Structural analysis and computer-based design of novel inhibitors for human protein-kinase-C-related enzyme PRK1

Dissertation

zur Erlangung des
Doktorgrades der Naturwissenschaften (Dr. rer. nat.)
der Naturwissenschaftlichen Fakultät I Biowissenschaften
der Martin-Luther-Universität Halle-Wittenberg,

vorgelegt von
Frau Inna Slynko
geb. am 25.10.1986 in Cherkassy, Ukraine

Gutachter/in:

1. Prof. Dr. Wolfgang Sippl, Halle
2. Prof. Dr. Mike Schutkowski, Halle
3. Prof. Dr. Gerhard Wolber, Berlin

Halle (Saale), 2015.06.16

Abstract

Protein kinase C-related kinase 1 (PRK1) has been shown to be involved in the regulation of androgen receptor signaling and has been identified as a novel potential drug target for prostate cancer therapy. Due to the absence of the PRK1 crystal structure multiple PRK1 homology models were generated in order to address the problem of protein flexibility. An in-house library of compounds tested on PRK1 was docked into generated homology models. In most cases the correct pose of the inhibitor could be identified by ensemble docking, while there was still a challenge of finding a reasonable scoring function able to rank compounds according to their biological activity. Thus, we estimated the binding free energy for our dataset of structurally diverse PRK1 inhibitors using the MM-PB(GB)SA and QM/MM-GBSA methods after energy minimization in implicit solvent. The QSAR model was designed on the basis of calculated binding free energy scores in order to predict the pIC_{50} values of novel compounds. The obtained results demonstrate that a good correlation between calculated and experimental pIC_{50} values could be derived. Furthermore, the developed approach was tested on a set of diverse PRK1 inhibitors taken from literature, which resulted in a significant correlation. The method is computationally inexpensive and can be applied as post-docking filter in virtual screening as well as for optimization of PRK1 inhibitors. Next, the developed approach was applied in a prospective way for the virtual screening of two external datasets containing compounds which were not tested on PRK1 previously. A number of highly potent PRK1 inhibitors were identified by this screening, proving the outstanding performance of the method. Finally, four crystal structures of PRK1 that were recently released were used to evaluate the accuracy of the predicted homology model of PRK1. The comparison shows a high similarity of the experimental data to the PRK1 homology models, especially the inhibitor binding was accurately modeled. The experimental data further support our findings about the PRK1 flexibility.

Key words: kinase, PRK1, epigenetics, prostate cancer, homology model, docking, virtual screening, QSAR, MM-PBSA, MM-GBSA

Acknowledgments

Foremost, I would like to express my sincerest gratitude to my supervisor Prof. Dr. Wolfgang Sippl for continuous support and guidance during all period of my PhD study - for responding very quickly to my application and to all my further questions, for accepting me as a part of his research group and sharing his immense knowledge, experience and enthusiasm during the course of my thesis, for assisting me in internship and job applications, for reading and correcting all the reports, articles and thesis, and of course for being kind, friendly and patient Doktorvater, who always finds time for his students.

I would like to thank our collaborators from the Institute of Pharmaceutical Sciences in Freiburg, who contributed a lot to our project - for Prof. Jung and his research group, especially for Tobias Rumpf for conducting numerous in vitro tests for PRK1.

My sincere thanks also goes to Prof. Dr. Gerhard Ecker and all the members of EUROPIN consortium for providing the research activities and comprehensive training for PhD students and for the opportunity to establish contacts with many professionals in the field of drug discovery.

I am also very grateful for Prof. Dr. Hillisch and Dr. Andreas Göller for the chance to gain the experience in pharmaceutical industry and to expand my knowledge in the area of computational chemistry during my internship at Bayer HealthCare.

My special thanks go to our Medicinal Chemistry group in Institute of Pharmacy in Halle. I am very grateful to Jelena Melesina and Berin Karaman for checking the thesis and to all colleagues who enriched my life on a personal and professional level. We shared not only scientific discussions, but also many memorable moments - grilling or canoeing together, traveling to Freyburg and Freiburg, spending bright Christmas and defense parties. It was great to meet so many intelligent and nice people from around the world, to try exceptional Turkish, Vietnamese, Thai, Egyptian, Syrian, Indian, German or Italian sweets and other specialties. Finally, I would like to express my exceptional gratitude to Jelena Melesina for being my “authorized person” and helping with different organizational issues, for supporting me in good and bad times and of course for sharing great “sugar parties” full of useful fruits and cakes.

Last but not the least, I would like to express my sincere appreciation to my parents and the family in Ukraine for all the support they have provided me over the years, and to my dear people in Germany - family Hartwig who made me feel like at home in a foreign country and with whom I have spent almost every German holiday.

It was my pleasure to follow this way and to spend the years of PhD exactly in this country, in this city and to be surrounded by these special people.

Contents

1 Introduction **1**

1.1 Kinase 1

1.1.1 Kinases and Their Role in Signal Transduction 1

1.1.2 Kinase Structure 2

1.1.3 Types of Kinase Inhibitors 4

1.1.3.1 Type I Inhibitors (ATP-Competitive) 5

1.1.3.2 Type II Inhibitors 7

1.1.3.3 Type III Inhibitors 7

1.2 Protein Kinase C Related Kinase 1 8

1.2.1 Structure of PRK1 8

1.2.2 PRK1 Function 9

1.3 Virtual Screening in Drug Design 12

1.3.1 Molecular Docking 12

1.3.2 Improvement of Virtual Screening Predictions: Ensemble Docking 12

1.3.3 Improvement of Virtual Screening Predictions: Rescoring of Docking Results 13

2 Materials and Methods **15**

2.1 Ligand-Based Virtual Screening 15

2.2 Homology Modeling 16

2.2.1 PRK1 Homology Modeling Using PKC-theta Template 17

2.2.2 PRK1 Binding Pocket Refinement 19

2.3 Structure-Based Virtual Screening 22

2.3.1 Molecular Docking 22

2.3.2 Pharmacophore-Based Virtual Screening 23

2.4 Molecular Dynamics Simulations 25

2.5 MM-PBSA and MM-GBSA Approaches for BFE Calculations 27

2.5.1 BFE Calculations for Compounds from Biomol Kinase Library . . 28

2.5.2 BFE Calculations for Rescoring of Docking Results 29

2.6 PRK1 *in vitro* Assay 30

3 Results and Discussion 32

3.1 PRK1 Inhibitors Identified by *in vitro* Screenings 32

3.2 PRK1 Homology Modeling 34

3.2.1 PRK1 Homology Modeling using PKC-theta as a Template 34

3.2.2 Refinement of PRK1 Homology Model with Ligands in the Binding Pocket 35

3.3 Molecular Dynamics and BFE Calculations for Top-Scored Biomol Compounds 38

3.4 Initial Virtual Screenings to Search for PRK1 Inhibitors 41

3.4.1 Similarity-Based Virtual Screening for Isoquinolines HA-1077 and H-7 42

3.4.2 Similarity Search for Pyrrolopyrimidine Derivative CP-690550 . . 43

3.4.3 Pharmacophore-Based Virtual Screening for PD-0166285 44

3.5 Post-Processing of Docking Results Using BFE Calculations 44

3.5.1 Studied Dataset 45

3.5.2 Docking Into a Single Protein Conformation 48

3.5.3 Docking Into an Ensemble of PRK1 Conformations 50

3.5.4 BFE Calculation for Rescoring of Docking Results 51
3.5.4.1 Accuracy of Docking/Scoring Methods in Enrichment Performance 52
3.5.4.2 Estimating the Biological Activity of PRK1 Inhibitors . 53
3.5.5 Validation of the Results 56
3.6 Application of the Model on External Datasets 63
3.6.1 Selleck Kinase Inhibitor Set 63
3.6.2 GSK Published Kinase Inhibitor Set 64
3.7 PRK1 Crystal Structures Released 68
3.7.1 Comparison of PRK1 Crystal Structures to Homology Models . . . 68
3.7.2 Comparison of Docking and Crystallographic Poses of Ro-318220, Lestaurtinib and Tofacitinib 71

4 Conclusions **77**
4.1 Application of Computer-Based Approaches to Search for Novel PRK1 Inhibitors 77
4.1.1 Structure-Based and Ligand-Based VS 77
4.1.2 Prediction of the Biological Activity 78
4.1.3 Comparison of Different VS Approaches to Search for PRK1 Inhibitors 79
4.2 Analysis of PRK1 Crystal Structures 82
4.3 Design of New PRK1 Inhibitors 83

Bibliography **87**

Appendix **112**

A Supplementary Information **112**

List of Figures

1.1 Kinase catalytic reaction 1

1.2 The structure of a kinase catalytic domain 3

1.3 The schematic representation of characteristic ATP binding mode 4

1.4 Different type of kinase inhibitors 6

1.5 Schematic representation of the structural features of PRK1 protein . . . 8

1.6 Schematic representation of PRK1 function 11

2.1 Sequence alignment between PRK1 and PKC-theta 18

2.2 The evaluation of PRK1 homology model with DOPE potential 19

2.3 PROCHECK stereochemical analysis of PRK1 homology model 20

2.4 RMSD plot for MD simulations of the PRK1 homology models 21

2.5 Pharmacophore model of compound PD-0166285 25

3.1 PRK1 inhibitors known from the literature as of 2010 33

3.2 PRK1 inhibitors identified by in-house *in vitro* screenings 34

3.3 Comparison of the generated PRK1 homology models 37

3.4 Comparison of the ATP-binding pockets of two PRK1 homology models - hm_STU and hm_2esm_l 38

3.5 Correlation plot for compounds from Biomol database, for which MD simulation was performed 41

3.6 Isoquinoline derivatives H-7, HA-1077 and inhibitors identified by similarity search using ligand-based fingerprint method MOLPRINT 2D 42

3.7 PRK1 inhibitors CP-690550 and other pyrrolopyrimidine derivatives, which were identified by similarity search using MACCS fingerprints combined with docking into PRK1 homology model 43

3.8 PD-0166285 and inhibitors identified by pharmacophore-based virtual screening 44

3.9 Representative PRK1 inhibitors from DS1, showing different hinge-binding scaffolds 48

3.10 Enrichment plot showing the percentage of actives found at a given percentage of the ranked database for DS1 and DS2 49

3.11 Receiver operating characteristic (ROC) curves for docking of training (DS1) and test set (DS2) into each single and ensemble of PRK1 homology models 51

3.12 Regression and corresponding correlation coefficients r^2 between observed activities and different scores are shown for the ensemble docking results of DS1 54

3.13 Regression and corresponding correlation coefficients r^2 are shown for DS1 between observed and predicted by QSAR_model compound activities . . 56

3.14 The representative subset of PRK1 inhibitors from DS2 58

3.15 Regression between the observed pK_d of PRK1 inhibitors from the test set DS2 versus predicted pIC_{50} by QSAR_model_3 60

3.16 Conformations of compound LY-317615 before and after minimization of the complex 60

3.17 PRK1 conformational flexibility on example of G-loop and C-tail motions 69

3.18 The superimposition of PRK1 homology models and crystal structures by all C_{alpha} atoms 70

3.19 The superimposition of PRK1 homology models and crystal structures by pocket residues using C_{alpha} atoms 70
3.20 The superimposition of PRK1 homology model hm_2jed and crystal structure 4OTH by pocket residues using C_{alpha} atoms 71
3.21 Comparison of the binding mode of compound Ro-318220 observed in crystal structure and the top-scored by glide_SP pose from ensemble docking 73
3.22 Comparison of the binding mode of compound lestaurtinib observed in crystal structure and the top-scored by glideSP pose from ensemble docking 74
3.23 Comparison of the binding mode of compound tofacitinib observed in crystal structure and the top-scored by glide_SP pose from ensemble docking 75

4.1 Kinome interaction maps for PRK1 inhibitors staurosporine and CP-690550 (tofacitinib) tested in a comprehensive analysis of kinase inhibitor selectivity by Davis at al.[19] 84
4.2 Structure and binding mode of PRK1 inhibitor tofacitinib (CP-690550) . 85

A.1 RMSD plots for 14 Biomol compounds for which MD simulation was performed 113
A.2 PRK1 inhibitors identified by comprehensive assay set for protein kinases by Davis et al.[19] (test set, DS2) 115

List of Tables

1.1 Examples of PKC structures available in the PDB, which share high sequence identity with PRK1 9

3.1 Model names, binding pocket template structures and inhibitors used for refinement of the homology models 36

3.2 MM/PB(GB)SA calculations for 14 top-scored by GlideSP compounds from Biomol set 40

3.3 PRK1 inhibitors from DS1 along with experimental activity and standard error of mean 46

3.4 Summary of r^2 correlation coefficients for different scoring methods using single homology models or ensemble of six homology models 53

3.5 Comparison of the measured pIC_{50} versus the predicted values by the QSAR_model for the most potent PRK1 inhibitors from DS1 56

3.6 Biological activity of 20 PRK1 inhibitors published by Davis et al.[19] . . 57

3.7 Results of the validation of QSAR_model_3 on the test set DS2 62

3.8 Comparison of measured data versus predicted pIC_{50} by QSAR_model_3 for the most potent PRK1 inhibitors from DS2 62

3.9 Twelve top-ranked compounds from Selleck database by predicted pIC_{50} by QSAR_model_3 66

3.10 Twenty five top-ranked compounds from GSK PKIS database by predicted pIC_{50} by QSAR_model_3, which were tested on PRK1 *in vitro* 67

3.11 Crystal structures of PRK1 68
3.12 The root mean square deviation (RMSD) between crystallographic pose and docked pose of ligands Ro-318220, lestaurtinib and tofacitinib 72

4.1 The overview of the screening performance using different approaches for a search of PRK1 inhibitors 80

A.1 Compounds from ChEMBL database, which were tested on PRK1 before the year 2011 and have reported activity 114

Acronyms

ACC antiparallel coiled-coil. 8

AML acute myelogenous leukemia. 73

AR androgen receptor. 9–11

ATP γ-phosphate of adenosine-5'-triphosphate. 1–5, 7, 36, 37, 39, 48, 49, 63, 68, 82–85

AUC area under the curve. 50

BFE binding free energy. 13, 14, 27–30, 32, 39–41, 52–54, 63, 65, 78, 79, 81, 82

BIM bisindolylmaleimide. 32, 63

BLAST basic local alignment search tool. 17, 35

CDK cyclin-dependent kinase. 63

CML chronic myeloid leukemia. 7

DMSO dimethyl sulfoxide. 64

DOPE discrete optimized protein energy. 17, 19

FDA Food and Drug Administration. 4, 63, 75

FEP free-energy perturbation. 13

GSK GlaxoSmithKline. 64, 81, 82

HTS high-throughput screening. 15, 24

IP intellectual property. 5

JAK Janus kinase. 63, 74, 83, 84

LIE linear interaction energy. 13

MD molecular dynamics. 13, 14, 18, 25, 26, 28, 29, 40

MM-GBSA molecular mechanics generalized Born surface area. 13, 14, 28, 78

MM-PB(GB)SA molecular mechanics Poisson-Boltzmann (generalized Born) surface area. 13, 27–29, 52, 56, 59

MM-PBSA molecular mechanics Poisson-Boltzmann surface area. 13, 14, 28, 78

MPL magnesium-positioning loop. 2

NCBI National Center for Biotechnology Information. 17, 35

PDB protein data bank. 14, 16, 17, 19, 35, 38, 48, 68, 71, 75

PDK1 phosphoinositide-dependent kinase-1. 44

PKA protein kinase A. 35, 48, 83

PKB protein kinase B. 35

PKC protein kinase C. 5, 8, 9, 17, 19, 32, 35, 63, 71, 83

PKIS published kinase inhibitor set. 64, 82

PRK1 protein kinase C-related kinase 1. 8–11, 13, 14, 16, 17, 24, 30–35, 38, 39, 41–44, 63–65, 68–71, 80–86

PSA prostate specific antigen. 74

QM quantum mechanics. 14

QM/MM quantum mechanics/molecular mechanics. 14

QM/MM-GBSA quantum mechanics/molecular mechanics generalized Born surface area. 14, 52, 55, 56, 59, 78

QSAR quantitative structure activity relationship. 54, 55, 59, 61, 63, 79

RMSD root mean square deviation. 18, 37, 68, 69, 71, 72, 82

ROC receiver operating characteristic. 50, 59

RTKs receptor tyrosine kinases. 63

SAR structure activity relationship. 5, 76, 83

TI thermodynamic integration. 13

VS virtual screening. 12–15, 23, 24, 43, 50–52, 78, 80–82

CHAPTER 1

Introduction

1.1 Kinase

1.1.1 Kinases and Their Role in Signal Transduction

Protein kinases are enzymes that transfer a phosphate group from γ-phosphate of adenosine-5’-triphosphate (ATP) to protein substrates covalently attaching it to one of three amino acids (serine, threonine or tyrosine) that have a free hydroxyl group (see Figure 1.1).

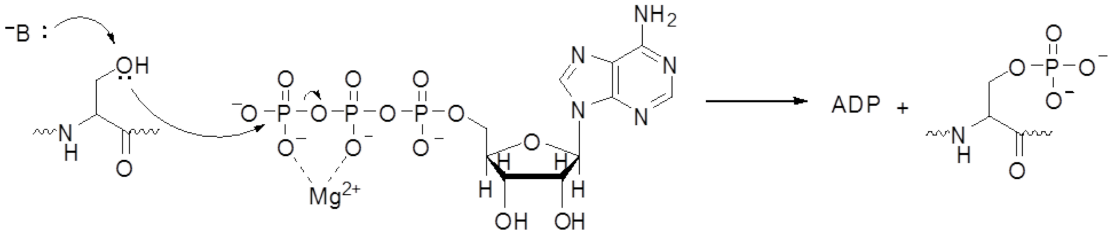


Figure 1.1: Kinase catalytic reaction. The mechanism of phosphate group transfer catalyzed by kinase is shown as phosphorylation of a serine residue.

Phosphorylation of proteins by kinases has been shown to play a significant role in signal transduction cascades by regulating versatile cellular processes such as cell proliferation, differentiation and apoptosis.[1, 2] These processes show a remarkable degree of coordination and the activity of kinases is highly regulated. Deregulation of kinase function or abnormal phosphorylation has been implicated in various pathological conditions such as cancer, infectious diseases as well as in metabolic, immunological and neurological disorders. Therefore, protein kinases have become one of the most prospective therapeutic drug targets.[1–3] A majority of cancers usually arise from events that activate kinases, cause their overexpression or disable their intracellular inhibition.[4] Because of its important physiological and pathological role, the regulation of kinase activity is an important therapeutic strategy for the treatment of many diseases.

1.1.2 Kinase Structure

Being one of the largest and diverse protein families, kinases mediate a large number of biological processes. Manning et al. have defined 518 kinases encoded in the human genome including tyrosine, serine/threonine and dual-specificity kinases.[5]

Despite the remarkable sequence diversity within this enzyme class, they share a universal secondary structure of the catalytic domain. Typically, kinase consists of two subunits – a catalytic and a regulatory domain. The catalytic domain consists of N- and C-terminal lobes connected by the so-called hinge region (see Figure 1.2). The N-terminal lobe is made of a 5-stranded β-sheet, a conserved $\alpha-$helix (helix C) and a magnesium-positioning loop (MPL). The area between them forms the active site harboring the ATP molecule. The ATP-binding pocket together with less conserved adjacent pockets is often considered as the target site for drug design. The C-terminal lobe is larger than the N-terminal domain and it is mainly α-helical.

The activation loop is a flexible loop, which plays a role in kinase regulation. Typically, kinases are strongly activated when a characteristic residue at the activation loop (serine, threonine or tyrosine) is phosphorylated. The N terminus of the activation loop contains the conserved DFG (Asp-Phe-Gly) motif, which can adopt so-called "in" and "out" conformations switching an active or inactive kinase state.

The glycine-rich loop (G-loop or P-loop) is another highly conserved kinase motif. It contains a typical sequence GxGxxG, which gives it an extra flexibility. The G-loop can adopt its conformation depending on the compound occupying the ATP-binding site.

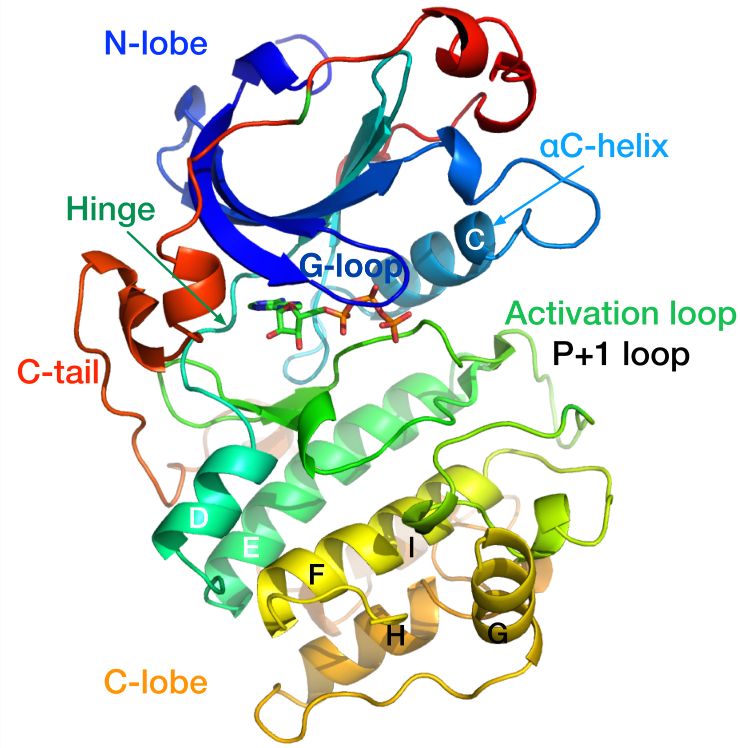


Figure 1.2: The structure of a kinase catalytic domain is exemplary shown for PKC-iota (PDB ID: 3A8W) in complex with ATP (green sticks). The protein is displayed as ribbon colored through the rainbow spectrum from blue at the N terminus to red at the C terminus.

The structure of the kinase catalytic domain is highly conserved across the protein kinase genome. Moreover, the residues within the ATP binding cleft are highly conserved among various kinase families. This makes the search for highly selective ATP-competitive inhibitors challenging. Nevertheless, there are certain differences in regions adjacent to the binding cleft, which ATP does not occupy (see Figure 1.3). Such less conserved parts of the protein provide an opportunity for the design of selective ATP-competitive inhibitors. Thus, a detailed knowledge of the protein structure is needed to assist the rational design of kinase inhibitors.

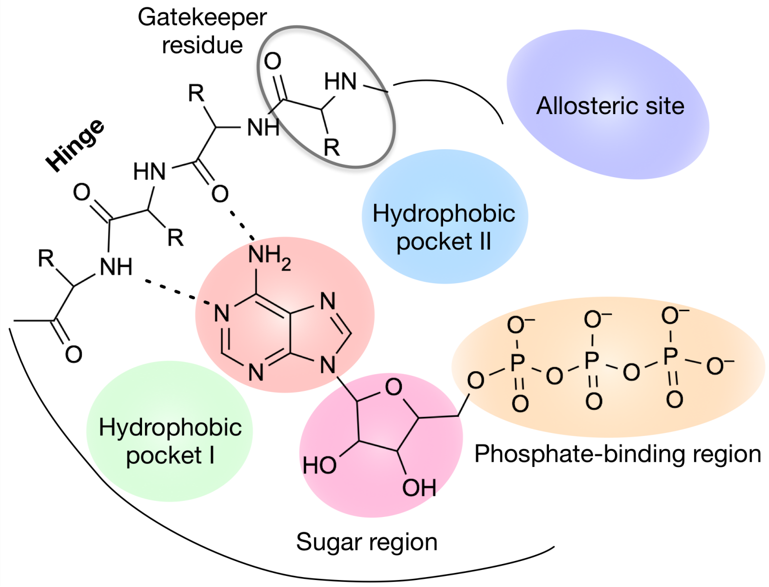


Figure 1.3: The schematic representation of characteristic ATP binding mode divided into subregions. The main structural elements are labeled.

1.1.3 Types of Kinase Inhibitors

Protein kinases are considered as a very promising drug targets.[3] During the last years, more than 20 kinase inhibitors have been approved by the Food and Drug Administration (FDA) for the treatment of various cancers,[6] and even more compounds are now in clinical trials. Nevertheless, kinase inhibitors are facing few major challenges in cancer therapy. One is that cancer's adaptive behavior can lead to the development of drug resistance. To overcome this problem new strategies, more effective than selective inhibitors, are now considered (e.g. combination of few therapies, using of multi-targeted or "group-selective" agents).[7–9]

Other problems are poor kinase selectivity and off-target effects. The majority of known kinase inhibitors address the ATP binding pocket, which is highly conserved within the human kinome. The strategies to optimize selectivity can include targeting of less conserved regions adjustment to ATP, targeting other pockets or stabilizing inactive kinase conformation. Three major types of kinase inhibitors are discussed below.

