

Contents

Zusammenfassung.....	1
Summary.....	3
Chapter 1: General introduction	5
1.1 Protein structure and folding	5
1.2 Experimental and computational determination of protein structure	7
1.3 Protein binding and the underlying physicochemical principles.....	10
1.4 Post-translational modification of proteins	12
1.5 Aims and organization of the thesis	15
Chapter 2: Protein-compound interactions with respect to binding specificity.....	17
Abstract.....	17
2.1 Introduction	18
2.2 Material and Methods.....	20
2.2.1 Compound- protein target datasets.....	20
2.2.2 Compound classification and property calculation	21
2.2.3 Compound-promiscuity propensity ratio calculation	22
2.2.4 Amino acid residue compositional propensities of protein binding pockets.....	22
2.2.5 Enzyme classification entropy and pocket variability analysis.....	23
2.2.6 Binding mode prediction models	23
2.2.7 Metabolite pathway, process, and organismal systems enrichment analysis.....	24
2.3 Results	24
2.3.1 Compound-protein target dataset	24
2.3.2 Physicochemical properties of metabolites and drugs bound to proteins	26
2.3.3 Characterization of compound binding promiscuity	28
2.3.4 LogP and compound binding promiscuity	30
2.3.5 Protein target-centric investigation of binding events.....	31
2.3.6 Functional and compositional diversity of target proteins and binding sites.....	33
2.3.7 Prediction of compound promiscuity using physicochemical properties.....	37
2.3.8 Metabolite pathway, process, and organismal systems enrichment analysis.....	40
2.4 Discussion	41

Chapter 3: Regulatory role of phosphorylation on compound binding..... 45

Abstract..... 45

3.1 Introduction 46

3.2 Materials and Methods 48

3.2.1 Phosphorylation site and compound binding site datasets 48

3.2.2 Compound subsets: metabolites and drugs 49

3.2.3 Spatial relationships between C-sites and P-sites 50

3.2.4 Randomization of P-sites 51

3.2.5 Compound structures and physicochemical properties 52

3.2.6 Other resources and methods 52

3.3 Results 52

3.3.1 Protein structure dataset with compound binding sites and phosphorylation sites..... 52

3.3.2 Non-random spatial distance distributions of C- and P-sites 53

3.3.3 Serine/ threonine vs. tyrosine P-site distributions relative to C-sites differ 55

3.3.4 P-sites are enriched around enzyme catalytic sites 56

3.3.5 P-site distributions differ depending on net-charge of bound compounds..... 57

3.3.6 Observations independent of potentially co-varying variables protein size and compound size, partial exception for P-site density..... 60

3.4 Discussion 62

Chapter 4: Spatial investigation of homo- and heterotypic post-translational modifications 67

Abstract..... 67

4.1 Introduction 68

4.2 Materials and Methods 71

4.2.1 Post-translational modifications and protein structures 71

4.2.2 Redundancy reduction..... 72

4.2.3 Co-occurrences..... 72

4.2.4 Homo- and heterotypic PTM site distances 72

4.2.5 Randomization of PTM sites..... 73

4.2.6 Multidimensional scaling..... 74

4.2.7 Protein shape factor..... 74

4.2.8 Other resources 74

4.3 Results 74

4.3.1 Protein structures and PTMs 74

4.3.2 Co-occurrence of PTM types 76

4.3.3	PTM site distances on protein surfaces	77
4.3.4	A global map of PTM locations on protein surfaces.....	83
4.3.5	Protein size and shape independence of PTM pair distributions.....	85
4.4	Discussion	86
Chapter 5:	General discussion	89
5.1	High-resolution protein structures from the PDB.....	89
5.2	Redundancy reduction of data	91
5.3	Physicochemical characteristics of protein-compound interactions.....	92
5.4	Mutual structural and functional relationships	94
5.5	Conformational changes of protein structure	98
5.6	Concluding remarks	99
Chapter 6:	Supplementary information	101
	Supplementary information to Chapter 2: “Protein-compound interactions with respect to binding specificity”	101
	Supplementary information to Chapter 3: “Regulatory role of phosphorylation on compound binding”	111
	Supplementary information to Chapter 4: “Spatial investigation of homo- and heterotypic post-translational modifications”	121
	List of figures and tables	125
	List of abbreviations	129
	Bibliography	131
	Eidesstattliche Erklärung.....	141