vařeková and Jaroslav Koča	
5.1	PatternQuery.	43
5.1.1	PatternQuery Explained.	44
5.1.2	Thinking in PatternQuery	45
5.1.3	Basic Principles of the Language.	46
5.2	MetaPocket 2.0.	51
5.2.1	Serotonin Receptor Example	51
5.3	Note on Pattern Comparison.	52

5.4	Exercises.	53
5.4.1	PatternQuery	53
5.4.2	MetaPocket	55
	References.	56
6	Detection of Channels	59
	Lukáš Pravda, Karel Berka, David Sehnal, Michal Otyepka, Radka Svobodová Vařeková and Jaroslav Koča	
6.1	Introduction and Motivation	59
6.1.1	Bunyavirus Polymerase Example.	62
6.1.2	Aquaporin Example.	63
6.2	MOLE - Channel Analysis Tool.	64
6.3	Identification of Channels Using MOLEonline.	64
6.3.1	Setup.	64
6.3.2	Geometry Properties	65
6.4	Exercises.	67
	References.	67
 Part III Characterization		
7	Characterization via Charges	73
	Radka Svobodová Vařeková, David Sehnal, Stanislav Geidl and Jaroslav Koča	
7.1	Introduction and Motivation	73
7.2	Dinitrotoluene Example	73
7.3	Charge Calculation Approaches	74
7.4	Charge Visualization.	75
7.5	Formats for Saving of Charges.	77
7.6	Exercises.	77
	References.	79
8	Channel Characteristics	81
	Lukáš Pravda, Karel Berka, David Sehnal, Michal Otyepka, Radka Svobodová Vařeková and Jaroslav Koča	
8.1	Physicochemical Properties.	81
8.1.1	Hydropathy	81
8.1.2	Polarity	82
8.1.3	Mutability	82
8.1.4	Charge.	83
8.2	Characterization of Channels Using MOLEonline	84
8.2.1	Results Analysis	84
8.3	Common Errors in Channel Calculation and Characterization	87
8.3.1	No Channels Have Been Identified	87

8.3.2	A Lot of Different Channels Are Identified, However None of Them Seems to be Relevant to My Expectations	89
8.4	Exercises.	90
	References.	90

Part IV Complete Process of Data Extraction and Analysis

9	Complete Process of Data Extraction and Analysis	93
	Radka Svobodová Vařeková and Karel Berka	
9.1	Lectin Example (Validation, Extraction, Comparison, Charge Calculation)	93
9.1.1	Step 1: Detection of All Occurrences of the Binding Site	93
9.1.2	Step 2: Validation of the Obtained PDB Entries	95
9.1.3	Step 3: Analysis of Organisms and Proteins, from Which the Obtained Binding Sites Originate	95
9.1.4	Step 4: Analysis of Common Amino Acid Composition	96
9.1.5	Step 5: Analysis of Common 3D Structure Parts.	97
9.1.6	Step 6: Analysis of Charge Distribution	98
9.1.7	Methodology of Data Analysis	99
9.2	Cytochrome P450 Example (Database Search, Detection of Channels, Channel Characterization)	100
9.2.1	Database Search	101
9.2.2	Channels Detection	102
9.2.3	Channels Characterization	102
9.2.4	Solution.	102

Part V Conclusion

10	Concluding Remarks.	111
	Jaroslav Koča, Radka Svobodová Vařeková, Lukáš Pravda, Karel Berka, Stanislav Geidl, David Sehnal and Michal Otyepka	
11	Exercises Solution	113
	Jaroslav Koča, Radka Svobodová Vařeková, Lukáš Pravda, Karel Berka, Stanislav Geidl, David Sehnal and Michal Otyepka	
11.1	Structural Bioinformatics Databases of General Use.	113
11.2	Validation.	121
11.3	Detection and Extraction of Fragments.	125
11.3.1	PatternQuery	125
11.3.2	MetaPocket	129
11.4	Detection of Channels.	133
11.5	Characterization via Charges.	134

Contents	xi
11.6 Channel Characteristics.....	138
References.....	139
Glossary	141
Index	143

Structural Bioinformatics Tools for Drug Design
Extraction of Biologically Relevant Information from
Structural Databases

Koča, J.; Svobodová Vařeková, R.; Pravda, L.; Berka, K.;

Geidl, S.; Sehnal, D.; Otyepka, M.

2016, XIII, 144 p. 62 illus., Softcover

ISBN: 978-3-319-47387-1