1.1.3.1 Type I Inhibitors (ATP-Competitive)

The initial strategy used in drug discovery was aimed to find small molecules which mimic ATP binding. Such compounds target the ATP binding site of the enzyme in its active form and are referred to as *type I* inhibitors. The active form is characterized by an open conformation of the activation loop, when the conserved aspartate-phenylalanine-glycine (DFG) residues are situated at the beginning of the activation loop. An example of type I inhibitor is the natural product staurosporine (see Figure 1.4a). It inhibits Ca^{2+}-dependent protein kinase C (PKC) in the nanomolar range.[10] However, structure activity relationship (SAR) study showed that staurosporine and its analogs are non-selective inhibitors and target most of the kinases.[11] Since the specific interactions between an ATP-mimetic ligand and a protein kinase in its active site are often very similar for multiple kinases, it is clear why type I inhibitors suffer often from cross-reactivity. Moreover, such inhibitors have to compete with millimolar intracellular ATP levels.[12] Furthermore, recent immense investigations, aiming to find ATP-mimicking inhibitors, has led to a crowded intellectual property (IP) space, leaving few space for innovation.

Despite all before-mentioned problems, the ATP-competitive compounds still remain very promising type of kinase inhibitors and have a potential in drug discovery. In order to improve selectivity of type I inhibitors, the less conserved areas of the kinase domain can be targeted, such as the hydrophobic pocket, the entrance to which is controlled by the so-called "gatekeeper residue" (see Figure 1.3); solvent-exposed region; differences in conformations of glycine-rich loop (G-loop) or flexibility of hinge region. [13–16] Interestingly, some of type I inhibitors have very good selectivity profile, for example, JAK inhibitor CP-690550 (Tofacitinib) [17–19] or dibenzepinones as inhibitors of p38 kinase.[20]

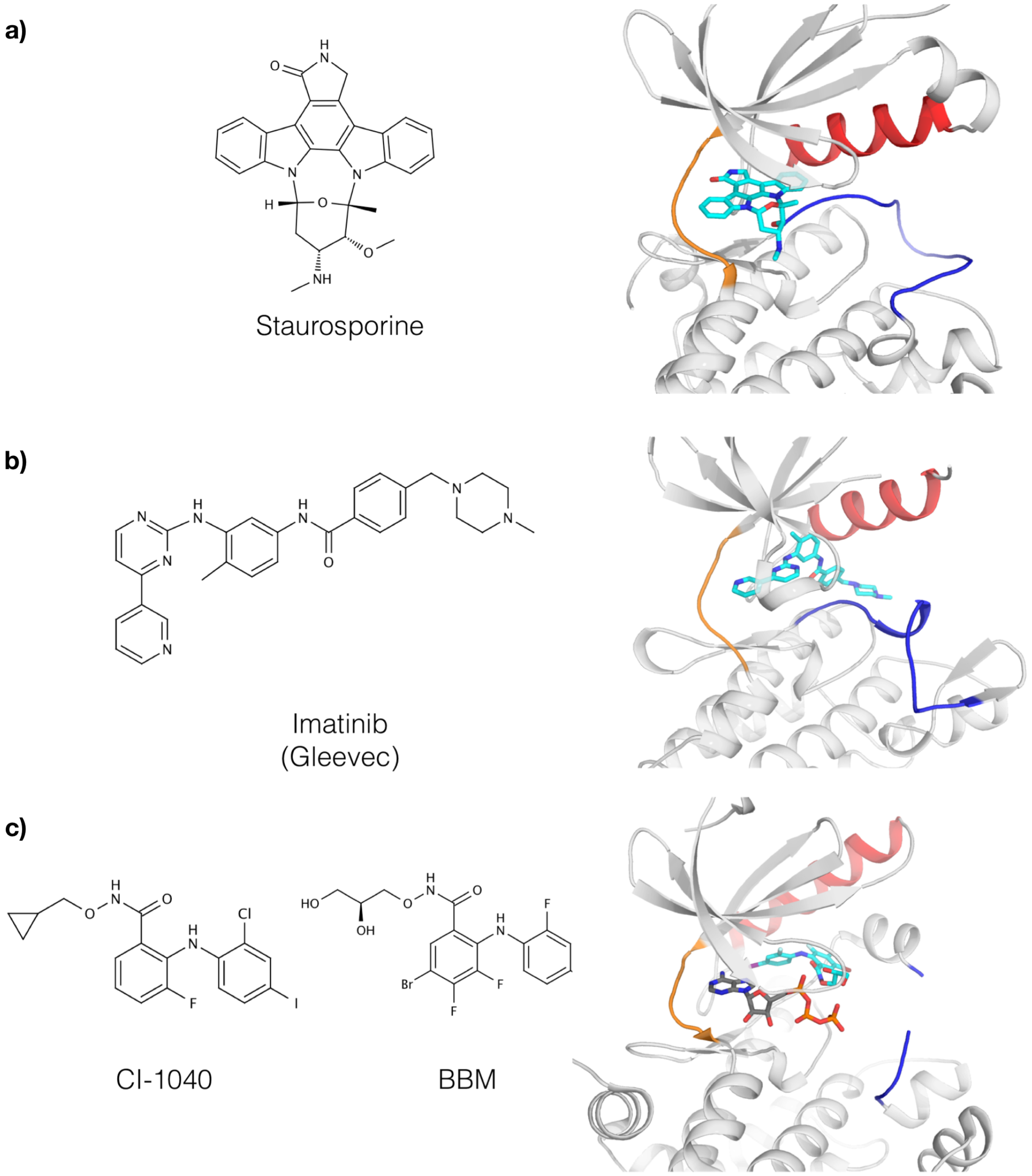


Figure 1.4: Different type of kinase inhibitors: **a)** non-selective type I inhibitor staurosporine in complex with PKC-theta (PDB ID: 1XJD); **b)** Abl kinase domain in complex with type II inhibitor imatinib (PDB ID: 2HYY); **c)** type III inhibitors CI-1040 and BBM and the binding mode of BBM bound to MEK1 in complex with ATP (PDB ID: 1S9J). Ligand carbon atoms are colored in cyan; ATP and protein ribbon - in dark gray. The ribbon of hinge region, αC-helix and activation loop is colored in orange, red and blue, correspondingly.

1.1.3.2 Type II Inhibitors

The first kinase inhibitor to reach the market was Imatinib (Gleevec, Novartis, see Figure 1.4b). Imatinib targets Bcr-Abl fusion protein and was approved for the treatment of chronic myeloid leukemia (CML).[21] Structural studies demonstrated that Imatinib stabilizes an inactive form of kinase, which is characterized by a closed conformation of the activation loop (DFG-out).[22] Such type of compounds, classified as *type II* inhibitors, usually occupy not only the ATP binding site like classical type I inhibitors do, but also extend their interactions to the hydrophobic pocket available in the inactive form. These inhibitors usually have advantageous pharmacological properties, however, the mutation resistance to type II inhibitors can quickly develop. In general, type II inhibitors are less promiscuous than ATP-competitive type I, but absolute selectivity has not been achieved.[19]

1.1.3.3 Type III Inhibitors

The emerged efforts directed towards the development of selective kinase inhibitors have lead to the new classes of inhibitors, which target less conserved regions outside ATP-binding pocket. Thus, a number of so-called *type III* inhibitors, which bind to an allosteric site of a kinase beyond the gatekeeper residue and do not interact with hinge region, were discovered.[23] The allosteric site is formed when the kinase adopts inactive (DFG-out) conformation. Type III inhibitors stabilize the inactive conformation, thus preventing the enzymatic activity of kinase. These compounds have the highest degree of selectivity, they can be designed for a particular kinase. Examples of allosteric inhibitors are shown on Figure 1.4c. These are MEK1 and MEK2 inhibitors such as BBM and CI-1040. BBM occupies a pocket adjacent to the ATP binding site as it is shown on Figure 1.4c.[24] Despite all advantages of type II or type III inhibitors, their design can be hampered by the difficulty to predict whether it is favorable or not for the individual kinase to adopt the DFG-out conformation and form a complex with such a compound.

1.2 Protein Kinase C Related Kinase 1

Protein kinase C-related kinase 1 (PRK1), also known as PKN1 or PKNα, is a serine/-threonine kinase and and belongs to the PKC super family. The PRK cDNA was first isolated in 1994 from a human hippocampal cDNA library.[25] Mammalian PRK/PKN family has three isoforms, which show different enzymatic properties, tissue distributions, and varied functions [26]:

PRK1	[=PKNα/PAK1/PKN1]
PRK2	[=PKNγ/PAK2/PKN2]
PRK3	[=PKNβ]

PRK1 is found in different eukaryotic organisms and its sequence is highly conserved during evolution.

1.2.1 Structure of PRK1

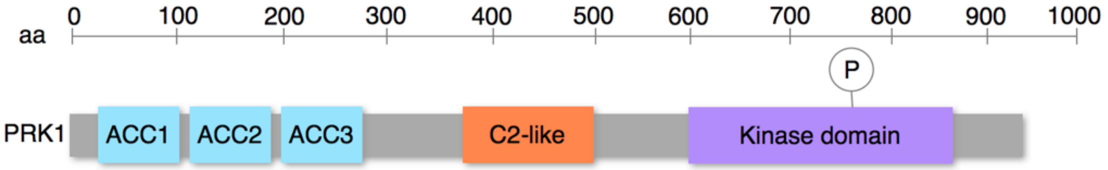


Figure 1.5: Schematic representation of the structural features of PRK1 protein (aa=amino acid).

Typically, PRKs consist of three highly conserved region (see Figure 1.5):

1) a unique regulatory region at the N-terminal half of the protein containing three homologous stretches rich in charged amino acids (antiparallel coiled-coil antiparallel coiled-coil (ACC) domains) followed by a leucine zipper-like sequence;
2) C2-like auto-inhibitory domain sensitive to arachidonic acid;
3) a single polypeptide chain with C-terminal catalytic domain (also known as kinase

domain).[26, 27] The catalytic domains of PRKs have high sequence homology to that of the PKC family members (see Table 1.1).

Table 1.1: Examples of PKC structures available in the PDB, which share high sequence identity with PRK1.

PDB ID	Protein kinase C	Ligand	Resolution, Å	Identities, %
1XJD	PKC-theta	Staurosporine	2.00	50
2JED	PKC-theta	Nvp-Xaa228	2.32	50
2I0E	PKC-beta	BIM	2.60	48
3IW4	PKC-alpha	Nvp-Aeb071	2.80	48
3A8X	PKC-iota	Phosphorylated peptide	2.00	41

1.2.2 PRK1 Function

PRK1 is a protein kinase involved in Rho- and androgen receptor-mediated signaling pathways.[26, 28–34] It is able to shuttle between the nucleus and the cytoplasm[35] and can be activated by phospholipids[36], arachidonic acid [37] and Rho GTPase.[27, 38] PRK1 and PRK2 have been reported to phosphorylate class IIa histone deacetylases (HDACs) -5, -7 and -9 at a threonine residue positioned within the nuclear localization signal of the protein.[39] The microarray analysis of PRK1 gene expression showed high levels of PRK1 in various malignancies, but especially in ovarian serous carcinomas.[40] Additionally, the role of PRK1 in the development of Germinal centers by regulating Akt kinase downstream has been described.[41] Furthermore, it was shown that activation of PRK1 kinase stimulates androgen receptor (AR) activity and is implicated in tumorigenesis.[32] Metzger et al. have demonstrated that PRK1 levels correlate with Gleason scores of prostate carcinomas[42] and knockdown of PRK1 gene or inhibition of PRK1 with known kinase inhibitors such as Ro-318220 blocks propagation of AR-induced tumor cell production (United States Patent Application 20070196882 Schüle and Metzger).

The AR is a nuclear receptor, that is activated by androgenic hormones testosterone or dihydrotestosterone (androgens). The main function of AR is to modulate gene expression which controls diverse biological functions including cell growth and differentiation, development, and function of male reproductive and accessory sex tissues. Androgen receptor signaling pathway plays an important role in the development and progression of prostate cancer[43, 44] which is the second leading cause of cancer deaths in men worldwide.[45] During androgen-dependent progression, the growth and survival of prostate cancer cells are critically dependent on androgens. An activation of AR initiates a series of events leading to the regulation of the target genes. When the AR is inactive, it is bound to heat-shock proteins in the cytoplasm of prostate cells. The binding of the androgen dihydrotestosterone into the receptor's binding site causes dissociation of AR from the heat shock proteins, dimerization and translocation of the complex into the nucleus, where it binds to the androgen-response elements, thereby activating genes involved in cell growth.[46]

Androgen ablation therapy, which is also called hormone therapy, is the most common clinical procedure, aiming to control/stop the growth of cancer cells. Most prostate cancers respond to this therapy. However, in some cases prostate cancer cells can recur even when using AR antagonists and reducing androgen levels. At this point, the prostate cancer becomes androgen-independent (also referred to as androgen-resistant or castration-resistant prostate cancer (CRPC)) and it can further progress and metastasize.

Despite the recent advances in research on prostate cancer, the molecular mechanism for the tumor reoccurrence is not entirely understood. The studies show that androgen receptors can be transactivated in the absence or very low amount of androgens. The aberrant AR signaling involves several, non-mutually-exclusive mechanisms including extracellular peptides such as Insulin-like growth factor, Epidermal growth factor and Interleukin-6, which can also activate the intracellular kinase signaling cascades responsible for cell proliferation and transcription initiation.[43]

Additionally, with biochemical assays it was speculated that PRK1 kinase controls the activity of AR in addition to ligands in hormone-dependent proliferation of prostate cancer cells.[32] In 2008 Schüle et al. have shown that PRK1 phosphorylates histone H3 upon ligand dependent recruitment to AR target genes.[42] In turn, phosphorylation of H3T11 increases demethylation of Lys 9 (K9) by Jumonji C (JmjC)-domain-containing protein (JMJD2C), which promotes AR-dependent gene expression and tumor cell proliferation.[47] Furthermore, PRK1 may directly phosphorylate JMJD2C, thereby stimulating its activity.[48] The function of PRK1 is schematically shown in Figure 1.6.

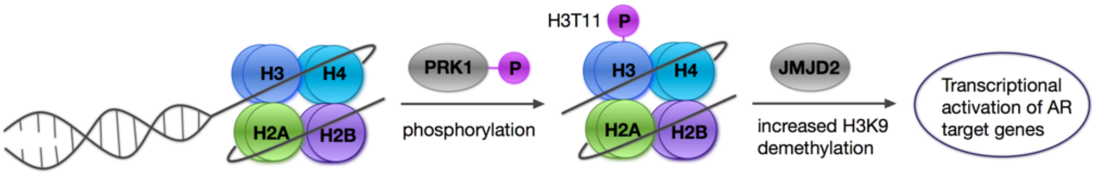


Figure 1.6: Schematic representation of PRK1 function. The structural unit of the chromatin, termed the nucleosome, comprises DNA wrapped around the octamer histone consisting of histones H3, H4, H2A and H2B. PRK1 phosphorylates histone H3 at threonine 11 (H3T11), which leads to an increased demethylation of histone H3 at lysine 9 (H3K9) and subsequent activation of androgen receptor target genes.

The role of PRK1 in androgen-independent prostate cancer is not entirely understood. It is known that PRK1 can be activated by Rho family GTPases and mediates such processes as cell migration and tumor cell invasion, which contribute to metastasis formation.[34, 49] In a recent study a novel mechanism for PRK1-controlled migration of androgen-independent prostate cancer cells was revealed, showing that PRK1 controls migration and invasion but not proliferation of PC-3M-luc2 cells.[50] Schüle et al. have demonstrated that PRK1 depletion in mice or inhibition with lestaurtinib leads to a dramatic decrease in the tumor cells ability to metastasize.[50] These findings indicate an important role of PRK1 in the hormone-resistant prostate cancer.

In summary, PRK1 is considered as a promising therapeutic target, and the discovery of potent and selective PRK1 inhibitor may lead to the development of clinical PRK1 modulators for the treatment of certain neoplasms. The availability of small molecule inhibitors of PRK1 that are selective and bioavailable will represent a major breakthrough in this field.

1.3 Virtual Screening in Drug Design

1.3.1 Molecular Docking

Molecular docking is one of the main techniques used in structure-based drug discovery.[51–55] Different studies have shown that docking-based virtual screening (VS) of large compound databases can be applied effectively for the identification of novel hits.[56–58] Despite its great potential, the method has a number of drawbacks and limitations.[59–61] The two major tasks of molecular docking include the determination of the correct ligand orientation within the protein active site and the assessment of the protein-ligand affinity. Whereas docking methods already show good results in the prediction of binding modes, the second task still remains a key challenge in computational chemistry. Considering the balance between the accuracy of calculations and the computational cost, scoring functions were developed on the basis of simplified empirical force fields or potentials of mean force using certain approximations, e.g. they usually do not take solvation effects or protein flexibility into account. Thus, accurate prediction of the binding affinity using docking methodologies remains an elusive goal and there is the need of more precise methods for its evaluation.

1.3.2 Improvement of Virtual Screening Predictions: Ensemble Docking

Various approaches were developed in order to improve the docking performance, e.g. docking to an ensemble of protein structures, normalizing docking scores, or using rescoring procedures. Different studies have shown that docking methods perform well in reproducing ligand binding poses for experimentally derived structures, however, they can fail if a protein structure was solved in the presence of a very different compound.[62] It is known, that during ligand binding conformational changes of a protein can occur. In fact, even small changes in a receptor structure can be important for ligand binding. Thus, using rigid receptor structures can hamper correct ligand docking, for example, the active ligands will not be docked correctly into a receptor or can be scored poorly. One of the approaches, which handle protein flexibility, is the so called "ensemble docking".[63–65]

Here, a ligand is docked into several protein structures/conformations in order to identify the best-scored pair of protein conformation and ligand binding mode. In the present work we used an ensemble of PRK1 homology models refined in the presence of active inhibitors from different chemical classes for docking studies.

1.3.3 Improvement of Virtual Screening Predictions: Rescoring of Docking Results

Correct ranking of compounds according to their binding affinity is another critical issue. Over the last years a large effort has been undertaken to address this question. The approaches vary from simple methods like rescoring of poses with external scoring functions, consensus scoring, normalization, etc.[66] to more sophisticated and power-demanding calculations of binding free energy (BFE) such as linear interaction energy (LIE), molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) or molecular mechanics generalized Born surface area (MM-GBSA), free-energy perturbation (FEP) and thermodynamic integration (TI).[67–70] Although some studies show successful examples of lead optimization using FEP and TI methods, they are still rarely applied in the drug discovery process, mainly due to the high computational costs. Nevertheless, the recent increase in computing power made it possible to apply some BFE calculation methods on quite large compound datasets. This allows using them not only on the late stages of drug discovery, but also as a tool for post-processing of VS results. The number of studies, reporting on the successful application of MM-PBSA, MM-GBSA methods for the estimation of protein-ligand binding affinities is increasing, and automated procedures for BFE estimation applicable for a large compound selection have been developed.[71–77]

Traditionally, molecular mechanics Poisson-Boltzmann (generalized Born) surface area (MM-PB(GB)SA) calculations are carried out using a number of snapshots derived from equilibrated molecular dynamics (MD) simulation. Various studies have demonstrated the efficacy of this approach for predicting binding affinities,[74, 75, 78, 79] however, it is usually applied for a small subset of molecules due to the limitations in computational power. Recent publications report that the calculation of BFE using a single snapshot derived from a short complex minimization in implicit solvent, in some cases combined with short MD simulation, can also result in a good correlation with experimental data.[80–84]

Furthermore, the combined quantum mechanics/molecular mechanics (QM/MM) scoring function implemented in AMBER[85] was tested for calculation of the binding affinity.[86, 87] This method allows treating a small part of the complex, e.g. the ligand and residues of the active site, with semi-empirical quantum mechanics (QM) calculations, while the remaining part is treated by classical molecular mechanics as a fixed charge background. Such a procedure considers electronic effects such as polarization [88–91] and charge transfer[92] for the selected part of the complex, which can play an important role in ligand binding to the target, and which are usually not considered in classical BFE calculations. Moreover, the support of QM/MM implicit solvent simulations was implemented into generalized Born surface area method, resulting in mixed quantum mechanics/molecular mechanics generalized Born surface area (QM/MM-GBSA) approach.[93]

Most of the studies using MM-PBSA, MM-GBSA approaches for BFE calculations show good results for congeneric series of compounds in a given receptor or for structurally diverse ligands, but with known binding modes taken from the protein data bank (PDB).[71, 75, 94] However, in a VS experiment usually a large set of diverse inhibitors is fitted into the protein structure and the ligand conformation is usually selected according to the best docking score. In the current study we wanted to test the performance of different BFE calculation methods such as MM-PBSA, MM-GBSA, QM/MM-GBSA on their ability to predict the relative binding affinity of structurally different PRK1 inhibitors from in-house and published datasets. Furthermore, we included inactive compounds in our studies in order to see which of the approaches performs best in discrimination. The BFE model showing the best performance in enrichment studies as well as in the prediction of the biological activity represents a valuable tool for optimization of PRK1 inhibitors.

CHAPTER 2

Materials and Methods

Starting from 1990s new methods like combinatorial chemistry and high-throughput screening (HTS) were widely used for searching of drugs. Although these techniques were proven to be successful in identification of new biologically active compounds, they have a number of limitations such as low success ratio, expensive and time consuming assay development and screening procedures. Hence, the alternative approaches driving the initial phases of drug discovery process were developed. One of these is virtual screening (VS) - a computational tool for the rapid *in silico* estimation of large chemical databases aiming to identify bioactive compounds prior to biological testing, thus, saving time and costs. Due to the recent advances in computer technology, VS has become an important part of modern drug design pipelines. Historically, VS approaches are divided into two categories - ligand- and structure-based. The next sections are describing methods used in the current work.

2.1 Ligand-Based Virtual Screening

When the structure of the biological target is not known or when there is not enough structural information about it, but there is one or more ligand available, then ligand-based computational methods can be employed. Ligand-based approaches rely on the “similar property principle” [95], according to which compounds that are more similar with respect to their chemical structures are more likely to possess similar properties.

Due to the lack of structural information about PRK1, ligand-based virtual screening methods were employed at the initial stages of the work. Based on 2D chemical fingerprints or 3D molecular shape-based representations of the initial hit, a similarity search was performed to find compounds that match a given query the best. MOLPRINT 2D,[96, 97] a fast fragment-based similarity screening method, has been applied for a search of compounds similar to PRK1 inhibitors HA-1077 and H-7 (see Figure 3.2a). The similarity search algorithm of MOLPRINT 2D is based on atom environments, information-gain-based feature selection, and a naive Bayesian classifier. The Tanimoto coefficient [98] is used to compare fingerprints and to give a quantitative measure of the similarity.

Following the same principle, MACCS key fingerprints (MOE 2012) [99] were used for similarity comparison and database filtering for another PRK1 inhibitor, namely CP-690550 (see Figure 3.2b). The Tanimoto similarity metric with 85% threshold was used.

2.2 Homology Modeling

Structure-based design refers specifically to the three dimensional structure of the protein. However, the structures of many proteins still remain unsolved. During the last years the number of protein sequences in Universal Protein Resource have dramatically increased. In release 2014_03 of 19 March 2014, UniProtKB/Swiss-Prot, a protein sequence database which includes reviewed, manually annotated entries, contained 542,782 entries.[100] At the same time, total amount of experimentally determined structures deposited in PDB was much lower (98,900 structures).[101] One possible solutions which can help to fill the huge gap between annotated sequences and available 3D structures, is homology modeling. This approach is based on the observation that proteins with similar sequences have similar structures. Generally, it comprises the following steps:

1. Search for an available template structure which shares sufficient sequence identity with a given target protein (30% and more).

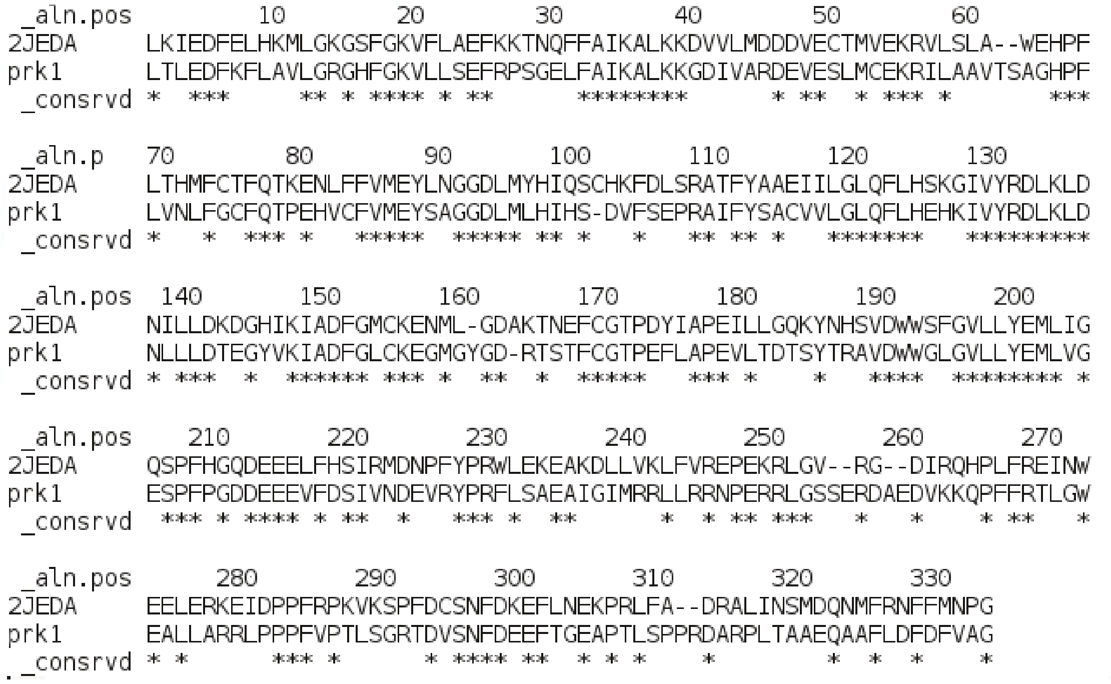
2. Sequence alignment of target and template proteins.
3. Building a model based on target-template alignment and 3D structure of template.
4. Refining and validation of the model.

These steps can be repeated until a sufficient model is achieved.[102] Furthermore, techniques like modeling using multiple templates, modeling of a protein-ligand complex and loop refining can improve the quality of the resulting model.

2.2.1 PRK1 Homology Modeling Using PKC-theta Template

Until August 2014, no three-dimensional structural information was available for the catalytic domain of PRK1. Thus, we applied comparative (homology) modeling techniques for PRK1 structure prediction. To identify which experimentally known 3D structure can serve as a template to generate a homology model of PRK1, a basic local alignment search tool (BLAST) search was carried out through the National Center for Biotechnology Information (NCBI) web site.[103, 104] The sequence of human PRK1 kinase domain (UniProt ID Q16512, residues 610-940) was used as a query. The structures available in the PDB, showing the highest similarity to the PRK1 sequence, belong to the PKC family. It was revealed that all these structures show the active kinase conformation. Among them, the crystal structure of PKC-theta (PDB code: 2JED, resolution 2.32 Å) was chosen as template as it shows the highest sequence identity with PRK1 (50%) and covers 99% of the whole sequence query. The homology model building was carried out using the MODELLER 9v8 software.[105]

The sequence alignment of PRK1 and PKC-theta was made using the default *align2d* parameters in MODELLER (see Figure 2.1). Based on the PKC-theta template structure and the alignment file, five models of PRK1 were generated. The model with the lowest value of the MODELLER objective function or the discrete optimized protein energy (DOPE) assessment score was chosen for further analysis.[106] The model was evaluated with the DOPE potential (see Figure 2.2). Further refinement was proceeded using Protein Preparation Wizard of Schrödinger Suite 2012.[107] Hydrogen atoms and partial charges were assigned and the model was energy minimized applying the OPLS2005 force field.[108, 109]



```
 _aln.pos            10        20        30        40        50        60
2JEDA       LKIEDFELHKMLGKGSFGKVFLAEFKKTNQFFAIKALKKDVVLMDDDVECTMVEKRVLSLA--WEHPF
prk1        LTLEDFKFLAVLGRGHFGKVLLSEFRPSGELFAIKALKKGDIVARDEVESLMCEKRILAAVTSAGHPF
 _consrvd   *  ***     ** * **** * **      ********      * **  * *** *       ***

 _aln.p     70        80        90       100       110       120       130
2JEDA       LTHMFCTFQTKENLFFVMEYLNGGDLMYHIQSCHKFDLSRATFYAAEIILGLQFLHSKGIVYRDLKLD
prk1        LVNLFGCFQTPEHVCFVMEYSAGGDLMLHIHS-DVFSEPRAIFYSACVVLGLQFLHEHKIVYRDLKLD
 _consrvd   *   *  *** *   *****  ***** ** *   *   ** ** *   *******   *********

 _aln.pos    140       150       160       170       180       190       200
2JEDA       NILLDKDGHIKIADFGMCKENML-GDAKTNEFCGTPDYIAPEILLGQKYNHSVDWWSFGVLLYEMLIG
prk1        NLLLDTEGYVKIADFGLCKEGMGYGD-RTSTFCGTPEFLAPEVLTDTSYTRAVDWWGLGVLLYEMLVG
 _consrvd   * ***  *  ****** *** *  **  *  *****   *** *    *   ****  ******** *

 _aln.pos      210       220       230       240       250       260       270
2JEDA       QSPFHGQDEEELFHSIRMDNPFYPRWLEKEAKDLLVKLFVREPEKRLGV--RG--DIRQHPLFREINW
prk1        ESPFPGDDEEEVFDSIVNDEVRYPRFLSAEAIGIMRRLLRRNPERRLGSSERDAEDVKKQPFFRTLGW
 _consrvd    *** * **** * **  *   *** *  **      *  * ** ***   *   *    * **   *

 _aln.pos        280       290       300       310       320       330
2JEDA       EELERKEIDPPFRPKVKSPFDCSNFDKEFLNEKPRLFA--DRALINSMDQNMFRNFFMNPG
prk1        EALLARRLPPPFVPTLSGRTDVSNFDEEFTGEAPTLSPPRDARPLTAAEQAAFLDFDFVAG
 _consrvd   * *      *** *      * **** **  * * *    *        *  *  *    *
```

Figure 2.1: Sequence alignment between PRK1 and PKC-theta (PDB code: 2JED) generated using *align2d* in MODELLER.

The quality of the model (further referred to as "hm_prk1") was analyzed by means of Protein Report tool implemented in Maestro v9.3.[107] Furthermore, the stereochemical analysis of the hm_prk1 with the PROCHECK [110] program (shown on Figure 2.3) confirmed that the model is reasonable. It shows that 93 % of the *phi* and *psi* angle values of the protein backbone are located in the most favored regions, and 5.6% in additionally allowed regions, two residues in generally allowed regions and one residue in the disallowed region. However, the outlier is not located in the binding pocket.

The stability of the derived homology model (hm_prk1) was examined by means of MD simulation using AMBER 12.[112] The root mean square deviation (RMSD) plot showed that the protein backbone atoms remained stable (near 2.5 Å) after the equilibration period (see Figure 2.4a).

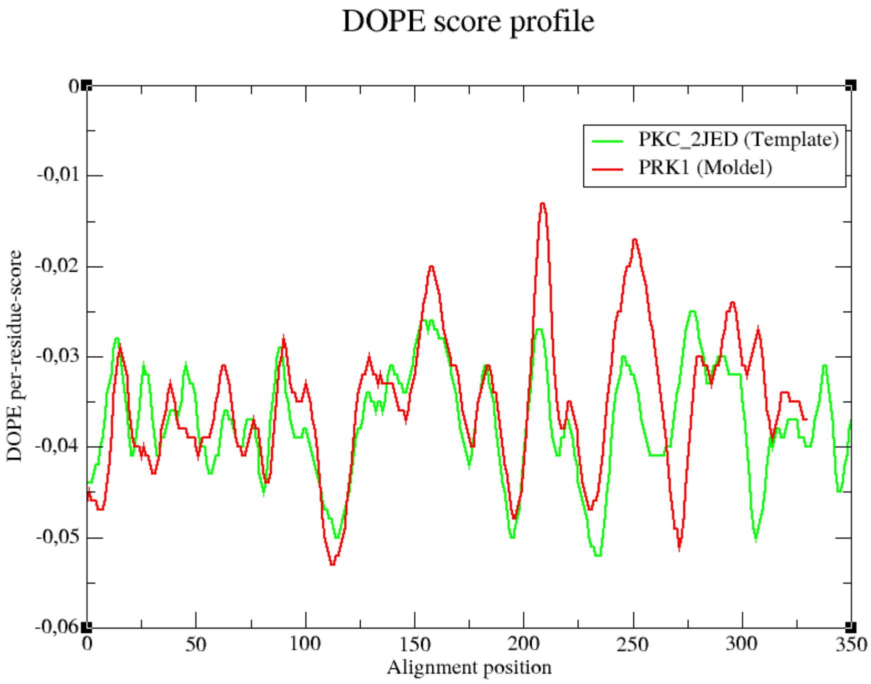


Figure 2.2: The evaluation of PRK1 homology model with DOPE potential.

2.2.2 PRK1 Binding Pocket Refinement

The binding site of the initially generated PRK1 model (hm_prk1) was refined by including structural information about the ligand (i.e. staurosporine) using MODELLER 9v8. For this the crystal structure of PKC-theta in complex with staurosporine at 2 Å resolution (PDB code: 1XJD) was taken from the PDB. The sequence of PRK1 was aligned to the initial model from the previous step (hm_prk1) and to the edited 1XJD structure containing residues of the binding site. Next, the alignment file and modeling script were modified for including the ligand into the refined homology model. The relative orientation of the ligand and the target was specified by restrained ligand-protein interactions within the binding site (hinge region residue Ser698). The conformation of the ligand was assumed to be rigid. Five models were generated and the best model was selected according to the lowest DOPE assessment score. Thus, a new homology model of PRK1 (hm_STU) with bound ligand was obtained. In a similar way, the binding site of PRK1 model (hm_prk1) was refined in the presence of other inhibitors (HA-1077,

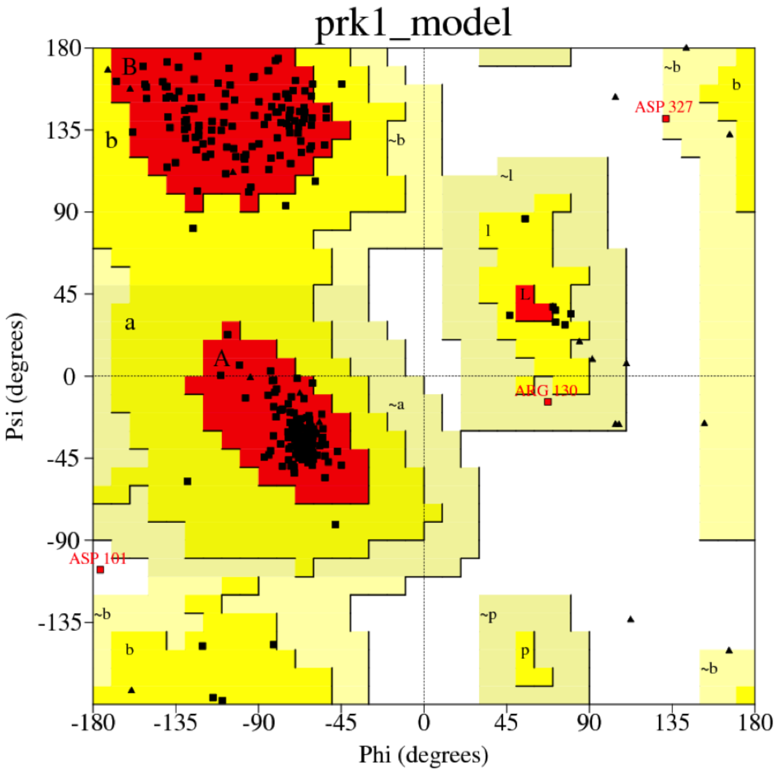


Figure 2.3: PROCHECK stereochemical analysis of the initial PRK1 homology model (hm_prk1, model was built using PKC-theta 2JED as a template). The Ramachandran plot shows the phi-psi torsion angles for all residues in the structure. The coloring on the plot represents the different regions as described by Morris et al. [111]: the red areas are "core" regions representing the most favorable combinations of phi-psi values. Usually, 90% or more of the residues in "core" regions represent good stereochemical quality.

Nvp-Xaa228, CP-690550,[19] GSK-690693[19]). The choice of the ligands for refinement was based on the structural diversity of the compounds as well as on different binding modes observed for these inhibitors, which could not be reproduced correctly by docking into a single protein structure. Furthermore, the residues 887-940 of the C-terminal loop were included in the model hm_2esm_l, whose structure was refined with the inhibitor HA-1077. Totally, six homology models were built. The stability of the models was confirmed through 10 ns of molecular dynamics simulation. All models gave relatively stable trajectories which were stabilized at $\text{RMSD}_{\text{heavy atoms}} = 2\sim3$ Å (see Figure 2.4b-f).

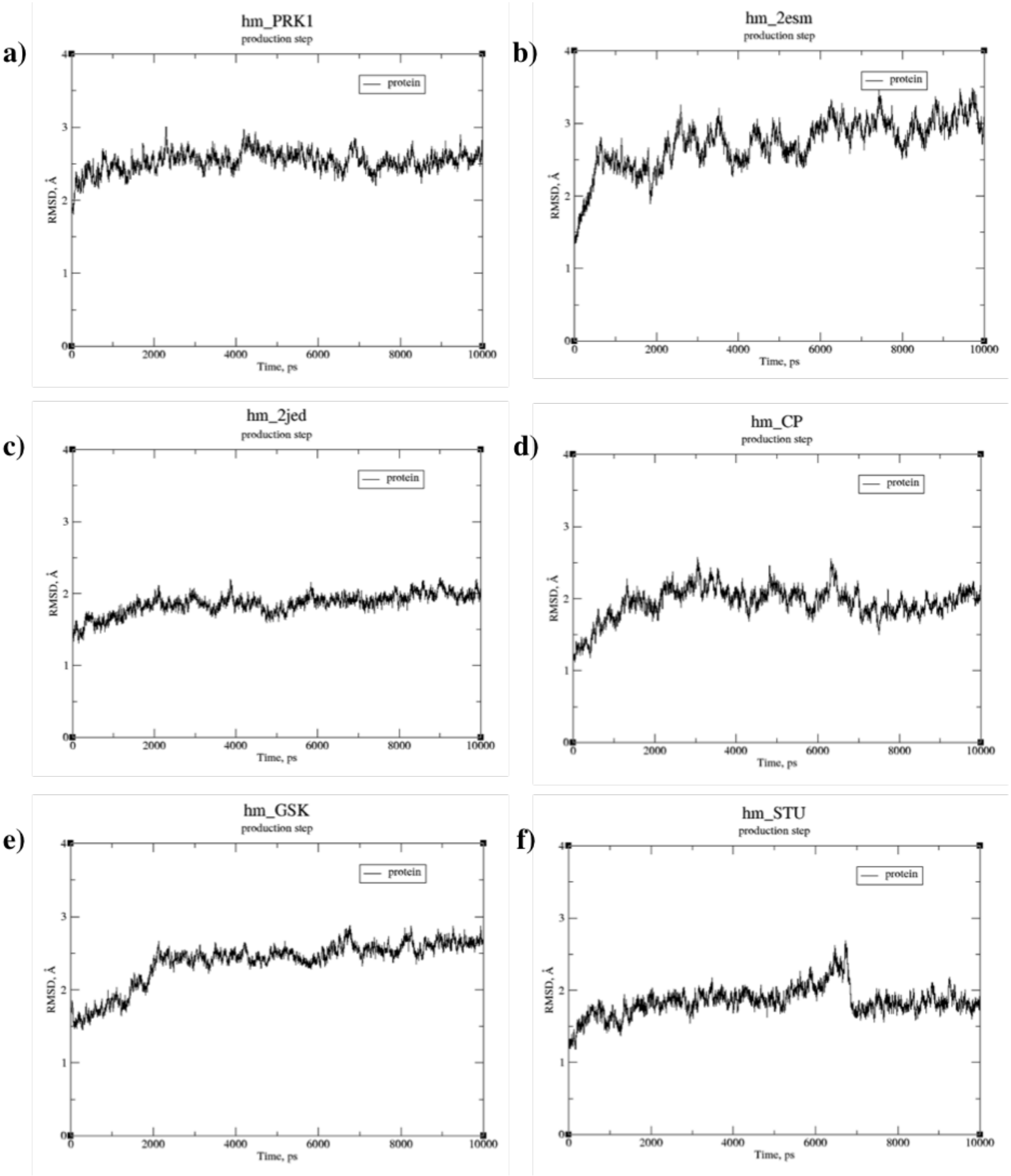


Figure 2.4: RMSD plot for MD simulations of different PRK1 homology models: **a)** hm_-prk1, initial PRK1 homology model generated using template PKC-theta 2JED; **b)** hm_2esm refined with ligand HA-1077; **c)** hm_2jed refined with ligand Nvp-Xaa228 (BIM1 analog); **d)** hm_CP refined with ligand CP-690550; **e)** hm_GSK refined with ligand GSK-690693; **f)** hm_STU refined with ligand staurosporine. The black line represents RMSD calculated for protein backbone heavy atoms (Cα, C, N).

2.3 Structure-Based Virtual Screening

2.3.1 Molecular Docking

Molecular docking is one of the main structure-based methods used in drug design and also in this study. It is a powerful and cost-effective technique, which intents to predict the geometry and strength of interaction between a ligand and a biological target. Docking algorithms are able to virtually screen large compound libraries in a reasonable amount of time, giving as a result a list of ranked compounds. This allows one to prioritize and select best-hit compounds for biological testing. Notwithstanding, limitations and drawbacks in molecular docking exist. For example, most of docking programs do not take into consideration protein flexibility, thus, neglecting ligand-induced conformational changes in proteins, which can be critical for the correct ligand binding. In general, the best docking quality is received when the docked compound is similar to the one in the used crystal structure.

There are many different docking tools based on various algorithms for scoring and pose generation, which have been investigated in recent years.[113–115] The two critical parts of docking process are the ability to predict the proper pose of a bound ligand and the correct binding strength, which should distinguish ligands from non-binding "decoys". Despite the efforts applied for improvement of various scoring functions, they still fail to predict binding affinities accurately.[51] However, a particular scoring function developed for a given protein target, can perform very well.[116]

Various approaches were suggested for the improvement of the docking performance:

- docking to several protein conformations, so-called "ensemble docking"
- rescoring experiment, where poses generated by docking with one scoring function are evaluated using another scoring function [117, 118]
- the development of scoring functions for a certain target class, etc.

A large number of docking programs have been developed in last years. In the current work the docking calculations were performed using the Glide program implemented in Schrödinger Suite 2012,[119] since different studies evaluating docking performance demonstrated that this program often shows the best docking accuracy.[113–115, 120–122] Similar observations were made for PRK1 docking taking other programs such as GOLD[123–125] or ParaDocks.[126, 127] The following paragraph describes the docking settings used in this work.

The receptor grid box coordinates were centered on the ligand in the binding pocket and the hydrogen bond constraint was defined at the hinge residue Ser698. All other parameters for the grid generation were kept as default. The ligands were prepared using LigPrep[128] within the Schrödinger utility MAESTRO by generation of ionization and tautomeric states at pH 7.4 and with less than 10 low-energy ring conformations. All ligands were energy-minimized using the MMFFs force field implemented in MAESTRO. The Glide standard precision mode (GlideSP) was used for flexible ligand docking. The options were set to penalize non-planar conformations of amide bonds and to enhance planarity of conjugated pi groups. In order to optimize highly strained ligand geometries, the post-docking minimization including five poses per ligand was applied. Compounds were docked into each of the protein structures and the top-scoring pose for each ligand was kept for further analysis. In case of ensemble docking, the docking results for different homology models were merged and the best-scored solution for each ligand was selected.

2.3.2 Pharmacophore-Based Virtual Screening

Another structure-based method used in this work is pharmacophore-based VS. The concept of pharmacophore is defined as "an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response".[129] The pharmacophore chemical features are represented by hydrogen bonds, charge interactions and hydrophobic areas. The pharmacophore-based methods are generally categorized as ligand-based and structure-based:

- In the *ligand-based pharmacophore* model the chemical features essential for bioactivity are extracted and superposed from a set of active molecules. This method requires a set of known reference compounds.
- The *structure-based pharmacophore* model requires the 3D structure of macromolecule or macromolecule–ligand complex and the pharmacophoric features are defined as interaction points between the macromolecular target and ligand.

At this stage of the work the structure-based pharmacophore modeling was the method of choice, because this tool provides detailed structural information, which is essential for understanding the protein-ligand interactions and it can be helpful in further lead optimization. Moreover, this approach has a number of advantages:

- it is universal (pharmacophore can represent not only bound, but also unknown molecules)
- it is simple (the method ensures computationally efficient virtual screening)
- it is comprehensive and editable (features can be added and removed, feature tolerance can be tuned)

The generated pharmacophore model can be used as a query for further virtual screening of 3D chemical databases in order to identify molecules with similar chemical features. This process is called "pharmacophore-based virtual screening". In the current study the *inte:ligand* software LigandScout[130] was used for pharmacophore generation and VS. It is an efficient tool, which has been successfully applied in several virtual and experimental HTS projects.[131]

In order to identify molecules with similarity to PD-0166285, an inhibitor identified by our *in vitro* screening, a structure-based pharmacophore[129] model was used. For the generation of the pharmacophore query we used the docking solution of compound PD-0166285 in the PRK1 model. The Inte:Ligand software LigandScout[130] was used for pharmacophore generation and virtual screening. A seven-point pharmacophore consisting of such chemical features as hydrogen bonds, charge interactions and hydrophobic areas was automatically generated using LigandScout. Furthermore, several features were enabled as optional, i.e. positive charge on the amine group and hydrophobic areas on Cl atoms of compound PD-0166285 (see Figure 2.5). Exclusion volumes were generated to take the shape of the binding pocket into account. The generated pharmacophore model was used as a query for further virtual screening of different 3D chemical databases.

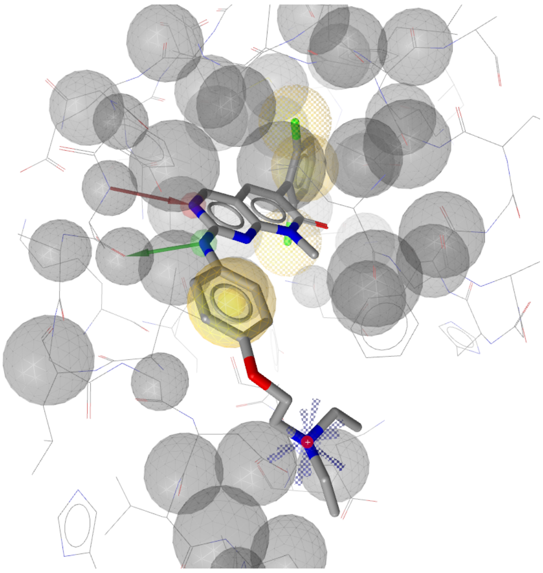


Figure 2.5: Pharmacophore model of compound PD-0166285 generated and refined using the program LigandScout. Hydrogen bond features are shown as green (donor) and red (acceptor) arrows; positive ionizable feature is shown in blue color; hydrophobic and exclusion volumes are indicated as yellow and grey spheres, respectively.

2.4 Molecular Dynamics Simulations

MD simulations can be used for the characterization of energetic and dynamics of protein-ligand complexes at atomic resolution. In MD simulations, Newton's equation of motion is solved numerically for an atomic resolution representation of the system including surrounding solvent and ions based on a force field description of intra- and intermolecular interactions. Both the protein and ligand molecule are fully flexible during the simulation, which is often neglected by the docking process.

In this work MD simulations were used to examine the stability of the homology models and of the derived PRK1-inhibitor complexes as well as for binding free energy estimation. The MD simulations were carried out using AMBER 12[112] and the AMBER ff99SB[132] force field. The molecular mechanics (MM) parameters for the ligands were assigned with antechamber, divcon and parmcheck modules.[133, 134] The explicit solvent model TIP3P was used.[135] The protein or protein-ligand complex was placed in a box of water molecules with a margin of 10 Å. Four steps of minimization each of 3000 iterations (500 of steepest descent and 2500 of conjugate gradient) were carried out. In the first step water molecules and ions were minimized while protein atoms (protein plus ligand atoms in case of complexes) were restrained to the initial coordinates with a force constant 10 kcal·mol^{-1}·Å^{-2}. In the second step, the protein side-chains were minimized together with the solvent (protein backbone and ligand atoms were restrained). During the next step, only a weak restrain of 0.01 kcal·mol^{-1}·Å^{-2} was applied on the protein backbone (ligand was released). Finally, all restraints were removed and the entire system was minimized.

The temperature of the system was then equilibrated at 300 K through 100 ps of MD using 2 fs time steps. A constant volume periodic boundary was set to equilibrate the temperature of the system by the Langevin dynamics[136] using a collision frequency of 2 ps^{-1}. During the temperature equilibration routine, the protein/complex in the solvent box was restrained to the initial coordinates with a weak force constant of 10 kcal·mol^{-1}·Å^{-2}. To avoid inaccurate calculations of pressure, the next 100 ps of MD were run at constant volume. The pressure of the solvated system was equilibrated at 1 bar in constant pressure periodic boundary by an isotropic pressure scaling method with pressure relaxation time of 1 ps.

Before running free MD simulations, the system was slowly equilibrated in four steps (each 100 ps) using constraints like in minimization steps. The time step was set to 2 fs with a cut-off of 9 Å for the non-bonded interaction, and SHAKE[137] option was employed to keep all bonds involving hydrogen atoms rigid. Electrostatic interactions were computed using the Particle Mesh Ewald method.[138] Each MD simulation was performed for 10 ns.

2.5 MM-PBSA and MM-GBSA Approaches for BFE Calculations

The free energy calculations were used to calculate the relative BFE of various ligands using the MM-PB(GB)SA methods. The MM-PB(GB)SA approach employs molecular mechanics (MM), the generalized Born (GB) or the Poisson Boltzmann (PB) model and solvent accessibility (SA) method to compute the free energy of a biomolecule by combining molecular mechanics calculations and continuum solvation models. The molecular dynamics coupled with MM-PB(GB)SA estimation is a valuable tool for the investigation of the experimental protein-ligand binding affinities and can serve as a powerful tool for the predicting of correct ligand ranking.[75]

According to the MM-PB(GB)SA method, the BFE between ligand and receptor ($\varDelta G_{bind}$) is computed as a sum of the changes of the gas-phase molecular mechanics energies ($\varDelta E_{MM}$), polar and nonpolar solvation energy ($\varDelta G_{solv}$) and conformational entropy ($-T\varDelta S$) upon binding, see Equation (2.1):[139, 140]

$$\varDelta G_{bind} = \varDelta E_{MM} + \varDelta G_{solv} - T\varDelta S \tag{2.1}$$

$$\varDelta E_{MM} = \varDelta E_{internal} + \varDelta E_{vdW} + \varDelta E_{ele} \tag{2.2}$$

$$\varDelta G_{solv} = \varDelta E_{PB/GB} + \varDelta G_{SASA} \tag{2.3}$$

$$\varDelta G_{SASA} = \gamma SASA + b \tag{2.4}$$

Correspondingly, $\varDelta E_{MM}$ includes the differences in bond, angle, dihedral energies ($\varDelta E_{internal}$), van der Waals ($\varDelta E_{vdW}$) and electrostatic energies ($\varDelta E_{ele}$), see Equation (2.2). The polar solvation contribution is usually computed by solving the generalized Born (GB) or Poisson-Boltzmann (PB) equation, while the nonpolar term is determined as a linear function of the solvent-accessible surface area (SASA),[141] see Equations (2.3) and (2.4).

The entropic contribution to binding $-T\triangle S$ can be calculated using normal-mode analysis on a set of conformational snapshots taken from MD simulations. However, such calculations are computationally expensive and usually give large error bars. Thus, the entropic contribution was not evaluated in this work assuming that it is equivalent for a congeneric ligand series.[67, 142, 143]

In this work, the single trajectory approach for MM-PB(GB)SA calculation is used. It means, that MD simulation is performed on the protein-ligand complex only, and the snapshots taken from this trajectory are used for the free energy calculation following the equations described above.

2.5.1 BFE Calculations for Compounds from Biomol Kinase Library

The MD and BFE evaluation were performed on 14 top-scored molecules received after the docking of the Biomol kinase library into the homology model of PRK1 (hm_STU) in order to see whether this approach is able to estimate the biological activity of compounds. The Biomol library was used because the experimental data were available from the previous work.[144] The protocol for complex preparation and MD simulations was already described in Section 2.4. The MD simulation for each complex was performed on an ensemble of PRK1-ligand complexes in water for 10 ns (RMSD plots for proteins and ligands are shown in Figure A.1, see Appendix). MM-PBSA and MM-GBSA calculations were carried out using a) the last frame of the minimized structure or b) the MD trajectory, where the last 5 ns (50 frames) were taken for the ligand-binding affinity estimation by the calculating energy components from equations 2.1-2.4 via MM-PBSA and MM-GBSA algorithms using the MMPBSA.py script implemented in AMBER 12 software package.[112] The internal, van der Waals and electrostatic terms, representing the gas-phase molecular mechanics energies ($\varDelta E_{MM}$) were calculated using the SANDER module with infinite cut-off for the long range non-bonded interactions. The polar solvation free energy was computed via Poisson-Boltzmann (PB) or generalized Born

(GB) solvation models using the modified GB model developed by Nguyen and Simmerling (igb=8),[145] and the PB model developed by Tan and Luo.[146] The salt concentration for GB and ionic strength for PB were set to 0.15 M; the interior (solute) and external (solvent) dielectric constants to 1.0 and 80, correspondingly; other parameters were kept default. The non-polar solvation term was calculated according to the equation 2.4 using solvent-accessible surface area (SASA) computed via the molsurf routine with default parameters γ=0.0072 and constant b=0. The entropic contribution was not calculated, assuming that investigated compounds may exhibit similar entropy contributions to binding, since they are congeneric and bind to the same protein binding site.[67]

2.5.2 BFE Calculations for Rescoring of Docking Results

Additionally, the BFE calculations were used in this work as a post-processing tool for re-ranking of docking poses and predicting of the binding affinity. In this case, the MM-PB(GB)SA calculation is done for a single snapshot of an energy-minimized protein-ligand complex obtained after ligand docking. Such a method requires much less computational resources than traditional BFE calculations after performing MD simulations, therefore it can be applied for the rescoring of thousands docking solutions after virtual screening. Furthermore, the rescoring using MM-PB(GB)SA has shown to be more accurate than simple scoring, since it considers protein flexibility and solvation effects. Recent studies have demonstrated that this method results in a significant correlation with experimental data and in many cases it can be applied for the identification of the correct ligand binding mode.[75, 81, 82, 142]

The next paragraphs describe the settings used for the BFE calculations for the rescoring of the docking results. The selected top-scored docking solutions for each compound with the corresponding model were subjected to further refinement by means of energy minimization and rescoring using three approaches for BFE calculations, namely MM-PBSA, MM-GBSA and QM/MM-GBSA. Force field parameters were assigned for ligand and protein using the Leap module in AMBER12.[112] General Amber Force Field

(GAFF)[133] and AM1-BCC[147] charges were used for ligands and AMBER ff99SB[132] force field for proteins. The minimization of protein-ligand complexes was done in an implicit solvent. The complex was minimized for 5000 steps of steepest descent and 5000 steps of conjugate gradient algorithm. Next, the binding free energy calculations were performed. Only one snapshot per complex derived after minimization was used for it. The MMPBSA.py script implemented in AMBER 12 was used for the calculation of energy components from equations (1)-(3). The internal, van der Waals and electrostatic terms, representing the gas-phase molecular mechanics energies (ΔE_{MM}) were calculated using the SANDER module with 16 Å cut-off for the long range non-bonded interactions. The polar solvation free energy was computed via implicit solvation models like Poisson-Boltzmann (PB) or generalized Born (GB) applying the following parameters: modified GB model developed by Onufriev et al.[148] (GBOBC2, igb=5), salt concentration for GB and ionic strength for PB of 0.15 M, internal dielectric constant of 1.0 for solute and 80 for implicit PB solvent, other parameters were kept default. The nonpolar solvation term was determined from the solvent-accessible surface area (SASA) computed via the molsurf routine following equation (4), where γ was set to 0.0072 and constant b to 0 (default values).

The combined quantum mechanics/molecular mechanics generalized Born (QM/MM-GBSA) calculations were done analogically to MM-GBSA with only one difference – the QM part was specified for the ligand and was treated with semi-empirical method (RM1),[149] while the remaining atoms were treated with ff99SB force field. Representing the ligand using QM method allows to eliminate general problem of ligand force field deficiencies. Analogically to the previous BFE calculations (see 2.5.1), the entropic contribution to binding was not evaluated.

2.6 PRK1 *in vitro* Assay

Screening was carried out using the LanthaScreen^{TM} Eu Kinase Binding Assay Kit (Invitrogen) with final assay concentrations of 5 nM for PRK1 (Proqinase, Freiburg), 2 nM LanthaScreen Eu-anti-GST Antibody (Invitrogen) and 10 nM Kinase Tracer 236 (Invitrogen). The assay was performed in 384-well microtiter plates (Perkin Elmer, Rodgau) with a final assay volume of 15 μL (1% (v/v) DMSO). Detection was performed with EnVision 2102 Multilabel Reader (PerkinElmer, Rodgau, Excitation: 340 nm, 1st Emission: 665 nm, 2nd Emission: 615 nm, Delay Time 100 μs, Integration Time 200 μs). IC_{50} values were determined using Graphpad Prism 5.0 (La Jolla, USA), where the half maximal inhibitory concentration (IC_{50}) is the concentration of compound at which PRK1 activity is inhibited by 50%.

CHAPTER 3

Results and Discussion

3.1 PRK1 Inhibitors Identified by *in vitro* Screenings

As of October 2010, when the project started, only few PRK1 inhibitors were described in literature. Some of the well-known PKC inhibitors such as staurosporine [150], Ro-318220 [42], bisindolylmaleimide (BIM) I and H-7 [151], as well as Rho kinase/ROCK inhibitors Y-27632 [152], HA-1077 and Y-32885 [153] were reported to inhibit also PRK1 (see Figure 3.1). However, these compounds are known as non-selective kinase inhibitors. Since they inhibit PRK1 in an ATP-competitive manner, they may interact with the catalytic domain of PRK1-related kinases in a similar way.

In order to identify novel PRK1 inhibitors, a focused library screening was initiated and performed by the group of Prof. Jung at the University of Freiburg (Institute of Pharmaceutical Sciences). The screening set included the Biomol kinase and phosphatase inhibitor library (n=84) as well as an in-house library of commercially available and generic compounds (n=200) which have been used to profile different kinases before. The initial screening was done at 100 nM inhibitor concentration. It identified highly potent nanomolar PRK1 inhibitors such as staurosporine and related analogs - K252a, lestaurtinib, PKC-412, Ro-318220 and BIM I.[144] Later on, few compounds from Biomol dataset were retested at higher concentrations based on the results of docking and BFE

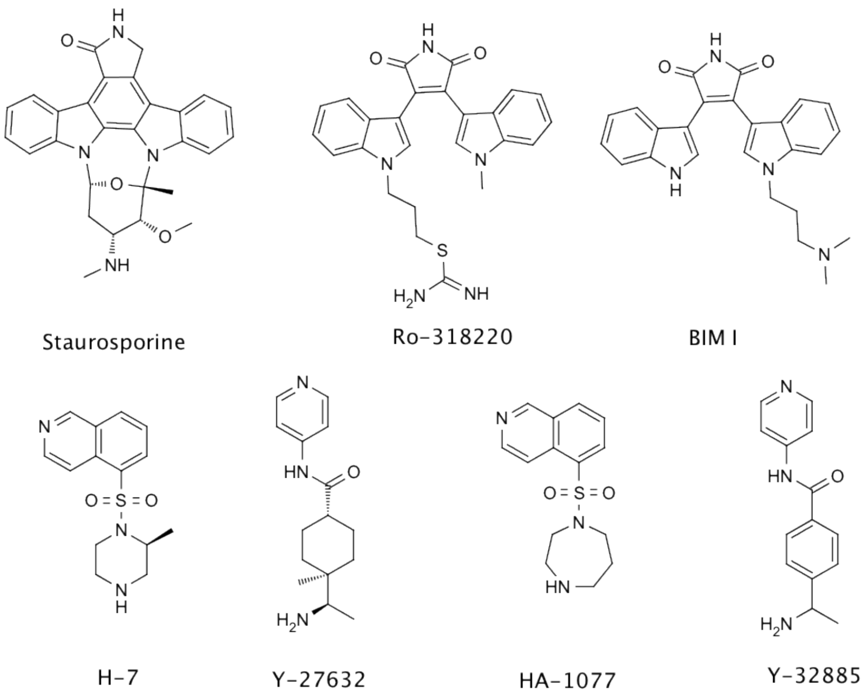


Figure 3.1: PRK1 inhibitors known from the literature as of 2010.

calculations (see Section 3.3 for more details), which identified low micromoloar inhibitors H-7 and HA-1077 (see Figure 3.2a and Table 3.3 for activities).

Succeeding *in vitro* screenings performed over the next few years identified a number of PRK1 inhibitors with different scaffolds and potency, see Figure 3.2. Their activities are summarized in Table 3.3. Depending on the information available at a certain time, different *in silico* approaches (ligand- and structure-based) were applied at various project stages. They are described in more detail in the following sections.

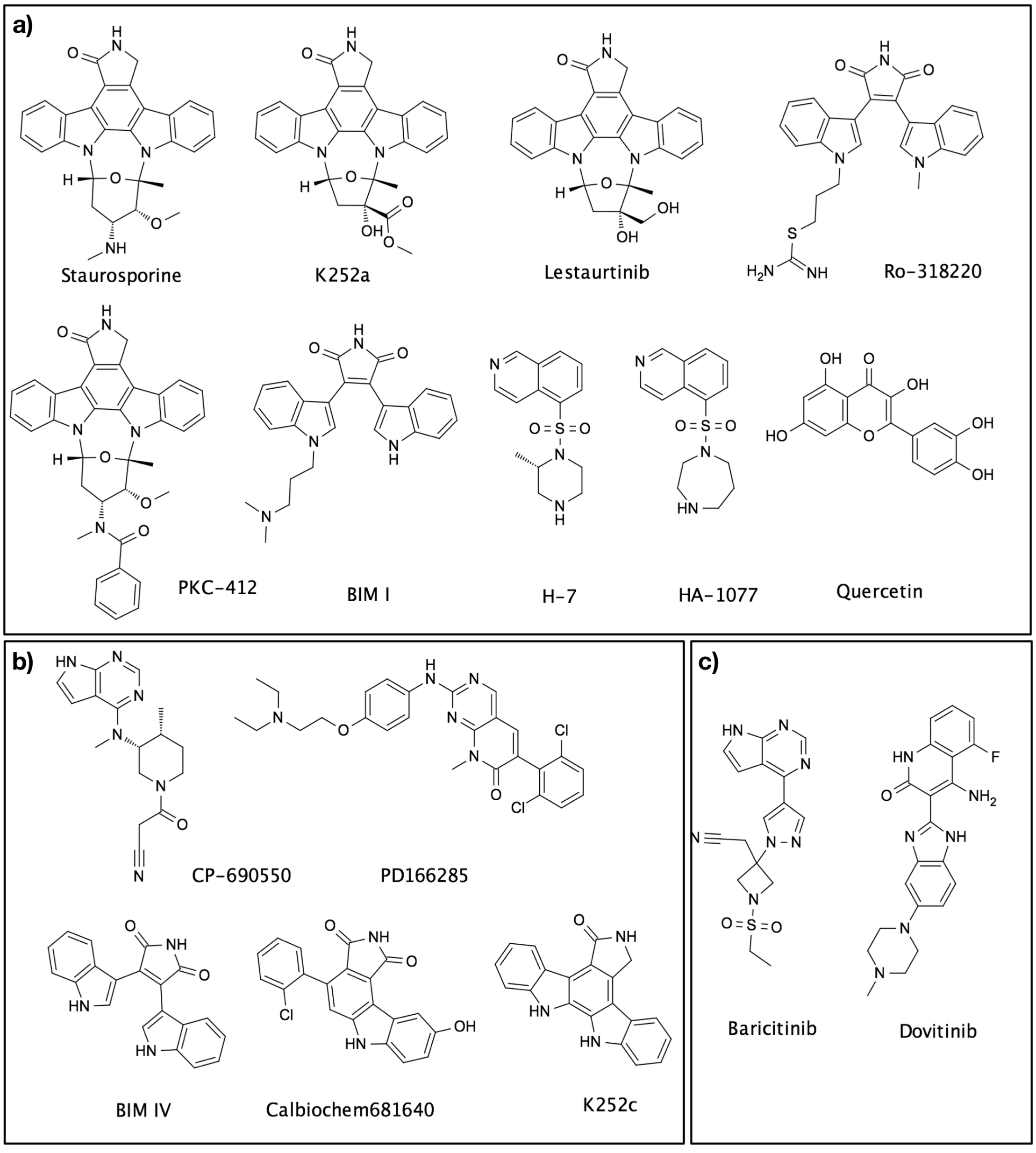

a)
Staurosporine
K252a
Lestaurtinib
Ro-318220
PKC-412
BIM I
H-7
HA-1077
Quercetin
b)
CP-690550
PD166285
BIM IV
Calbiochem681640
K252c
c)
Baricitinib
Dovitinib

Figure 3.2: PRK1 inhibitors identified by in-house *in vitro* screenings performed by the group of Prof. Jung at the University of Freiburg (Institute of Pharmaceutical Sciences): **a)** focused library screening of Biomol database 2010; **b)** screening 2012; **c)** screening 2013.

3.2 PRK1 Homology Modeling

3.2.1 PRK1 Homology Modeling using PKC-theta as a Template

The three dimensional structure of the PRK1 molecule is the prerequisite for structure-based drug design. Despite its early identification and importance in cancer research there was no crystal structure of PRK1 catalytic domain available until recently (PRK1 structures 4OTD, 4OTI, 4OTG, 4OTH were released in the PDB on 27th August, 2014, see Section 3.7). Thus, the homology modeling methodology was used to predict the 3D structure of PRK1.

Studies have shown that protein structure can often be predicted with an accuracy equivalent to a low-resolution X-ray structure, when a template with over 30% of protein sequence identity is given.[154] Since PRK1 shares high sequence similarity with PKC family (around 50%, see Table 1.1) and approximately 40% sequence similarity with other kinase families such as protein kinase A (PKA) and protein kinase B (PKB), its structure can be reliably predicted.

The NCBI BLAST search identified structures available in the PDB which share the high similarity with PRK1 and which belong to the PKC family. Among them, crystal structure of PKC-theta (PDB code: 2JED) was chosen as a template as it shows high sequence identity with PRK1 (50%), good resolution (2.32 Å) and complete set of coordinates. The PRK1 homology model (hm_prk1) was built, refined and evaluated as described in Methods Section 2.2.1.

3.2.2 Refinement of PRK1 Homology Model with Ligands in the Binding Pocket

The secondary structure prediction is an easy task for the program when a highly similar template sequence is given. However, the conformation of side chain residues can differ from ones in the crystal structure. In fact, even small changes in the receptor structure can be important for ligand binding. Since in most drug discovery frameworks the protein is kept rigid, it is necessary to develop a high-quality model of the binding site.

Initially, a homology model (hm_prk1) was generated based on the template PKC-theta (PDB code: 2JED), which shares high sequence identity with PRK1. However, not all PRK1 inhibitors could be docked into this initial model correctly. The reason for that can be that the binding pocket of the homology model adopts another conformation that certain compound requires for tight binding. It is known that during ligand binding conformational changes of the protein can occur. Proteins adapt their structure to the bound ligand. Thus, the docking method can fail to recognize an active ligand when a protein structure was solved in the presence of a very different compound. One possible solution to overcome this problem is to build a homology model, which would have a conformation of the binding pocket adopted to the presence of a given ligand.

Thus, the conformation of residues in ATP-binding pocket of model hm_prk1 was adjusted to the presence of several structurally different inhibitors. As a template for this refinement the conformation of binding pocket and ligand was taken from crystal structures of homologous kinases. In this way, six PRK1 homology models were generated (see Table 3.1 for details). All models were obtained by refinement of the initial PRK1 homology model (hm_prk1) by including active ligands into the binding pocket using the program MODELLER 9v8 as described in Methods Section 2.2.2. Using an ensemble of induced fit models increases conformational sampling of the binding site, which leads to improved pose prediction and scoring performance.

Table 3.1: Model names, binding pocket template structures and inhibitors used for refinement of the homology models. The initial model was generated on the basis of PKC-theta as template structure.

Model name	**Crystal structure used for refinement (PDB ID)**	**Binding pocket template**	**Inhibitor**
hm_2esm/ hm_2esm_l	2ESM	ROCK1	HA-1077
hm_2jed	2JED	PKC-theta	Nvp-Xaa228 (BIM1 analog)
hm_CP	3LXN	TYK2	CP-690550
hm_GSK	3D0E	AKT2	GSK-690693
hm_STU	1XJD	PKC-theta	Staurosporine

The superposed structures of the six PRK1 homology models are shown in Figure 3.3. The models were superimposed by backbone Cα atoms, as shown in Figure 3.3a, which resulted in RMSD values in the range of 1.26-2.55 Å. The ATP-pocket residues superimposed by all atoms in the binding site are shown in Figure 3.3b (RMSD in the range of 0.74-2.50 Å). All models share a high similarity of the ATP binding site and only minor deviations are observed for some side chains and the G-loop.

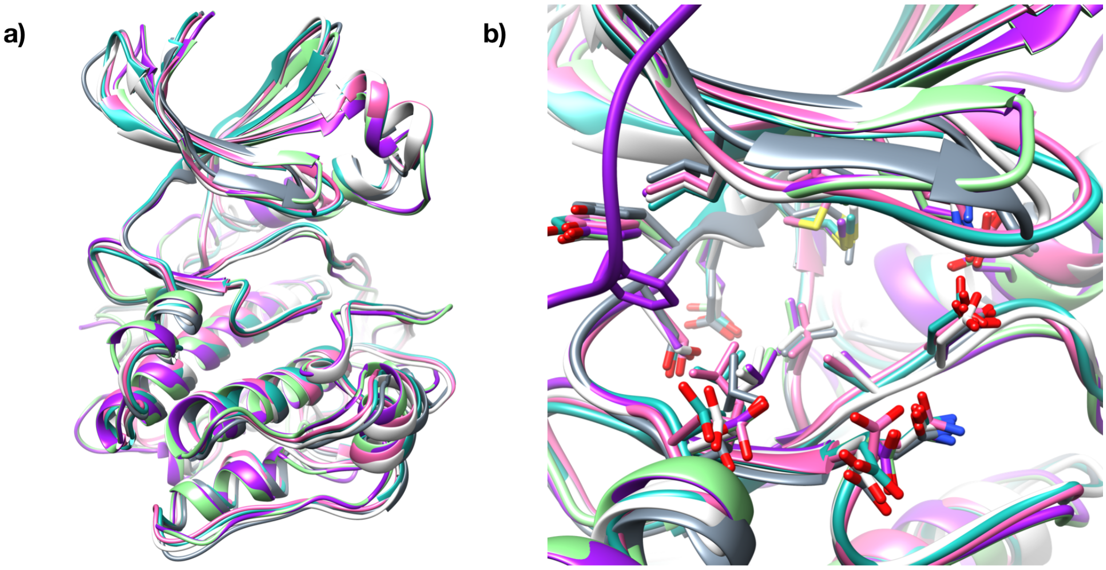


Figure 3.3: Comparison of the generated PRK1 homology models: **a)** structures of six PRK1 homology models, shown as colored ribbons, after superimposition on the backbone Cα atoms (the C-tail is omitted for clarity); **b)** conformation of ATP-binding site residues from different PRK1 homology models superimposed by using all pocket residues (the C-tail flanking ATP-binding pocket is shown).

It is worth to note that PRK1, as well as other members of AGC kinase family (e.g. PKA, PKC, RSK, SGK, GRK and PKB) are characterized by the presence of a C-terminal regulatory region (C-tail),[155] which is involved in the regulation of enzyme activity.[156, 157] The C-tail is flanking the ATP-binding pocket and can insert the conserved phenylalanine residue into it (see Figure 3.3b), while many other kinases have open solvent channel at this place. However, the C-terminus of the catalytic domain of AGC kinases is often disordered and not visible in the electron density. Due to the flexibility and lack of structural information, the C-tail (residues 887-940) was not considered in most models except in one (hm_2esm_1). The docking experiments showed that certain conformations of this loop could hinder the docking of some inhibitors which extend its interactions into the solvent-accessible channel (e.g. TAE-684, VX-680 or sunitinib). Furthermore, this loop can undergo large rearrangements, thus, it is often not resolved in crystal structures of protein kinases, what impedes its modeling. Nevertheless, the analysis of complexes with isoquinoline derivatives available in PDB showed that a certain conformation of Phe904 of the C-terminal loop (see Figure 3.4) can play an important role in pose identification and scoring of this type of inhibitors, which is discussed later. In order to take this possibility into account, we decided to keep one model (hm_2esm_l) containing C-tail residues.

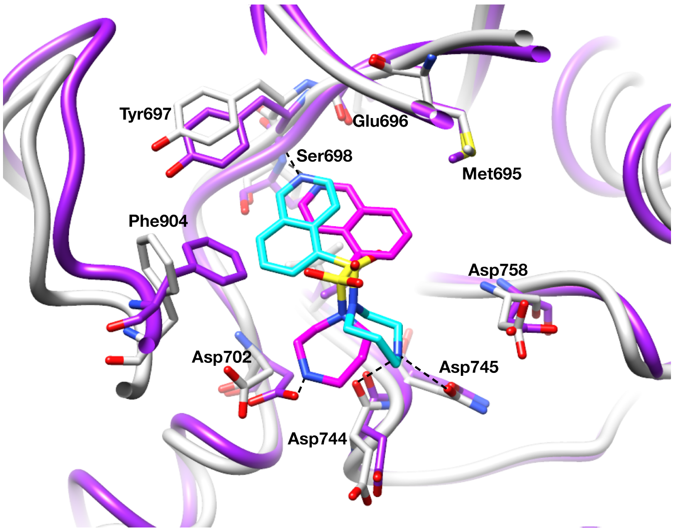


Figure 3.4: Comparison of the ATP-binding pockets of two PRK1 homology models - hm_STU refined with staurosporine in the binding site (light grey ribbon and sticks; the C-terminal flexible loop, modeled from PKC theta 2JED, is shown in figure for comparison, but it was deleted due to uncertainty in its structure) and hm_2esm_l refined in the presence of HA-1077 (purple ribbon and sticks). The docking solutions of compound HA-1077 with hm_STU and hm_2esm_l are shown in cyan and magenta sticks, respectively. Hydrogen bonds between inhibitor and the kinase are displayed as dashed lines.

3.3 Molecular Dynamics and BFE Calculations for Top-Scored Biomol Compounds

As described in Section 3.1 the initial screen of Biomol kinase and phosphatases inhibitor library identified only staurosporine analogs as highly potent PRK1 inhibitors (IC_{50} < 100 nM). Thus, in the next step, we performed *in silico* docking studies for Biomol database in order to test and validate initially generated PRK1 homology models. The binding modes of known kinase inhibitors received after the docking into PRK1 were compared to those solved crystallographically with related kinases. Short conclusions include following:

1) Staurosporine and its analogs have massive rings (indole, carbazole), which usually extend into the binding pocket much farther than the adenosine moiety of ATP and are known to cause conformational changes of neighboring enzyme residues known as "induced fit".[158] Thus, these compounds could not have been docked correctly into initial PRK1 homology model (hm_prk1) generated using default Modeller settings. Thereby, the binding pocket of the homology model was refined for the presence of staurosporine and hm_STU was generated (see Sections 2.2.2 and 3.2.2 for details), which resulted in better docking poses for this compound class.
2) The compounds with high molecular weight (here staurosporine and analogs) usually show good docking scores. It is common limitation of scoring functions to favor towards larger molecules [159] and accurate binding-affinity prediction still remains a challenging issue in molecular docking.[160]

Interestingly, the docking of Biomol database into PRK1 homology model refined with staurosporine (hm_STU) revealed another compound class (isoquinoline derivatives) among the top-scored molecules. In order to estimate the stability of the complexes and ligand binding affinity we decided to perform molecular dynamics simulation with subsequent BFE calculations for fourteen compounds from the Biomol screening set top-ranked by GlideSP score. The methodology is described in detail in Methods Sections 2.4 and 2.5.1. Since the bioassay threshold of the initial screening was set to 100 nM, several compounds have been retested at the concentration up to 5 μM based on their low values of calculated binding affinities according to generalized Born or Poisson-Boltzmann approaches (GBTOT and PBTOT score, respectively) as well as their availability. Thus, four compounds were retested (SP600125, HA-1077, H-7 and SB216763) and two of them, namely isoquinoline derivatives HA-1077 and H-7, were shown to inhibit PRK1 with IC_{50} values of 1945 nM and 658 nM (pIC_{50} 5.71 and 6.18, correspondingly), confirming the binding free energy calculation results, see Table 3.2.

Table 3.2: MM/PB(GB)SA calculations for 14 top-scored by GlideSP compounds from Biomol set. The BFE calculations were done using 50 frames taken from MD simulation according to generalized Born or Poisson-Boltzmann approaches (the scores are abbreviated as "GBTOT md 50f" and "PBTOT md 50f", respectively) as well as using only one frame after the complex minimization before the MD was performed ("GBTOT Emin 1f" and "PBTOT Emin 1f"). The compounds which were retested on PRK1 are selected with a red frame.

Name	**pIC_{50}**	**glideSP score**	**GBTOT md 50f**	**PBTOT md 50f**	**GBTOT Emin 1f**	**PBTOT Emin 1f**
			kkal/mol			
Staurosporine	9.10	-12.17	-58.48	-56.78	-53.77	-57.66
Lestaurtinib	8.07	-11.74	-45.58	-37.69	-55.81	-47.08
K252a	8.49	-11.04	-47.77	-38.44	-49.53	-29.55
Ro-318220	7.11	-10.82	-41.05	-31.01	-46.95	-42.37
PKC-412	7.85	-10.68	-46.78	-37.51	-49.11	-32.06
SP600125		-9.53	-29.09	-23.20	-27.16	-24.79
H-9		-8.91	-20.81	-36.06	-36.04	-56.44
HA-1004		-8.46	-25.08	-29.58	-36.74	-34.61
HA-1077	5.71	-8.44	-26.82	-29.34	-30.68	-37.79
H-8		-8.33	-20.22	-21.41	-20.61	-25.69
PD-98059		-8.23	-23.47	-15.94	-25.20	-18.91
H-7	6.18	-8.18	-20.29	-25.79	-27.42	-29.33
LY-294002		-8.12	-32.36	-28.83	-29.82	-25.57
SB216763		-7.99	-36.82	-27.19	-39.34	-28.94
r^2		**0.86**	**0.90**	**0.75**	**0.81**	**0.22**

Furthermore, it is worth noting that BFE scores were able to discriminate highly potent (IC_{50}<100 nM) PRK1 inhibitors from moderately potent, thus, low nanomolar inhibitors staurosporine and analogs usually show GBTOT scores in a range of -40 kkal/mol or lower, and PBTOT scores less than -30 kkal/mol. Interestingly, GlideSP scoring gave a good correlation ($r^2 = 0.86$) with experimental activities, see Figure 3.17a. Typically, docking scoring functions have been designed for generating an accurate poses for ligands and database enrichment. Nevertheless, studies have shown that in some cases the scoring function, e.g. Glide, can result in a good correlation with experimental activities, especially, for simple cases when the set of ligands share a common structure.[161, 162]

In addition, a strong correlation ($r^2 = 0.90$) is observed between the calculated BFE scores "GBTOT md 50f" and experimentally-derived pIC_{50} values for seven PRK1 inhibitors from Biomol kinase set, see Figure 3.17b.

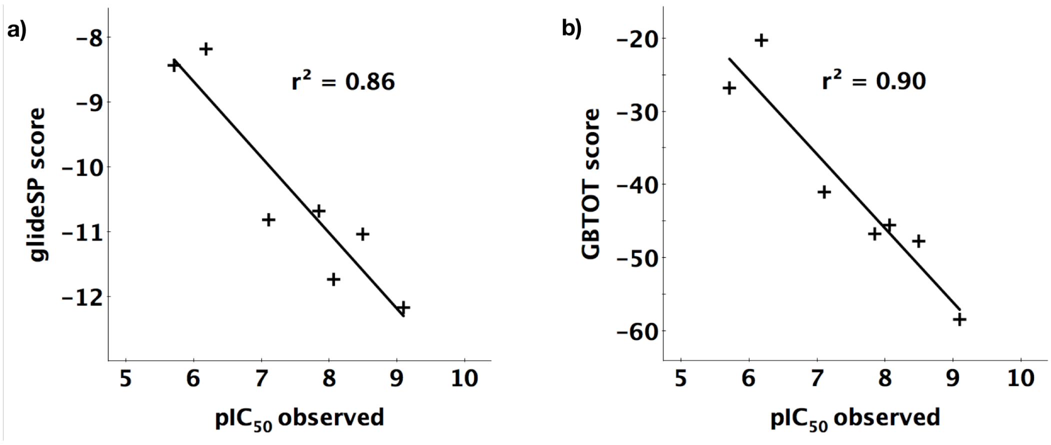


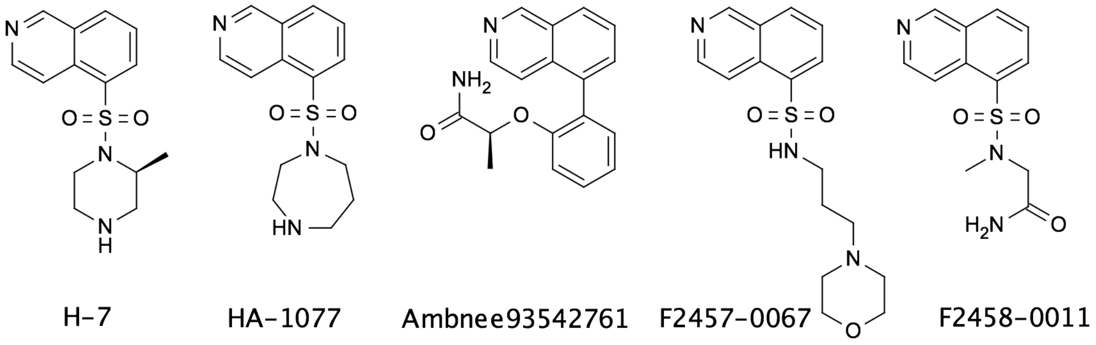
Figure 3.5: Regression and corresponding correlation coefficients r^2 are shown between the observed pIC_{50} of 7 compounds from Biomol database top-scored by GlideSP versus: **a)** glideSP docking score; **b)** binding free energy score (GBTOT) calculated by MM-GBSA method using 100 snapshots after molecular dynamics simulation.

3.4 Initial Virtual Screenings to Search for PRK1 Inhibitors

Considering that at the start of the project most of the known PRK1 inhibitors from the literature or in-house library screening were staurosporine analogs, we decided to search for compounds with a distinct scaffold. Therefore we took the isoquinolines HA-1077 and H-7, the pyrrolopyrimidine CP-690550 and the pyridopyrimidinone derivative PD-0166285, which were identified by *in vitro* screenings at different project stages, and carried out similarity- and pharmacophore-based virtual screenings.

3.4.1 Similarity-Based Virtual Screening for Isoquinolines HA-1077 and H-7

First, a fast fragment-based similarity screening method MOLPRINT 2D,[96, 97] which is based on atom environments, was applied for a search of compounds similar to the PRK1 inhibitors HA-1077 and H-7. This method is based on ligand information only and follows the molecular similarity principle, i. e. the compounds that are more similar with respect to their chemical structures are more likely to possess similar properties. After the screening of ZINC [163] drug-like database 1000 hits with the highest Tanimoto score were selected. After the correction of protonation states and removing of duplicates from the database, the 825 compounds left were docked with GlideSP into the PRK1 model refined with ligand HA-1077 (hm_HA). The binding poses of the best-ranked molecules were visually analyzed and 23 compounds were selected for purchase. However, only four of them were finally available from commercial suppliers for biological testing. These compounds, like HA-1077, comprise an isoquinoline scaffold and were evaluated for PRK1 inhibition. Experimental results revealed that three of these compounds (Ambnee93542761, F2458-0011, F2457-0067, see Figure 3.6) are weak micromolar inhibitors of PRK1 with IC_{50}s below 100 μM (hit rate 75 %), showing worse activity on PRK1 than the initial compounds HA-1077 and H-7 used in this virtual screening. Their activities are summarized in Table 3.3.

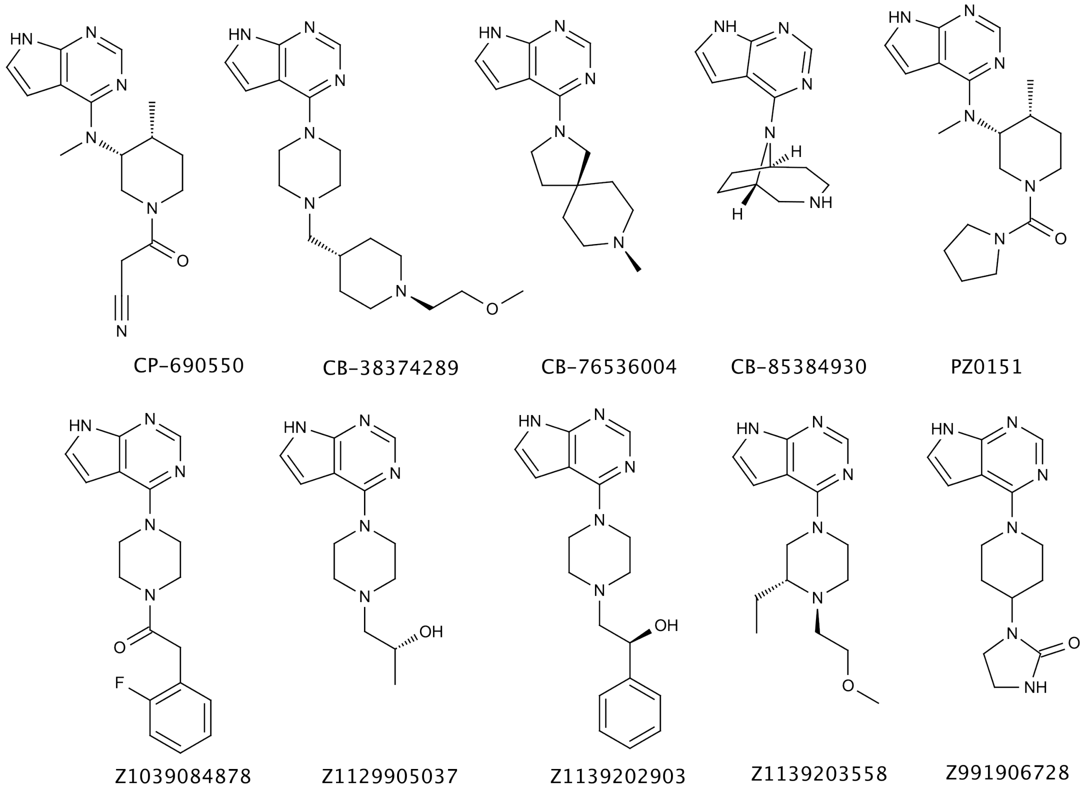


H-7 HA-1077 Ambnee93542761 F2457-0067 F2458-0011

Figure 3.6: Isoquinoline derivatives H-7, HA-1077 and inhibitors identified by similarity search using ligand-based fingerprint method MOLPRINT 2D.

3.4.2 Similarity Search for Pyrrolopyrimidine Derivative CP-690550

A further ligand-based VS method using MACCS (Molecular ACCess System) keys fingerprints implemented in MOE [99] was applied for a search of compounds similar to the PRK1 inhibitor CP-690550 (Tofacitinib). Screening of the ZINC [163] drug-like database identified 500 compounds, which were subsequently docked into the PRK1 homology model refined in the presence of ligand CP-690550 (hm_CP). The best-scored solutions were visually inspected and 27 hits were selected. Among them 18 compounds were ordered and tested *in vitro* for PRK1 inhibition. As a result, nine pyrrolopyrimidine analogs (hit rate 50 %) showed inhibitory activity in the range of 1-56 μM (see Figure 3.7 and Table 3.3).



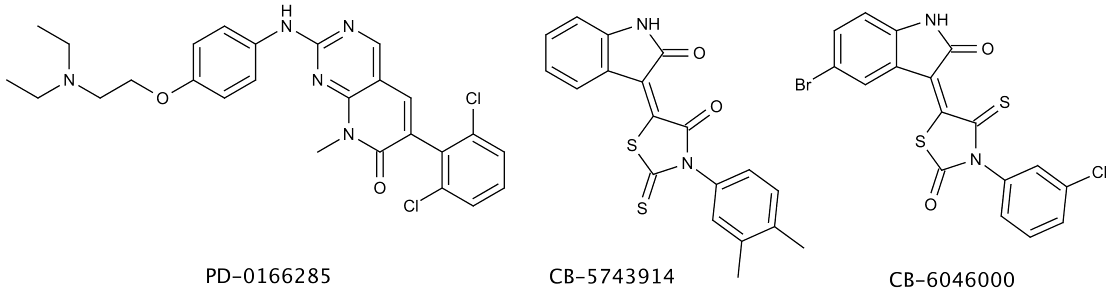
CP-690550 CB-38374289 CB-76536004 CB-85384930 PZ0151

Z1039084878 Z1129905037 Z1139202903 Z1139203558 Z991906728

Figure 3.7: PRK1 inhibitors CP-690550 and other pyrrolopyrimidine derivatives, which were identified by similarity search using MACCS fingerprints combined with docking into PRK1 homology model.

3.4.3 Pharmacophore-Based Virtual Screening for PD-0166285

With intention to identify further hits and to increase the chemical diversity of our dataset, we applied a structure-based pharmacophore model generated from another PRK1 inhibitor with a 8-methylpyrido[2,3-d]pyrimidin-7(8H)-one scaffold, namely PD-0166285. The ligand was docked into the initially generated PRK1 model (hm_prk1) and the predicted pose was compared with a reference crystal structure of an analogous compound with the related kinase phosphoinositide-dependent kinase-1 (PDK1) (PDB code: 1M52). The pharmacophore model (see Figure 2.5) was generated using LigandScout[130] as it is described in Methods Section 2.3.2. Various chemical libraries from three different suppliers (ChemBridge, Life Chemicals and Enamine) were used for the screening process. The compounds with a Pharmacophore Fit Score of over 55 were selected for further docking experiments. The docking to the PRK1 model was done using GlideSP. Seven of the top-scored molecules comprising several novel scaffolds were purchased for further biological validation; two of them (compounds CB-5743914 and CB-6046000, see Figure 3.8) showed an IC_{50} in the low micromolar range (see Table 3.3).



PD-0166285 CB-5743914 CB-6046000

Figure 3.8: PD-0166285 and inhibitors identified by pharmacophore-based virtual screening.

3.5 Post-Processing of Docking Results Using BFE Calculations

The virtual screenings described in the previous sections have shown that the methods used here were able to identify compounds which inhibit PRK1, however, they are still not accurate enough to discriminate highly active PRK1 inhibitors from moderately or low active ones. The results clearly demonstrate that the choice of compounds for testing based simply on docking score and visual inspection may be not sufficient criteria to find highly potent PRK1 inhibitors such as the staurosporine derivatives. Thus, there is still a need of finding reasonable methodology, which should be able to distinguish the biological activities of compounds.

With the aim to prioritize potential activities of compounds, an approach was developed, which combines a structure-based virtual screening together with rescoring methodology using binding free energy calculations. It can be applied as an efficient tool for lead structure identification and prioritizing compounds for further biological validation, and it is described in the following sections.

3.5.1 Studied Dataset

As of 2010, when we started the project, there were a few PRK1 inhibitors available in public databases such as ChEMBL (38 compounds, see Appendix Table A.1). Only one inhibitor (CHEMBL38380 = fasudil, HA-1077)[164] had the half maximal inhibitory concentration (IC_{50}, nM) reported, other compounds had determined dissociation constant (K_d, nM, 11 inhibitors including staurosporine analogs lestaurtinib, midostaurin, ruboxistaurin),[165] residual activity (RA, %, 3 compounds, quinazolines published by Tasler et al.)[166] or inhibition values (%, 23 compounds).[167–170]

Since our study includes the prediction of relative binding free energies, the ideal dataset for this purpose should contain homogenous data, where compounds were tested under the same assay conditions and have measured activities of the same type. Thus, we decided to use as a training set our in-house library of compounds, which were tested on PRK1 in a LanthaScreen^{TM} competitive binding assay and have measured IC_{50} values (see Section 2.6 for assay conditions). It is worth to note, that most of the PRK1 inhibitors registered in ChEMBL (2010) with reported IC_{50} or K_d values, were included in our datasets (test or training set, see the descriptions below).

Table 3.3: PRK1 inhibitors from DS1 along with experimental activity and standard error of mean (SEM). IC_{50} values were determined in a PRK1 *in vitro* assay as described in the Methods Section.

Name	**Supplier**	**IC_{50}, nM**	**±SEM, nM**	**pIC_{50}**	**Screen** [a]
Staurosporine	Biomol	0.8	0.2	9.10	1
K252a	Biomol	3.2	0.6	8.49	1
Lestaurtinib	Biomol	8.6	0.9	8.07	1
PKC-412	Biomol	14.2	1.5	7.85	1
Ro-318220	Biomol	78.3	8.6	7.11	1
CP-690550	Sigma-Aldrich	129	29	6.92	3
BIM IV	Sigma-Aldrich	154.2	49	6.81	3
BIM I	Biomol	579	140	6.24	2
H-7	Biomol	658.5	456	6.18	2
PZ0151	Sigma-Aldrich	1060	140	5.97	5
Calbiochem681640	Calbiochem	1441	220	5.84	3
HA-1077	Biomol	1945	465	5.71	2
CB-5743914	ChemBridge	2940	420	5.53	6
CB-6046000	ChemBridge	5350	1880	5.27	6
PD-0166285	Tocris	5517	2471	5.26	3
Z1039084878	Enamine	6210	3180	5.21	5
CB-76536004	ChemBridge	13090	2660	4.88	5
Z991906728	Enamine	26400	2770	4.58	5
Ambnee93542761	Ambinter	29330	4110	4.53	4
Z1139202903	Enamine	31470	4590	4.50	5
CB-85384930	ChemBridge	34020	11820	4.47	5
CB-38374289	ChemBridge	42560	11650	4.37	5
Z1129905037	Enamine	53380	5970	4.27	5
F2458-0011	Ambinter	53730	9590	4.27	4
Z1139203558	Enamine	55730	11520	4.25	5
F2457-0067	Ambinter	70300	7340	4.15	4
Quercetin	Biomol	Inhib.[b] $> 5\ \mu M$			3
K252c	Sigma-Aldrich	Inhib.[b] $> 1\mu M$			3

[a] (1) compounds identified by *in vitro* screening of Biomol kinase and phosphatase inhibitor screening set at 100 nM threshold concentration; (2) compounds from Biomol screening, for which full IC_{50} measurement was carried out; (3) other known kinase inhibitors tested *in vitro* on PRK1; (4) inhibitors identified by similarity search using isoquinoline derivative H-7 (5) inhibitors identified by similarity search using CP-690550; (6) compounds identified by pharmacophore-based virtual screening using PD-0166285 as query. Compounds are named by trade name or by supplier name with code (most of compounds were purchased from commercial suppliers ChemBridge, Life Chemicals, Amber Chemical, Enamine).

[b] Compounds showed inhibition of PRK1, however, full IC_{50} measurement was not performed.

The compounds for training set were collected from a number of *in vitro* screenings carried out in our laboratory (Data Set 1, DS1). As it was described in Section 3.1, initially, the Biomol kinase and phosphatase inhibitor screening set (84 compounds), containing some generic kinase inhibitors, was tested at 100 nM threshold concentration.[144] This screen has identified the highly potent PRK1 inhibitors staurosporine, Ro318220, K252a and lestaurtinib as well as some moderately active inhibitors (e.g. HA-1077, H-7 and BIM I, see Figure 3.2a). Later on, further *in vitro* tests identified PRK1 inhibitors showing diverse chemical structures (e.g. PD-0166285 and CP-690550, see Figures 3.2b-c). With the intention to identify more diverse inhibitors and to increase the range of activity within the dataset, several virtual screening campaigns were initiated. Since staurosporine analogs represent a well-studied class of non-selective kinase inhibitors, we decided to focus our attention on compounds with distinct scaffolds and binding modes. Therefore we took the isoquinolines HA-1077 and H-7, the pyrrolopyrimidine CP-690550 and the pyridopyrimidinone PD-0166285 and carried out similarity and pharmacophore-based virtual screenings as it is described in Section 3.4. The identified hits were docked into the PRK1 homology models using GlideSP. The top-ranked binding poses were visually analyzed and some compounds were purchased and evaluated for PRK1 inhibition. Several of them showed an inhibition in the micromolar range (e.g. isoquinolines Ambnee93542761, F2458-0011, F2457-0067; pharmacophore hits CB-5743914 and CB-6046000; pyrrolopyrimidiness PZ0151, Z1039084878 and others; see Table 3.3).

The virtual screenings and *in vitro* tests, applied on different stages of the project, expanded our dataset to 328 compounds (28 actives and 300 inactives). Among them 26 compounds were shown to inhibit PRK1 with IC_{50} values in the range of 0.8 nM - 70 μM (see Table 3.3). Additionally, two compounds showed inhibition of PRK1 at 5 μM (quercetin) and 1 μM (K252c). PRK1 inhibitors representing diverse hinge-binding scaffolds, which compose the training set DS1, are shown in Figure 3.9. For the structures of all 28 PRK1 inhibitors see Figures 3.2, 3.6, 3.7 and 3.8.

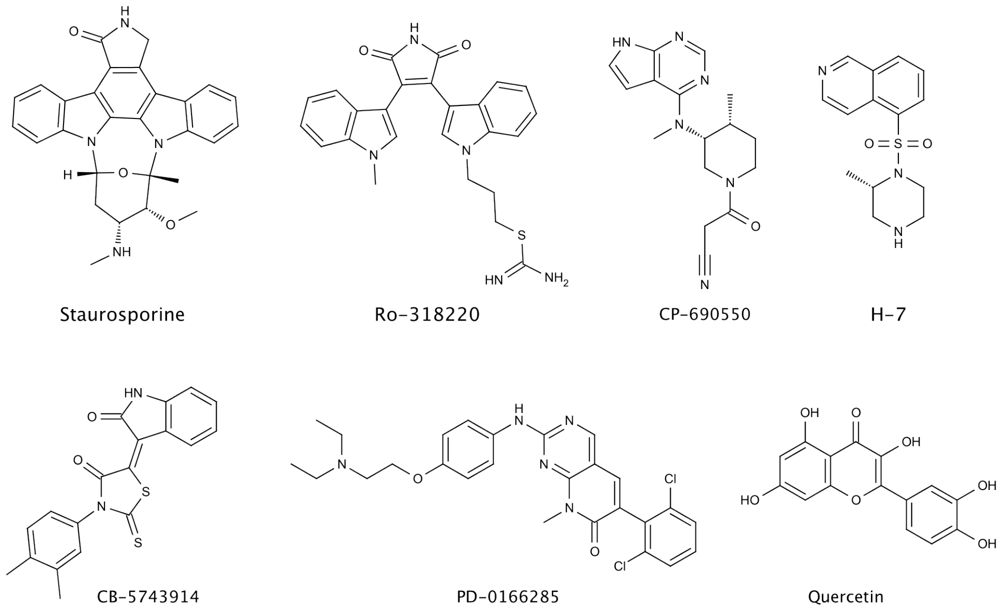


Figure 3.9: Representative PRK1 inhibitors from DS1, showing different hinge-binding scaffolds.

3.5.2 Docking Into a Single Protein Conformation

In the docking studies we first investigated the performance of each homology model separately. Because there was no experimentally determined structure for PRK1 available, the ability to reproduce the correct docking pose of known inhibitors was evaluated by comparing with binding modes taken from crystal structures of homologous kinases with the same or a closely related inhibitor. The results show that a proper pose can be identified by Glide for a given inhibitor when the correct protein conformation is used. However, the particular model can be limited to the usage for compounds similar to the one used in induced-fit modeling. For example, the side chain of Phe904 within the C-terminal flexible loop (C-tail) is exposed into the ATP-binding pocket in homology model refined for HA-1077 (hm_2esm_l, see the Section 3.2.2 which describes the homology model refinement and Figure 3.4 for details). The phenylalanine makes typical "edge-to-face" aromatic interaction with the ligands HA-1077 and H-1152P in crystal structures with PKA or Rho kinase (PDB structures 1Q8U, 1Q8W, 2ESM, 2GNI) as well as in the homology model of PRK1 with HA-1077. Such a conformation of Phe904 plays an important role in the identification of the correct pose and scoring of compounds similar to HA-1077 during the docking. At the same time, the size of the pocket of hm_2esm_l is too small to accommodate large ligands like staurosporine and its analogs. Conversely, the homology model refined with staurosporine (hm_STU) has a larger ATP-binding pocket with Phe904 pointing outside of it, thus, the flipped HA-1077 conformation is usually observed as a result of docking to hm_STU (see Figure 3.4).

The observations clearly show that using only one protein conformation can result in

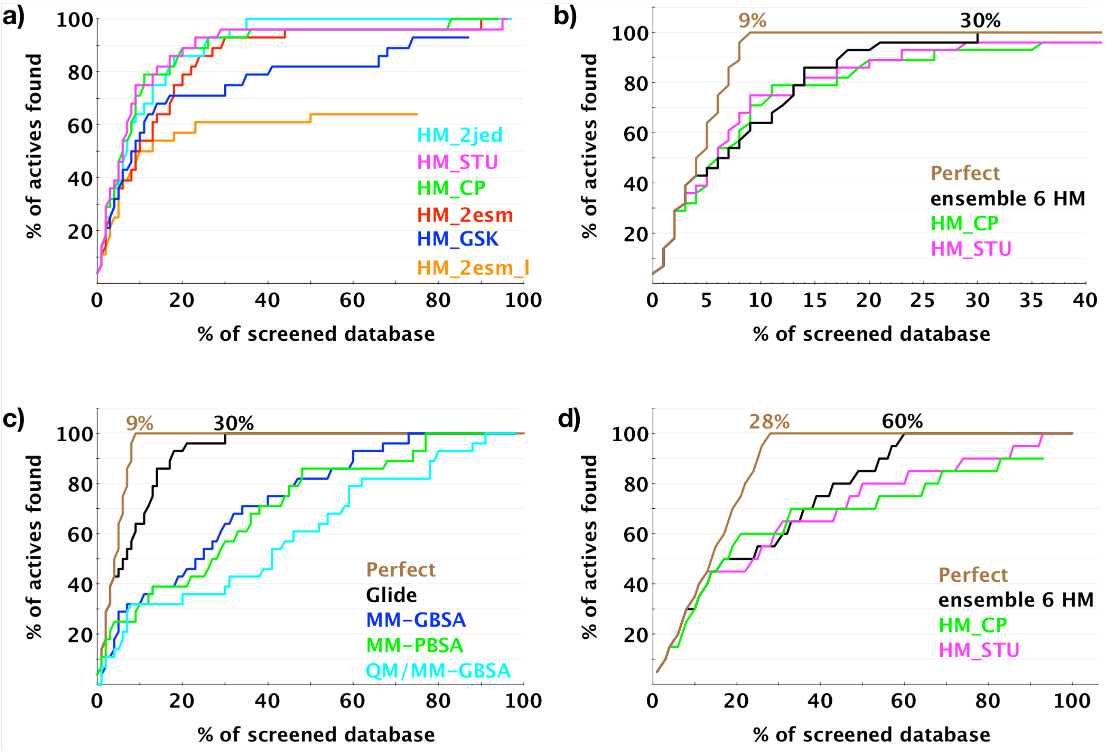


Figure 3.10: Enrichment plot showing the percentage of actives found at a given percentage of the ranked database for DS1 (**a-c**) and DS2 (**d**). The plot compares the performance of: **a)** docking of DS1 into each single homology model; **b)** ensemble docking of DS1 using all six homology models versus the two best single models; **c)** ensemble docking of DS1 using different scoring methods – docking score (GlideSP) or BFE scores (MM-PBSA, MM-GBSA, QM/MM-GBSA); **d)** ensemble docking of DS2 versus the two best single models. The theoretical perfect curve, when all known PRK1 binders are ranked among the top of database, is shown in brown. In some cases the percentage of active compounds or screened database is not reaching 100%, implicating that Glide did not return poses for some compounds.

an inaccurate prediction of ligand binding poses and, subsequently, in low enrichment factors, especially when a diverse set of ligands is docked (see Figure 3.10a). Furthermore, the incorrectness on the stage of docking can mislead further rescoring experiments. It is worth to note that although some models yield lower enrichment than others, they still can be used in VS or ensemble docking since they can represent a more unique protein conformation that is able to recognize certain types of compounds.

3.5.3 Docking Into an Ensemble of PRK1 Conformations

There are several reasons why docking programs fail to identify the correct ligand pose. Besides the limitations of common scoring functions, the reason may lie in using a wrong protein conformation. The docking methods can fail to recognize an active ligand when some residues in the binding pocket occupy the ligand binding space. To overcome this problem we docked all compounds tested on PRK1 from DS1 in all six PRK1 homology models, see Section 3.2.2 for the details of homology modeling. An ensemble of protein conformations provides a degree of receptor flexibility, which is often not considered in docking experiments. The optimal docking mode was then selected according to the best GlideSP score. The analysis of ensemble docking results shows an overall improvement in the ranking of active ligands, which are distributed among 30% of the top-ranked compounds in the database (see Figure 3.10b). In comparison, the best single homology model (hm_2jed) has actives distributed among the top 35% of the database. Even though in this case the difference is not dramatic, it is important that using multiple models leads to better pose prediction for structurally diverse groups of inhibitors. However, if the best performing model for a certain compound class is known, it is computationally favorable to use single receptor docking.

Furthermore, the docking results were evaluated by plotting receiver operating characteristic (ROC) curves, which show the distribution of true positives versus false positives. The corresponding area under the curve (AUC), which describes the quality of enrichment, was calculated. These data are summarized in Figure 3.11, illustrating that the ensemble docking approach (AUC=0.96) outperforms single protein models (AUC=0.45 ∼ 0.94).

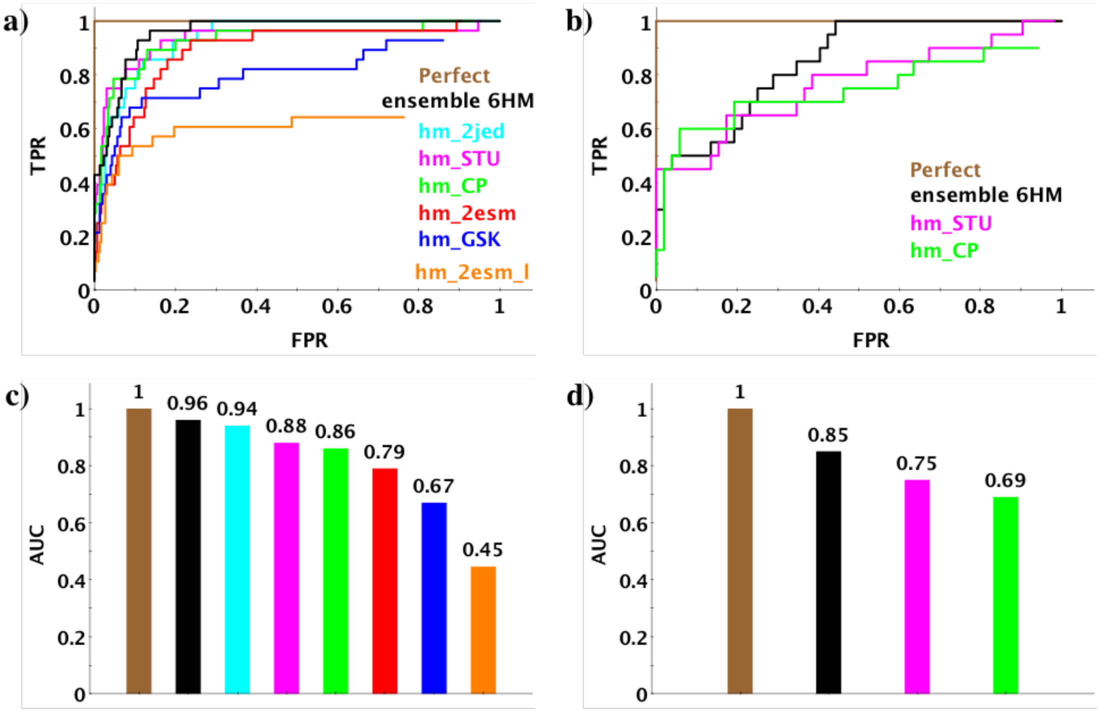


Figure 3.11: Receiver operating characteristic (ROC) curves showing the fraction of true positives (TPR = true positive rate) versus the fraction of false positives (FPR = false positive rate). The plot compares the performance of: **a)** docking of training set DS1 into each single homology model and ensemble of six homology models; **b)** ensemble docking of DS2 versus two best performing single models. The theoretical perfect curve, when all known PRK1 binders are ranked among the top of database, is shown in brown. In some cases TPR or FPR is not reaching maximum value of 1, implicating that Glide did not return poses for some compounds. Correspondingly, AUC values for **c)** DS1 and **d)** DS2 are confirming that ensemble docking outperforms docking into individual protein models (different models are colored as it is shown on ROC curves).

3.5.4 BFE Calculation for Rescoring of Docking Results

The results of previous VS experiments and biological tests showed that relying only on the docking score might be insufficient for correct ligand scoring according to the biological activity. Even though some of the hits derived from ligand- and pharmacophore-based screening were active on PRK1, they appeared to be weaker PRK1 inhibitors compared to the corresponding query compounds. The ability to estimate the activity of a compound is one of the critical VS tasks, which aims to reduce the time and costs needed for inhibitor development. Thus, we decided to test the performance of some BFE assessment methods on known PRK1 inhibitors.

DS1, containing 26 active PRK1 inhibitors with available IC_{50} values, was used for the calculation of binding free energies using a fast post-processing approach. The top Glide docking solution for each compound in the corresponding homology model was selected and submitted for the energy minimization using an implicit solvent model. After this, the binding free energy was evaluated using the MM-PB(GB)SA or QM/MM-GBSA protocol for one snapshot derived after the minimization. In order to investigate the performance of different scoring methods on enrichment and correlation with biological activity, the results for different homology models were combined and the best pose of each ligand was selected according to Glide score.

3.5.4.1 Accuracy of Docking/Scoring Methods in Enrichment Performance

The virtual screening performance was evaluated by the ability to enrich known actives among the top ranked compounds in the database comprising experimentally validated active and inactive PRK1 inhibitors. The enrichment plot shown in Figure 3.10c indicates that rescoring of poses with MM-PB(GB)SA or QM/MM-GBSA performed worse in the discrimination of active from inactive compounds in comparison to the Glide score. Interestingly, several studies report a better discrimination of known actives from decoys while using MM-PB(GB)SA rescoring methodologies as compared to the AutoDock score.[74, 171] However, the screening performance is usually target- and dataset-dependent, which probably can explain success in some cases and failure in others. Nevertheless, the work by Graves et al.[172] shows contradicting results - rescoring introduces many false positives, especially among the top ranked ligands, as compared to simpler docking protocols. The reason may lie in larger errors when internal energies of the receptor structure are introduced. Also wrong ligand parameterization or electrostatics treatment might be responsible for larger errors. In addition, minimization of the complex allows the adjustment of the binding site for bigger ligands, including false positives.

3.5.4.2 Estimating the Biological Activity of PRK1 Inhibitors

We have analyzed the performance of different approaches for their ability to correctly estimate the experimental inhibition data. The results for some single homology models show low correlation with experimental pIC_{50} values (see Table 3.4), demonstrating that the protein conformation plays an important role in the ligand pose prediction and BFE calculation. Even though, some improvement of correlation was observed for the models with good enrichment (e.g. hm_2jed and hm_CP), when the rescoring with MM-PBSA was done ($r^2 = 0.47$, $RMSE = 1.03$, $q^2_{LOO} = 0.37$ and $r^2 = 0.51$, $RMSE = 0.99$, $q^2_{LOO} = 0.43$, for hm_2jed and hm_CP, respectively).

Table 3.4: Summary of r^2 correlation coefficients for different scoring methods using single homology models or ensemble of six homology models.

Correlation coefficient ($\mathbf{r^2}$)				
Model name	**glide_SP**	**MM-GBSA**	**MM-PBSA**	**QM/MM-GBSA**
hm_2esm_l	0.51[a]	0.15	0.02	0.17
hm_2esm	0.18	0.45	0.63	0.59
hm_2jed	0.05	0.29	0.47	0.44
hm_CP	0.32	0.25	0.51	0.40
hm_GSK	0.14	0.39	0.37	0.57
hm_STU	0.30	0.32	0.31	0.54
ensemble	0.42	0.42	0.61	0.53

[a] Note: Glide did not return poses for some of compounds

Interesting results have been observed for the rescoring of ensemble docking solutions (see Figure 3.12). The best correlation coefficient ($r^2 = 0.61$, $RMSE = 0.89$, $q^2_{LOO} = 0.56$) was obtained when the MM-PBSA method was applied for post-processing (see Figure 3.12c).

Furthermore, we investigated the effect of other descriptors like partial charges or

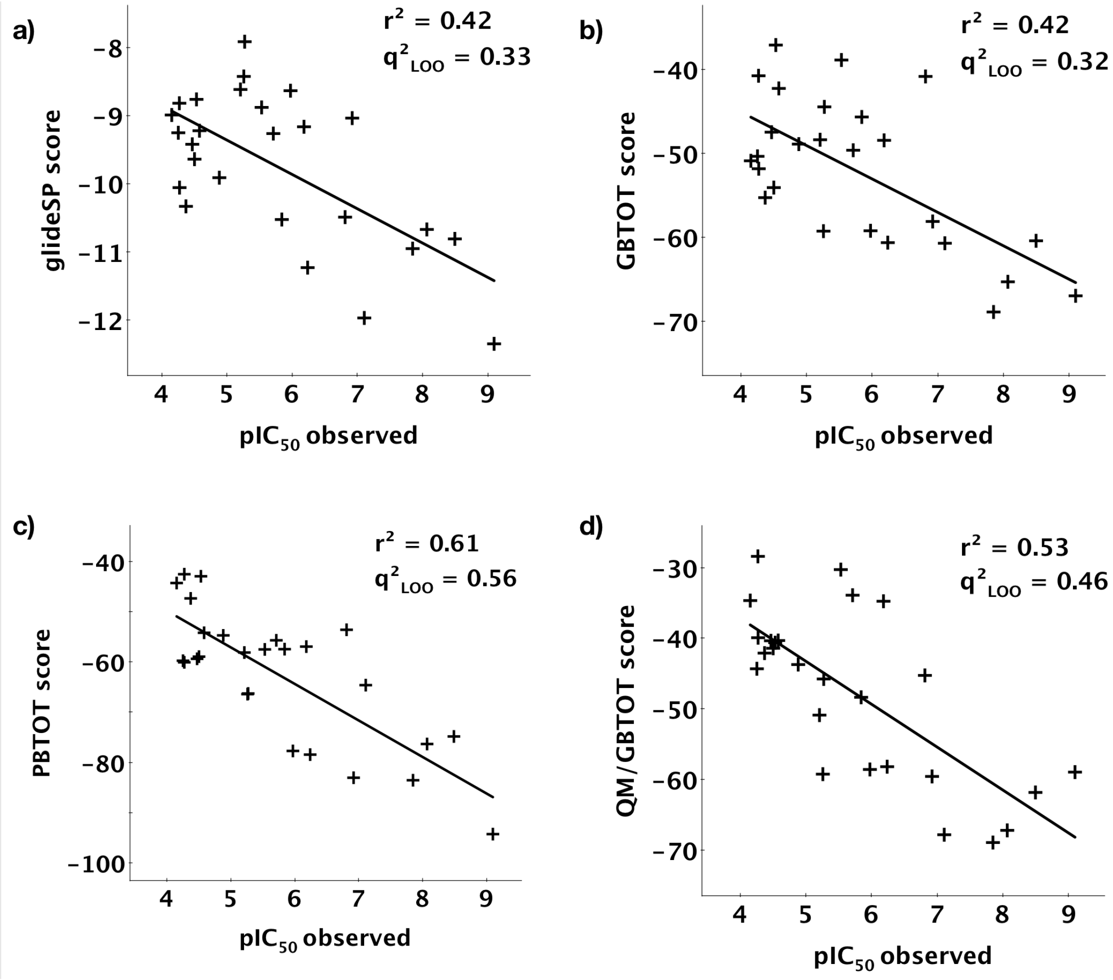


Figure 3.12: Regression and corresponding correlation coefficients r^2 are shown for the ensemble docking results. The plots represent the correlation between observed pIC$_{50}$s for 26 compounds of DS1 and different scores: **a)** docking score glide_SP; binding free energy scores calculated by **b)** MM-GBSA method (GBTOT score); **c)** MM-PBSA method (PBTOTscore); **d)** QM/MM-GBSA method (QM/GBTOT score).

number of rotatable bonds on the predicted BFE scores. In order to include the partial charges we constructed a simple predictive quantitative structure activity relationship (QSAR) model based on two descriptors, namely MM-PBSA score "PBTOT" and partial charges "FCharge" (QSAR_model_1), using MOE[99] PLS methodology, which resulted in correlation coefficient $r^2 = 0.66$ with root mean square error $RMSE = 0.82$ and cross-validated correlation coefficient $q^2_{LOO} = 0.57$.[173, 174] Next, we tested the number of rotatable bonds “brotN” on the QSAR model as descriptor containing information on molecular flexibility and compound’s conformational space. The most interesting results were found when this descriptor was combined with QM/MM-GBSA score “QM/GBTOT” (QSAR_model_2), which led to an improvement of the correlation to $r^2 = 0.68$, $RMSE = 0.80$ and $q^2_{LOO} = 0.61$. Other tested descriptors, such as logP, solvent accessible surface area or number of heavy atoms, did not improve the quality of the QSAR model (results are not shown).

Furthermore, different combinations of descriptors were tested on the model’s predictive power. Finally, the QSAR_model_3 containing four descriptors (MM-PBSA score “PBTOT”, QM/MM-GBSA score “QM/GBTOT”, number of rotatable bonds “brotN” and Glide score “glideSP_score”) was developed, which resulted in a correlation $r^2 = 0.78$, $RMSE = 0.66$, $q^2_{LOO} = 0.70$ between the predicted and experimental pIC_{50} (Figure 3.13). The following equation describes the estimated linear model (QSAR_model_3):

$$\begin{aligned} pIC_{50} = &-1.77015 - 0.36664 \times glideSP\_score - 0.03968 \times PBTOT \\ &- 0.04309 \times QM/GBTOT - 0.17230 \times b\_rotN, \end{aligned} \tag{3.1}$$

where the relative importance of descriptors is 0.776953 (glideSP_score), 1.000000 (PBTOT), 0.975233 (QM/GBTOT) and 0.706790 (b_rotN).

Interestingly, the most potent inhibitors (IC_{50} <100 nM) have a calculated pIC_{50} higher or close to 7, which is in correspondence with the experimental values for these compounds (see Table 3.5). Thus, the value of $pIC_{50}>7$ could be used as a threshold for the predicted pIC_{50} in order to prioritize compounds for further synthesis or purchase. Such a strategy of rescoring the docking solutions and calculating a predicted pIC_{50} using the developed QSAR model can be especially suitable on the stage of compound optimization, where highly potent inhibitors are of interest.

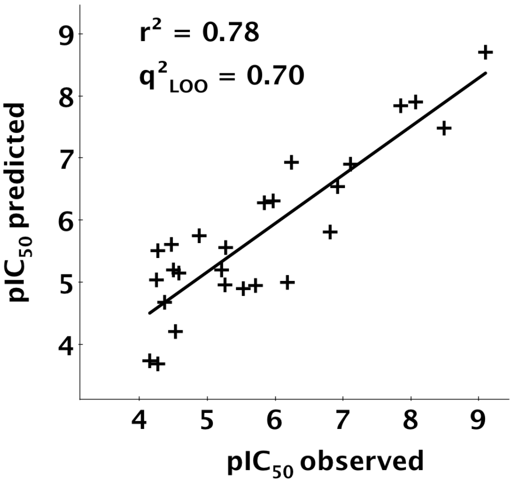


Figure 3.13: Regression and corresponding correlation coefficients r^2 are shown between the observed pIC_{50} versus predicted pIC_{50} by QSAR_model_3 using four descriptors (glideSP-PBTOT-QM/GBTOT-brotN).

Table 3.5: Comparison of the measured pIC_{50} versus the predicted values by the QSAR_-model_3 for the most potent PRK1 inhibitors from DS1.

Name	**IC_{50}, nM**	**pIC_{50}**	**Predicted pIC_{50} (QSAR_model_3)**
Staurosporine	0.8	9.10	8.70
K252a	3.2	8.49	7.48
Lestaurtinib	8.6	8.07	7.90
PKC-412	14.2	7.85	7.84
Ro-318220	78.3	7.11	6.90
CP-690550	120	6.92	6.54

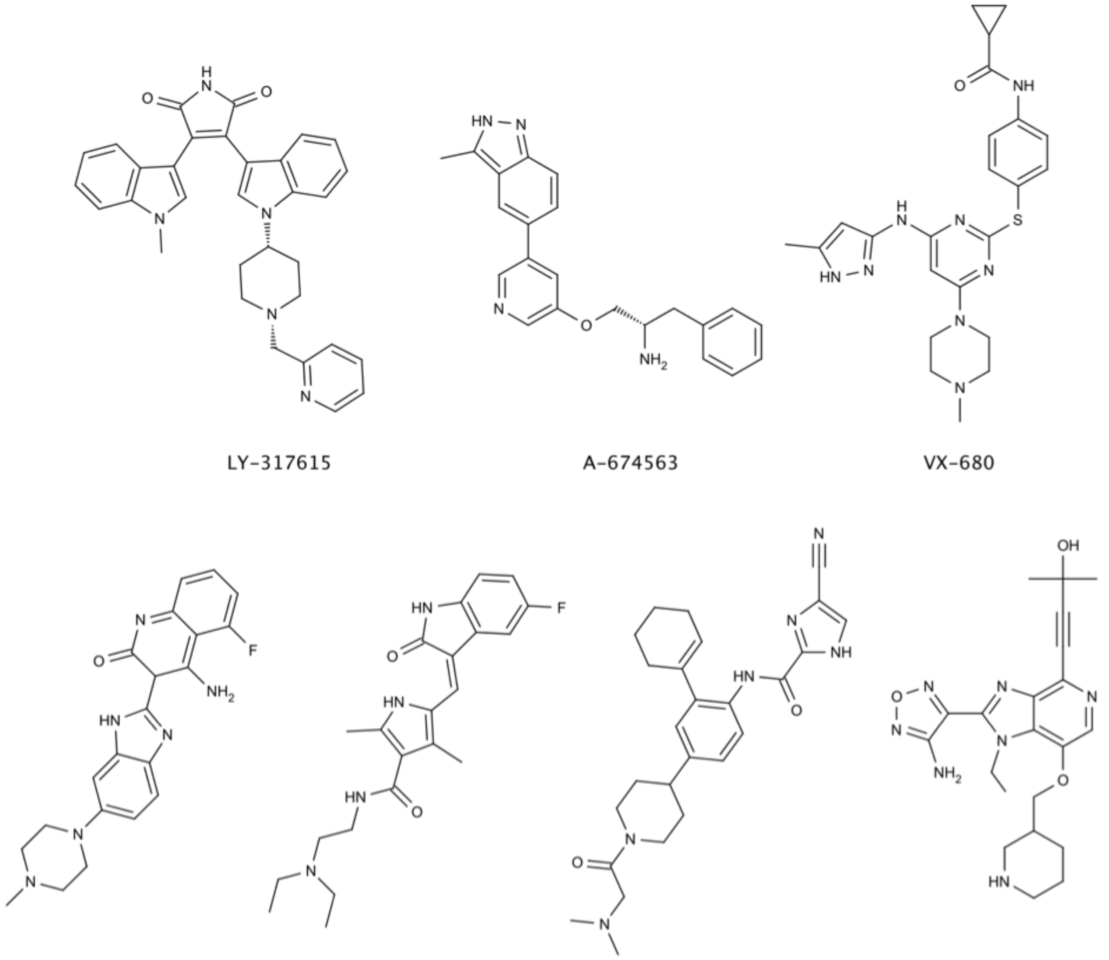
3.5.5 Validation of the Results

The performance of our protocol for virtual screening and rescoring using MM-PB(GB)SA, QM/MM-GBSA methodologies was further evaluated on an external test dataset (DS2) containing 20 diverse PRK1 inhibitors (representative inhibitor structures can be found in Figure 3.14) and 52 inactive compounds.[19] Twenty compounds from DS2 were shown to inhibit PRK1 in the range between 1.3 nM and 5 μM (see Table 3.6). It is worth to note that inhibitors from the second dataset are more chemically diverse and have less significant variation in their inhibitory potency in comparison to DS1, which poses a number of challenges in docking and rescoring experiments. The structures of all 20 PRK1 inhibitors from DS2 can be found in Figure A.2, see Appendix.

Table 3.6: Biological activity of 20 PRK1 inhibitors published by Davis et al.[19]

Name	K_d, nM	pK_d
Staurosporine	1.3	8.89
CEP-701	5.3	8.28
PKC-412	9.3	8.03
GSK-690693	34	7.47
CP-690550	170	6.77
CHIR-258	180	6.74
A-674563	210	6.68
TAE-684	340	6.47
LY-333531	350	6.46
KW-2449	580	6.24
R406	660	6.18
Sunitinib	710	6.15
BIBF-1120_deriv	740	6.13
TG-101348	1100	5.96
VX-680	1500	5.82
SU-14813	1900	5.72
JNJ-28312141	2400	5.62
SKI-606	2900	5.54
Flavopiridol	3500	5.46
LY-317615	5100	5.29

Even when the inhibitor data of Davis et al. were not obtained in the same assay we used, it can be assumed that the values from both assays are comparable. This



LY-317615 A-674563 VX-680

CHIR-258 Sunitinib JNJ-28312141 GSK-690693

Figure 3.14: The representative subset of PRK1 inhibitors from DS2.

is supported by very similar values obtained for the standard inhibitors staurosporine ($IC_{50} = 0.8 \pm 0.2$ nM, $K_d = 1.3$ nM) and PKC-412 ($IC_{50} = 14.2 \pm 2.5$ nM, $K_d = 9.3$ nM). Correspondingly, all compounds of DS2 were docked to all six PRK1 homology models and then the top Glide solutions were rescored using binding free energy calculations after complex minimization. The procedure was done as described in Methods Section 2.5.2.

Similar to the docking of the first dataset, using an ensemble of protein conformations improves the enrichment compared to rigid receptor docking (see Figure 3.10d) and performs better in identifying a ligand conformation similar to the bioactive one (as compared with known crystal structures). The area under the ROC curve (AUC) for ensemble docking results is equal to 0.85, which outperforms the best single homology models hm_STU and hm_CP (AUC=0.75 and 0.69, correspondingly, see Figure 3.11).

The correlation between the experimental pIC_{50} and those predicted by the developed QSAR_model_3 was only moderate ($r^2 = 0.45$, $RMSE = 0.71$, $q^2_{LOO} = 0.29$, $r^2_0 = 0.40$), nevertheless, some trends could be observed (see Figure 3.15a). Therefore, we analyzed the model for outliers, which were detected by large Z-scores. The Z-score represents the absolute difference between the predicted value of the model and the observed biological activity divided by the square root of the mean square error of the dataset. The Z-scores were calculated using the QSAR module of MOE. A compound was considered an outlier if the Z-score was higher than 2.5. Only one outlier was identified, namely compound LY-317615. The comparison of structures of PRK1 with the inhibitor LY-317615 before and after the minimization shows that the refining of the complex led to conformational changes within the protein-inhibitor complex, which resulted in a strong interaction pattern between ligand and receptor (Figure 3.16). This could in part explain the highly favorable scores for this compound. However, LY-317615 inhibits PRK1 with $K_d = 5100$ nM and it represents the least potent inhibitor in DS2, suggesting that the binding mode of this compound could be incorrectly predicted or other factors could play a role during the ligand binding, e.g. entropy changes or presence of water in the binding pocket.

Removal of compound LY-317615 improved the results - the correlation coefficient r^2 became equal to 0.68 and the cross-validation correlation coefficient q^2_{LOO} to 0.59 (see Figure 3.15b), which is slightly lower than for the validation set. The removal of three duplicates from DS2 (staurosporine, PKC-412 and CP-690550) resulted in correlation $r^2 = 0.49$, $RMSE = 0.51$ and $q^2_{LOO} = 0.26$. It is worth to note, that DS2 has a more narrow range of activities (1.3 nM - 5 μM) in comparison to DS1 (0.8 nM - 70 μM), which may be one possible reason for the weaker statistical data. As the results indicate, the MM-PB(GB)SA, QM/MM-GBSA rescoring approach is not accurate enough to

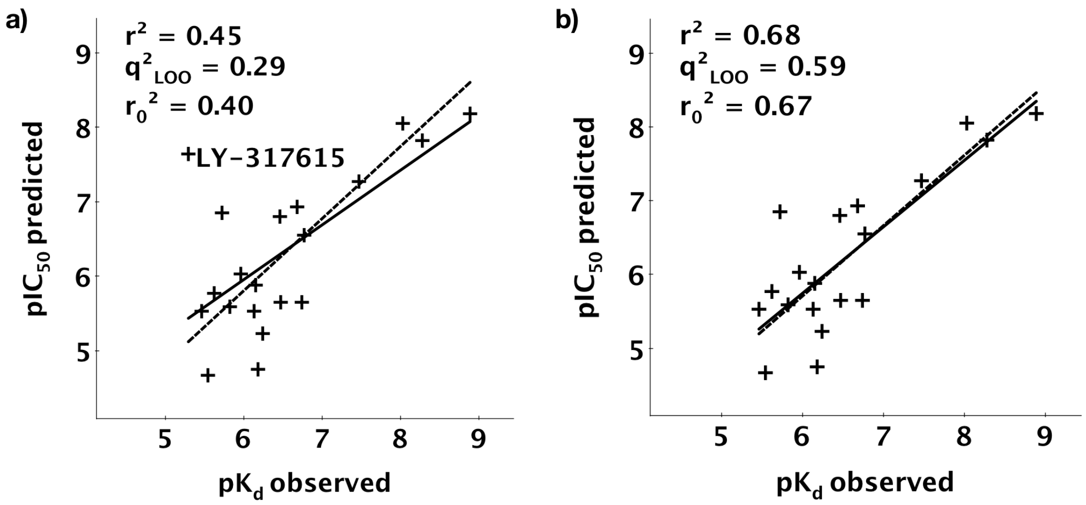


Figure 3.15: Regression between the observed pK_d of PRK1 inhibitors from the test set DS2 versus predicted pIC_{50} by QSAR_model_3: **a)** using the data of all 20 active compounds; **b)** after removing the outlier (compound LY-317615). The optimal regression line of the model is shown as solid black line and the regression through the origin of the coordinate system is shown as dashed black line.

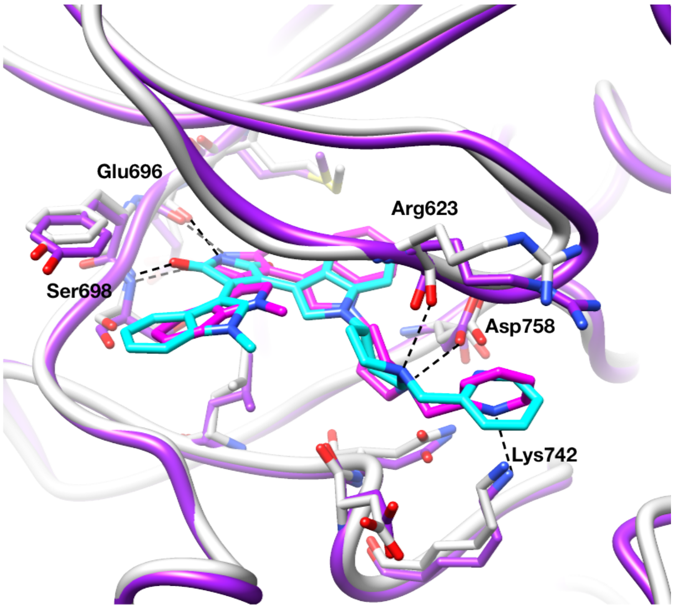


Figure 3.16: Conformations of compound LY-317615 before (light grey protein and cyan ligand) and after (purple protein and magenta ligand) minimization of the complex along with interacting residues of the binding site are shown. Hydrogen bonds between the inhibitor and the kinase are displayed as dashed lines.

distinguish the compounds with similar binding affinities, but it can be used to separate strong PRK1 inhibitors from moderate or weak binders.

Since the data of DS1 and DS2 were obtained in different assays and reported K_d values of DS2 can differ from IC_{50} values of DS1, we additionally generated a QSAR model for DS2. Correspondingly, the same methodology and descriptors were used as for DS1 (QSAR_model_3, MM-PBSA and QM/MM-GBSA scores, number of rotatable bonds and Glide score). The resulting QSAR model showed slightly different coefficients for the individual descriptors. This resulted in a correlation coefficient of $r^2 = 0.53$, $RMSE = 0.66$, and $q^2_{LOO} = 0.12$ for the 20 compounds of DS2 with reported activities. After the removal of one outlier (LY-317615, same outlier as observed in the prediction) the correlation was improved to $r^2 = 0.73$, $RMSE = 0.49$ and $q^2_{LOO} = 0.53$ (data not shown).

Furthermore, the predictive ability of the QSAR_model_3 was tested using Golbraikh-Tropsha criteria.[175–177]

They consider that a model with a good predictive power should satisfy the following conditions:

1) squared correlation coefficient between predicted and observed activities $r^2 > 0.6$;
2) squared cross-validation correlation coefficient $q^2 > 0.5$;
3) one of coefficients of determination for regressions through the origin (either predicted vs. observed activities r_0^2 or observed vs. predicted activities $r_0^{'2}$) should have value close to r^2;
4) $(r^2 - r_0^2)/r^2$ or $(r^2 - r_0^{'2})/r^2 < 0.1$ and $0.85 < k$ or $k' > 1.15$, where k and k' are slopes of the regression line through the origin.

The calculated parameters for the test dataset and the satisfactory conditions are summarized in Table 3.7, indicating that the derived model has a good predictive power.

Notably, the obtained data show that the most potent inhibitors (K_d<100 nM) from

Table 3.7: Results of the validation of QSAR_model_3 on the test set DS2.[a]

Criteria	Calculated value	Satisfactory value
r^2	0.677	>0.6
q^2_{LOO}	0.586	>0.5
r_0^2	0.675	$(r_0^2 \text{ or } r_0'^2) \approx r^2$
$r_0'^2$	0.574	
k	0.952	$0.85 \leq (k \text{ or } k') \geq 1.15$
k'	1.042	
$\lvert r_0^2 - r_0'^2 \rvert$	0.101	< 0.3
$(r^2 - r_0^2)/r^2$	0.003	one of this criteria < 0.1
$(r^2 - r_0'^2)/r^2$	0.152	

[a] Golbraikh-Tropsha criteria were used.

DS2 have a predicted pIC$_{50}$ more than 7 (see Table 3.8), which is in accordance with the results for DS1. These observations demonstrate that the model can be used for the prediction of biological activity and especially for the design of highly potent inhibitors.

Table 3.8: Comparison of measured data versus predicted pIC$_{50}$ by QSAR_model_3 for the most potent PRK1 inhibitors from DS2.

Name	K$_d$ (PRK1), nM	pK$_d$ (PRK1)	pIC$_{50}$ predicted
Staurosporine	1.30	8.89	8.18
CEP-701	5.30	8.28	7.82
PKC-412	9.30	8.03	8.05
GSK-690693	34.00	7.47	7.27
CP-690550	170.00	6.77	6.55

3.6 Application of the Model on External Datasets

3.6.1 Selleck Kinase Inhibitor Set

In the previously described study the predictive ability of the model was tested in a retrospective way on a set of PRK1 inhibitors with known activities taken from the article of Davis et al.[19] It is a common procedure used to establish and validate the predictive power of QSAR model prior to its use for external predictions. However, true validation is prospective experimental validation. Thus, in the next study the docking-minimization-rescoring procedure was applied on external dataset of 355 commercially available kinase inhibitors from Selleck Chemicals.[178] This library contains structurally-diverse, medicinally active, and cell permeable kinase inhibitors, which target different kinase families including receptor tyrosine kinases (RTKs), PKC, Janus kinase (JAK), cyclin-dependent kinase (CDK), etc. Most of inhibitors are ATP-competitive and some of them have been approved by the FDA.

Thus, following the rescoring procedure, compounds were docked into six PRK1 homology models, poses were minimized using implicit solvent model and BFE calculation was performed. Noteworthy, only few compounds among the top hits have predicted pIC_{50} more than 7 and these are already known PRK1 inhibitors such as staurosporine, GSK-690693, or related analog of BIM - enzastaurin.[19] Thus, we decided to purchase and test compounds distinct from the well explored staurosporine-like structures, namely, we selected 5 compounds with diverse scaffolds - S4 (GSK-690693), S5, S6, S7 (A-674563), S10 (formulas are not shown since the results are unpublished). As we can see from the results of biological screening (Table 3.9), all tested compounds showed binding to PRK1 in submicromolar or nanomolar range (hit rate 100%), with only one exception - weak binder S6, for which IC_{50} could not have been determined.

Interestingly, the measured K_d values for compounds GSK-690693 and A-674563 differ from corresponding IC_{50} values (34 versus 270.1±31.12 nM and 210 versus 149±20.30 nM), indicating that these measurements do not have direct relationship. According to

Cheng and Prusoff [179] K_d of an inhibitor is equal to the IC_{50} only when noncompetitive or uncompetitive kinetics apply. Despite these limitations, it is often assumpted in computational chemistry that K_d and IC_{50} values reflect the same rank order. Here, we could observe that even though the estimated K_ds for compounds GSK-690693 and A-674563 differ by several times from their calculated IC_{50}s, both are of the same submicromolar order. Furthermore, it is known that the IC_{50} value for a compound may vary depending on experimental and incubation conditions under which the value is determined.[180] Thus, for the correct interpretation of the results it is always better to compare IC_{50} values from the same experiment.

In summary, the presence of actual PRK1 inhibitors among the best-ranked compounds as well as experimental results for five tested compounds proved that rescoring methodology was able to identify potent nanomolar PRK1 inhibitors among the top hits, which were ranked superior to glide scoring.

3.6.2 GSK Published Kinase Inhibitor Set

The next library, on which the rescoring methodology was applied was taken from the GlaxoSmithKline (GSK) published kinase inhibitor set (PKIS). It contains 367 ATP-competitive protein kinase inhibitors, which were recently released by GlaxoSmithKline for public screening as a resource to develop probes for the untargeted kinome.[181] This library contains compounds which have been published in scientific literature and are accompanied with target and activity annotation, including data about inactive compounds. The GSK PKIS set is available free of charge in screening quantities (10 μL of a 10 mM solution in dimethyl sulfoxide (DMSO)) under agreement. The structures of compounds were downloaded from ChEMBL website [182, 183] and the rescoring methodology was applied as it was described previously. The top-scored molecules with the highest predicted pIC_{50} values were visually analyzed and 25 compounds ($pIC_{50\_predicted}$=4.84 $\sim$ 7.02) were selected for biological testing (see Table 3.10), eight of them (hit rate 32%) showed binding to PRK1. The formulas of compounds are not shown since the results are unpublished. The IC_{50} values calculated from the concentration-inhibition response curve for seven PRK1 inhibitors are ranged from 40 nM to 5 μM (see Table 3.10) with GSK2 and GSK5 being the most potent of these compounds (68.1±12.7 and 40±8.7 nM, respectively). These results are encouraging and can serve as a starting point for the next step of compound optimization in order to achieve a higher potency and selectivity to PRK1, which should involve close cross-disciplinary collaboration between computational, medicinal and biochemists and should be extended in further studies.

Furthermore, it is interesting to note that rescoring of docked molecules using BFE calculations resulted in ranking of active compounds higher in the hit list compared to glideSP score, e.g ranks of the some of good-scored compounds (glideSP = -10 and less) were lowered (pIC_{50} predicted around 5.5 $\sim$ 6) and otherwise, compounds with moderate docking scores (glideSP $\sim$ -9) got higher ranks in rescoring, see Table 3.10. These results suggest that despite of limitation of both docking and BFE calculation approaches, the combination of these two methodologies can be especially useful to assist in the selection of potentially interesting compounds for biological tests.

Table 3.9: Twelve top-ranked compounds from Selleck database by predicted pIC_{50} by QSAR_model_3.

Name	**Original target**	**K_d, nM**[a]	**pK_d**[a]	**pIC_{50} predicted**	**glideSP score**	**Binding, 1 μM**	**IC_{50}, nM**	**SEM, nM**	**pIC_{50}**
S1(Staurosporine)	PKC	1.30	8.89	8.15	-11.34				
S2(LY317615)	PKC	5100.00	5.29	7.86	-12.10				
S3	PKC			7.30	-11.56				
S4(GSK-690693)	Akt	34.00	7.47	7.27	-9.82	binding	270.10	31.12	6.57
S5	PDGFR			7.09	-10.51	binding	71.28	16.50	7.15
S6	CDK			6.93	-10.83	weak binding	n.d.	n.d.	n.d.
S7(A-674563)	Akt, CDK, PKA	210.00	6.68	6.90	-10.10	binding	149.00	20.30	6.83
S8	Flt, JAK, c-RET			6.71	-8.77				
S9	PLK			6.67	-9.60				
S10	Others			6.64	-10.47	binding	11.05	3.70	7.96
S11	Chk			6.60	-8.97				

[a] K_d and pK_d values for PRK1 are taken from Davis et al.[19]
n.d. = not determined

Table 3.10: Twenty five top-ranked compounds from GSK PKIS database by predicted pIC_{50} by QSAR_model_3, which were tested on PRK1 *in vitro*.[a]

Name	**glideSP score**	**pIC_{50} predicted**	**Binding at 1 μM**	**IC_{50}, nM**	**SEM, nM**	**pIC_{50}**
GSK1	-9.29	7.02	binding	181.30	41.80	6.74
GSK2	-9.07	6.71	binding	68.10	12.70	7.17
GSK3	-9.61	6.65	n.b.			
GSK4	-9.90	6.62	n.b.			
GSK5	-9.27	6.56	binding	40.00	8.70	7.40
GSK6	-8.59	6.47	n.b.			
GSK7	-10.38	6.45	n.b.			
GSK8	-10.89	6.33	binding	n.d.	n.d.	
GSK9	-9.00	6.30	binding	120.00	22.80	6.92
GSK10	-10.05	6.26	n.b.			
GSK11	-9.39	6.23	binding	3600.00	2800.00	5.44
GSK12	-9.37	6.12	binding	4900.00	1070.00	5.31
GSK13	-8.79	6.06	n.b.			
GSK14	-8.92	6.06	binding	710.00	74.00	6.15
GSK15	-10.05	6.03	n.b.			
GSK16	-9.09	5.89	n.b.			
GSK17	-10.15	5.64	n.b.			
GSK18	-8.62	5.61	n.b.			
GSK19	-10.28	5.58	n.b.			
GSK20	-10.39	5.57	n.b.			
GSK21	-10.37	5.52	n.b.			
GSK22	-9.07	5.46	n.b.			
GSK23	-8.62	5.33	n.b.			
GSK24	-6.85	4.88	n.b.			
GSK25	-9.31	4.84	n.b.			

[a] n.b = no binding
n.d. = not determined

3.7 PRK1 Crystal Structures Released

As it was mentioned in Section 3.2 there was no PRK1 crystal structure available when the project started. That is why homology models were used in the current work. However, in August 2014 a paper was published describing crystal structures of PRK1 in complex with clinical compounds lestaurtinib and tofacitinib as well as staurosporine analog Ro-318220 and apo-structure.[184] Later on, the structures were deposited in protein databank under PDB codes 4OTD, 4OTG, 4OTH and 4OTI, see Table 3.11 for more details.

Table 3.11: Crystal structures of PRK1.

PDB Code	Resolution, Å	Ligand
4OTD	2.00	apo-structure
4OTI	1.93	Tofacitinib
4OTG	2.60	Lestaurtinib
4OTH	1.80	Ro-318220

In their study Chamberlain et al.[184] highlight the major role of kinase plasticity, showing that PRK1 is able to adapt to ligands and exhibit various conformational changes with the main motions in the region of G-loop and C-tail. Such PRK1 disordering can be demonstrated by comparison of structures 4OTI, 4OTH, and 4OTG as shown on Figures 3.17a and 3.17b. It is worth noting that authors' observations concerning the C-terminus of the catalytic domain (C-tail), which can adopt different conformations depending on inhibitor-bound state and place phenylalanine in the ATP-binding site or out of it, are in correspondence with our previous findings (see Section 3.2.2).

3.7.1 Comparison of PRK1 Crystal Structures to Homology Models

Furthermore, we compared the crystal structures of PRK1 with our homology models. Figure 3.18 displays their superposition using all C_{alpha} atoms. The calculated RMSD values show that the overall difference between models and crystal structures range from

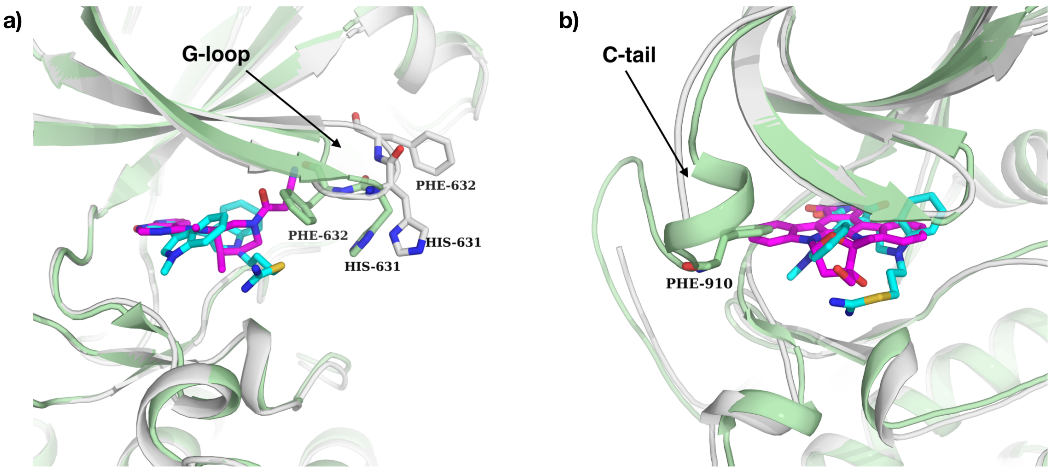
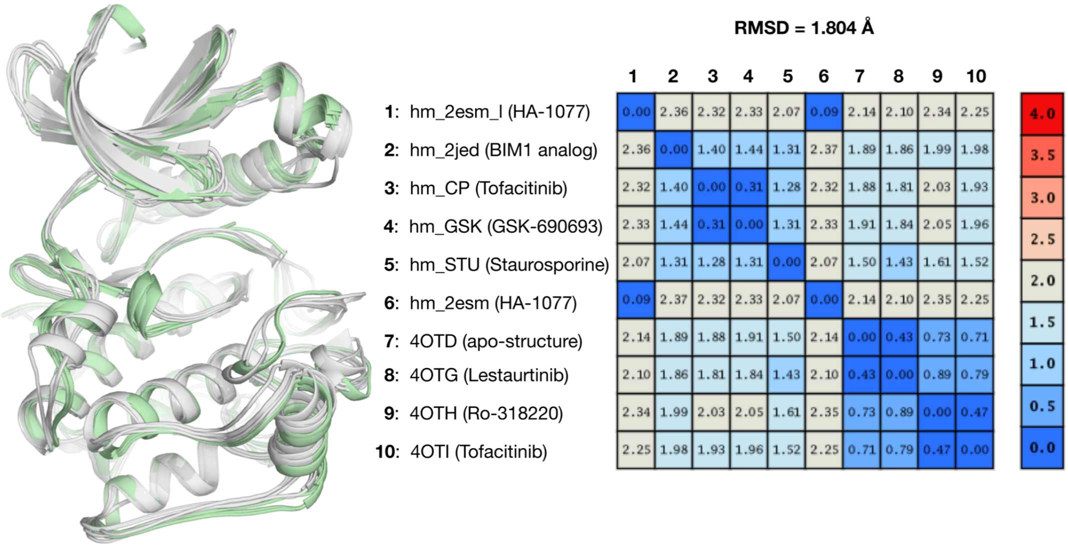


Figure 3.17: PRK1 conformational flexibility on example of G-loop and C-tail motions: **a)** superposition of PRK1 structures 4OTI and 4OTH co-crystallized with tofacitinib (light grey protein and magenta ligand) and Ro-318220 (pale green protein and cyan ligand), showing a large conformational shift of G-loop; **b)** conformational changes of C-tail is shown on superposition of structures 4OTG and 4OTH co-crystallized with lestaurtinib (light grey protein and magenta ligand) and Ro-318220 (pale green protein and cyan ligand). The binding of lestaurtinib to PRK1 caused substantial C-tail disordering, which is no longer visible in the electron density.[184]

1.43 Å (hm_STU with 4OTG) to 2.35 Å (hm_2esm with 4OTH). The most remarkable conformational differences can be observed at loop regions which are known to be flexible, e.g. at the activation loop or glycine-rich loop (G-loop).

Since in structure-based drug design the primary interest is in analyzing the binding site of a protein, it is important that the binding pocket of the homology model is accurately predicted. In order to analyze the differences within the binding pockets, they were superposed using C_{alpha} atoms as it is shown on Figure 3.19, demonstrating RMSD values in a range from 0.57 Å (hm_2jed with 4OTH) to 1.54 Å (hm_CP or hm_GSK with 4OTH). These values are not significantly different as those observed for the homology models or crystal structures themselves, and represent the receptor flexibility, demonstrating that binding site can adapt to different ligands. These findings are in agreement with our previous conclusions concerning protein flexibility, which is currently one of the major pitfalls in the structure-based drug design. Thus, in order

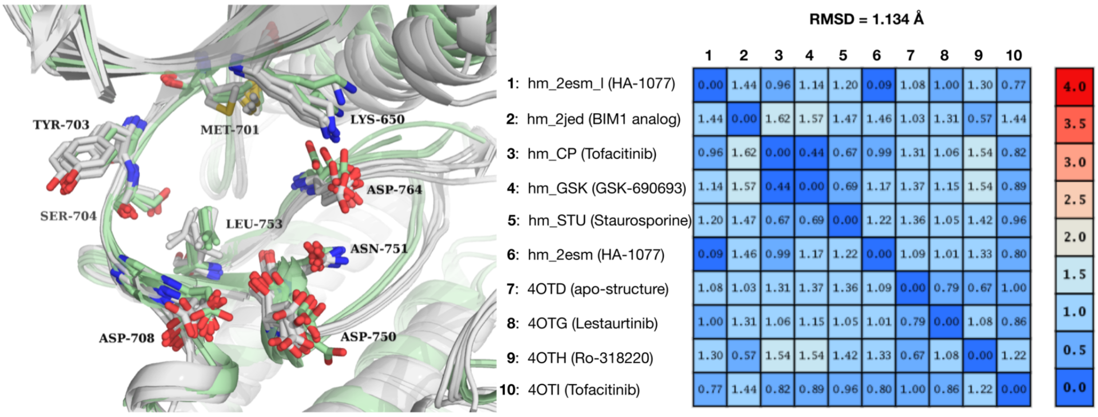


RMSD = 1.804 Å

	1	2	3	4	5	6	7	8	9	10
1: hm_2esm_I (HA-1077)	0.00	2.36	2.32	2.33	2.07	0.09	2.14	2.10	2.34	2.25
2: hm_2jed (BIM1 analog)	2.36	0.00	1.40	1.44	1.31	2.37	1.89	1.86	1.99	1.98
3: hm_CP (Tofacitinib)	2.32	1.40	0.00	0.31	1.28	2.32	1.88	1.81	2.03	1.93
4: hm_GSK (GSK-690693)	2.33	1.44	0.31	0.00	1.31	2.33	1.91	1.84	2.05	1.96
5: hm_STU (Staurosporine)	2.07	1.31	1.28	1.31	0.00	2.07	1.50	1.43	1.61	1.52
6: hm_2esm (HA-1077)	0.09	2.37	2.32	2.33	2.07	0.00	2.14	2.10	2.35	2.25
7: 4OTD (apo-structure)	2.14	1.89	1.88	1.91	1.50	2.14	0.00	0.43	0.73	0.71
8: 4OTG (Lestaurtinib)	2.10	1.86	1.81	1.84	1.43	2.10	0.43	0.00	0.89	0.79
9: 4OTH (Ro-318220)	2.34	1.99	2.03	2.05	1.61	2.35	0.73	0.89	0.00	0.47
10: 4OTI (Tofacitinib)	2.25	1.98	1.93	1.96	1.52	2.25	0.71	0.79	0.47	0.00

Figure 3.18: The superimposition of PRK1 homology models (grey ribbon) and crystal structures (pale green ribbon) by all C_{alpha} atoms.

to account for the induced fit effect in rigid-docking approach, it is rational to use few protein structures for the docking of structurally-diverse ligands, which was the case for PRK1 (see Section 3.5.3).



RMSD = 1.134 Å

	1	2	3	4	5	6	7	8	9	10
1: hm_2esm_I (HA-1077)	0.00	1.44	0.96	1.14	1.20	0.09	1.08	1.00	1.30	0.77
2: hm_2jed (BIM1 analog)	1.44	0.00	1.62	1.57	1.47	1.46	1.03	1.31	0.57	1.44
3: hm_CP (Tofacitinib)	0.96	1.62	0.00	0.44	0.67	0.99	1.31	1.06	1.54	0.82
4: hm_GSK (GSK-690693)	1.14	1.57	0.44	0.00	0.69	1.17	1.37	1.15	1.54	0.89
5: hm_STU (Staurosporine)	1.20	1.47	0.67	0.69	0.00	1.22	1.36	1.05	1.42	0.96
6: hm_2esm (HA-1077)	0.09	1.46	0.99	1.17	1.22	0.00	1.09	1.01	1.33	0.80
7: 4OTD (apo-structure)	1.08	1.03	1.31	1.37	1.36	1.09	0.00	0.79	0.67	1.00
8: 4OTG (Lestaurtinib)	1.00	1.31	1.06	1.15	1.05	1.01	0.79	0.00	1.08	0.86
9: 4OTH (Ro-318220)	1.30	0.57	1.54	1.54	1.42	1.33	0.67	1.08	0.00	1.22
10: 4OTI (Tofacitinib)	0.77	1.44	0.82	0.89	0.96	0.80	1.00	0.86	1.22	0.00

Figure 3.19: The superimposition of PRK1 homology models (grey ribbon and atoms) and crystal structures (pale green ribbon and atoms) by pocket residues using C_{alpha} atoms. The key residues in the binding pocket are shown. G-loop is omitted for clarity.The corresponding RMSD values are shown.

As we could see from the comparison of PRK1 binding pockets, the most accurate homology model is hm_2jed, which can be well superposed to crystal structure of PRK1 with Ro-318220, PDB code 4OTH (RMSD = 0.57 Å for binding pocket C_{alpha} atoms). This is not surprising since the hm_2jed was generated and refined using the highly similar template PKC-theta (PDB code 2JED) complexed with Ro-318220 analog - ligand NVP-XAA228.

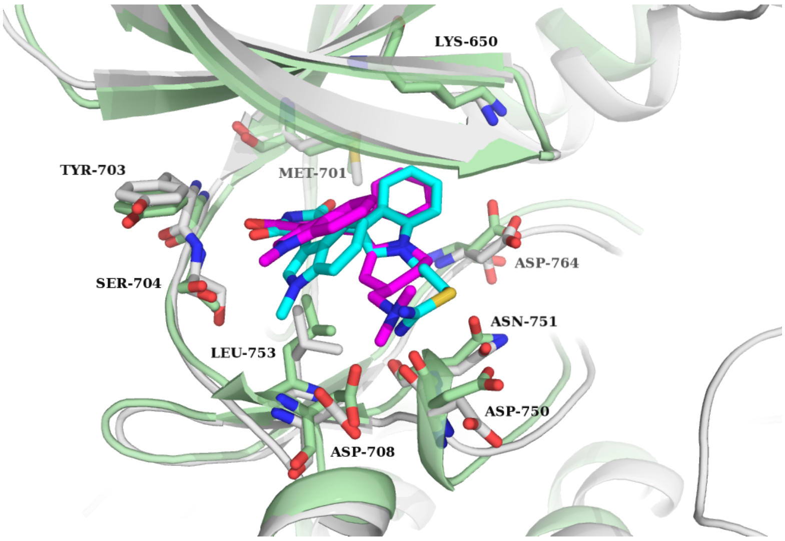


Figure 3.20: The superimposition of PRK1 homology model hm_2jed (grey protein and magenta ligand NVP-XAA228) and crystal structure 4OTH (pale green protein and cyan ligand Ro-318220) by the pocket residues using C_{alpha} atoms. The key residues in the binding pocket are shown. C-tail is omitted for clarity.

3.7.2 Comparison of Docking and Crystallographic Poses of Ro-318220, Lestaurtinib and Tofacitinib

Next, we compared the ligand conformations of Ro-318220, lestaurtinib and tofacitinib (CP-690550) identified crystallographically (PDB structures 4OTH, 4OTG and 4OTI, correspondingly) to the binding modes of these compounds predicted by the docking program Glide.

In general, the structures of PRK1 with above mentioned inhibitors solved crystallographically are consistent with binding modes for these ligands known from the literature.[184] As we can see from Figures 3.21 to 3.23 the top-scored poses from the ensemble-based docking (ligands are depicted in magenta) show overall agreement to the corresponding crystal structures (ligands are depicted in cyan). There is no significant difference in conformations of ligand cores which interact with the hinge region, however, there are some differences in conformations of ligand rotatable bonds (see docking solutions for Ro-318220 and lestaurtinib), which have also impact on the RMSD values, see Table 3.12. Usually, it is more challenging for the docking program to predict the binding positions correctly for a compound with many rotatable bonds, due to the increased number of degrees of freedom to be explored.[185]

Table 3.12: The root mean square deviation (RMSD) between crystallographic pose and docked pose of ligands Ro-318220, lestaurtinib and tofacitinib. The docking pose of each ligand was selected as the best scored ensemble docking solution, the corresponding homology model is indicated. The RMSD value was calculated using all heavy atoms.

Ligand	**RMSD, Å**	**Crystal structure PDB code**	**Homology model**
Ro-318220	2.61	4OTH	hm_2jed
Lestaurtinib	2.04	4OTG	hm_2esm
Tofacitinib	1.16	4OTI	hm_CP

Ro-318220 (bisindolylmaleimide IX), lestaurtinib and tofacitinib inhibit PRK1 with IC_{50}s of 78.3±8.6, 8.6±0.9, 129±29 nM, correspondingly (see Table 3.3). The staurosporine analog, Ro-318220 is a well-known specific PKC inhibitor[186] with no significant PKC isoform selectivity, which also shows potent inhibition against MAPKAP-K1b,[187] MSK1,[188] GSK3β[189] and S6K1.[190] The crystal structure of Ro-318220 in complex with PRK1 as well as docking solution of this compound into PRK1 homology model ($RMSD_{Ro\text{-}318220}$= 2.61 Å, see Table 3.12 and Figure 3.21) is consistent with known binding modes for staurosporine analogs, exhibiting classical hydrogen bond interactions with the hinge backbone (residues Glu702 and Ser704) as well as numerous hydrophobic interactions which indole rings make with non-polar residues of the binding pocket. In the crystal structure 4OTH the propyl carbamimidothioate flexible chain of Ro-318220 makes further hydrogen bond to Asp708 in the ribose pocket, whereas in the docking solution it is positioned between Asn751 and Asp764 replacing the position of the conserved water molecule which is observed in the X-ray structure. The waters were not considered in homology modeling due to the uncertainties in the exact water location as well as its role in ligand binding. Taking into account absence of water and conformational flexibility of propyl carbamimidothioate chain of Ro-318220, it is not surprising that docking program identified another solution able to fit into the binding site.

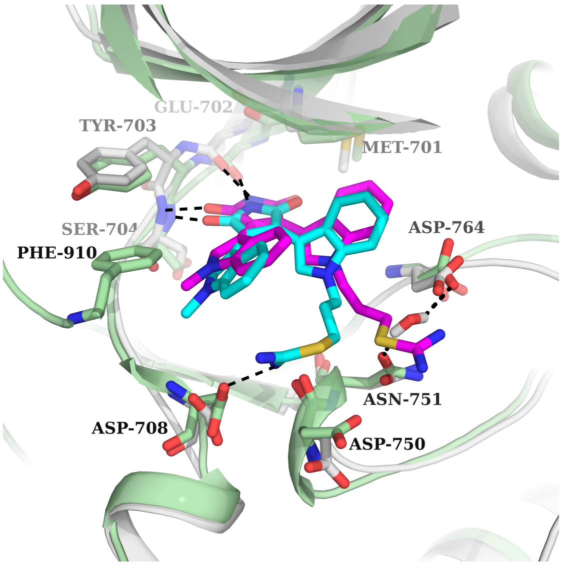


Figure 3.21: Comparison of the binding mode of compound Ro-318220 observed in crystal structure (PDB code 4OTH, protein is shown in pale green and ligand in cyan) and the top-scored by glideSP pose from the ensemble docking (hm_2jed is colored in grey and ligand in magenta). The conserved water interacting with Asn751 and Asp764 derived from the structure 4OTH is shown.

Lestaurtinib is another staurosporine analog discussed here. It has been described as orally available and potent inhibitor of kinases TrK[191], FLT3,[192] JAK2[193] and PRK1.[144] Lestaurtinib was initially developed for the treatment of acute myelogenous leukemia (AML) bearing FLT3 activating mutations,[194, 195] however, the results of phase III clinical trial showed no increase in response rates or prolong of survival of patients with FLT3 mutant AML in first relapse.[196] Additionally, lestaurtinib has been tested in phase II study (NCT00081601, see http://www.cancer.gov/clinicaltrials) for the treatments of the hormone refractory recurrent prostate cancer, showing no significant prostate specific antigen (PSA) response; however, it was speculated that PSA may be not an adequate indicator of response for prostate cancer treatment.[197] Thus, there was no adequate study performed to identify the response of prostate cancer on the treatment with lestaurtinib. The binding mode of lestaurtinib is similar to staurosporine or its analogs, e.g. Ro-318220, exhibiting two hydrogen bond interactions between lactam amide group and hinge residues Glu702 and Ser704, and numerous hydrophobic contacts between the large planar aromatic system of the inhibitor and non-polar residues in the binding pocket ($\mathrm{RMSD}_{\mathrm{lestaurtinib}}$= 2.04 Å, see Table 3.12 and Figure 3.22). The hydroxymethyl group of lestaurtinib is predicted by the docking program to interact with side chain of Asn751 (see magenta ligand on Figure 3.22), while in the crystal structure 4OTG it is pointing towards the glycine-rich loop (see cyan ligand on Figure 3.22).

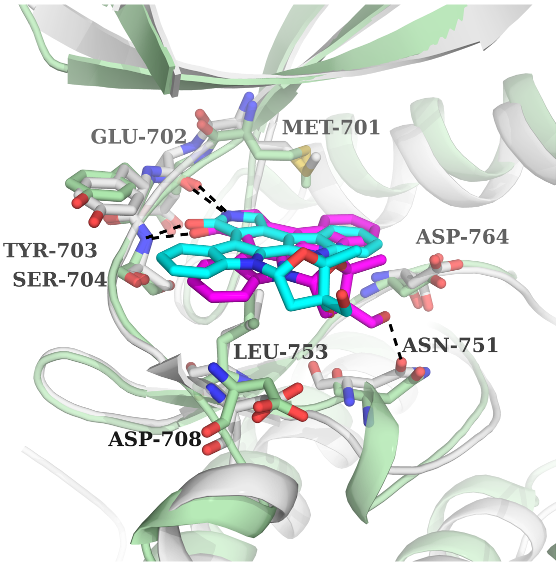


Figure 3.22: Comparison of the binding mode of compound lestaurtinib observed in crystal structure (PDB code 4OTG, protein is shown in pale green and ligand in cyan) and the top-scored by glideSP pose from ensemble docking (hm_2esm is colored in grey and ligand in magenta). G-rich loop is omitted for clarity.

Tofacitinib (CP-690550) is an ATP-competitive pan-JAK inhibitor,[17, 198] which is of particular interest due to its outstanding kinome selectivity for JAK, TYK and PRK kinases.[19, 165] In November 2012 it has been approved by the US FDA for the treatment of severe rheumatoid arthritis (RA) with inadequate response or intolerance to methotrexate therapy.[199–201] It is marketed by Pfizer under the trade names Xeljanz and Jakvinus. Tofacitinib has been also studied for the treatment of other autoimmune diseases, e.g. psoriasis[202] and inflammatory bowel disease,[203] as well as for the prevention of organ transplant rejection.[17, 204] The crystal structure of PRK1 with tofacitinib (PDB code 4OTI) and the top-scored docking solution with homology model hm_CP show almost identical conformations of tofacitinib with $\text{RMSD}_{\text{tofacitinib}}$ of 1.16 Å (see Table 3.12 and Figure 3.23).

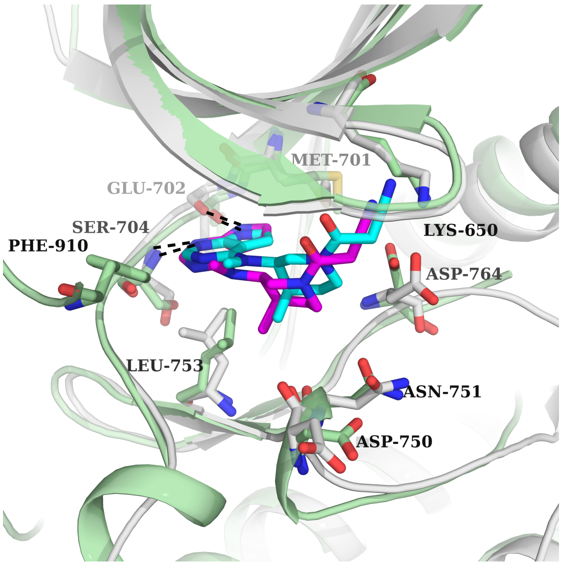


Figure 3.23: Comparison of the binding mode of compound tofacitinib observed in crystal structure (PDB code 4OTI, protein is shown in pale green and ligand in cyan) and the top-scored by glide_SP pose from ensemble docking (hm_CP is colored in grey and ligand in magenta).

This binding mode is also in agreement with previously solved X-ray structures of tofacitinib with JAK1, JAK3 and TYK2 tyrosine kinases (PDB codes 3EYG, 3LXK and 3LXN, respectively). The pyrrolopyrimidine scaffold of the inhibitor forms classical two hydrogen bond interactions to the hinge region of PRK1 (residues Glu702 and

Ser704), the 4-methyl-piperidine ring makes hydrophobic contacts with surrounding residues in the lipophilic cavity (e.g. Met701, Leu753, Leu627, Val635, Ala763) ideally filling the underlying pocket, and the terminal cyanoacetamide functionality is situated under the glycine-rich loop (G-loop) forming polar contacts with main chain of Gly634, Lys634, Val635, thus, stabilizing the G-loop which is known to be conformationally flexible and may form interactions with diverse ligands. The SAR studies for tofacitinib has shown that the cyanoacetamide side chain plays an important role in the binding and ensures an optimal balance between ligand affinity, selectivity and pharmacokinetic properties.[17] The above-described features of tofacitinib ensure the high degree of shape complementarity and, as a result, its exclusive kinome selectivity.

In summary, we could show that the homology models can be effectively used to reproduce known ligand binding modes, nevertheless, some differences between the predicted and experimental binding poses could be observed, as it is shown above. Various factors can affect the docking prediction, e.g. model quality, absence/presence of water, receptor and ligand flexibility. These aspects should be considered whenever possible.

CHAPTER 4

Conclusions

4.1 Application of Computer-Based Approaches to Search for Novel PRK1 Inhibitors

4.1.1 Structure-Based and Ligand-Based VS

The present study describes the application of computer-aided drug design approaches for the search of novel inhibitors of protein kinase C-related kinase 1 (PRK1). Various approaches were applied for this purpose: homology modeling, ligand- and structure-based virtual screening as well as binding free energy calculations for predicting the biological activity.

PRK1 has been shown to be implicated in numerous cellular processes and is particularly associated with prostate cancer, which makes it an attractive target for an anticancer drug development. Despite the profound effect of PRK1 on the propagation of androgen receptor driven cancer cell proliferation, there was little information available about the PRK1 structure and its inhibitors. Due to the absence of a 3D structure of PRK1 at the start of the project, we first developed a homology model of this enzyme. Docking studies for known active and inactive compounds from our in-house library were performed. We observed that the correct pose of kinase inhibitors could be identified in most cases by the docking programs, but this conformation was not always found to be top ranked, especially if the protein adopted another conformation than the specific ligand requires.

Thus, multiple protein models based on inhibitor-bound conformational states of PRK1 were introduced in order to address protein flexibility. The results showed improved docking performance in pose prediction and enrichment when an ensemble of protein conformations was used.

Next, structure-based and ligand-based virtual screenings were applied to search for novel PRK1 inhibitors. These included similarity search using ligand fingerprints MOLPRINT2D and MACCS keys followed by ligand docking as well as pharmacophore-based VS. The hits identified in VS experiments were further tested for their *in vitro* biological activity. Several PRK1 inhibitors with different scaffolds and potency have been identified by these initial screenings, see Section 3.4.

4.1.2 Prediction of the Biological Activity

Another question we addressed in this work was the prediction of the biological activity of novel compounds. The initial model based only on the docking scores could not explain the activity of compounds as shown by the results of our previous ligand-based or pharmacophore-based VS. The compounds, which were selected on the basis of favorable docking scores, were shown to be only moderately active. Thus, there is a need for more sophisticated methods that are able to evaluate the activity of compounds. The application of binding free energy calculations using MM-PBSA and MM-GBSA approaches after molecular dynamics simulation, which was performed for a small subset of 14 compounds from Biomol library (see Section 3.3), has proven the potential for activity prediction. However, such an approach is time-demanding and cannot be performed for a large dataset. In order to find the optimal balance between the invested time and the quality of the results, we tested the application of BFE calculations using a simplified method which excludes molecular dynamics and explicit solvent (see Section 3.5).

The approach presented in this work employs the estimation of binding free energy using MM-PBSA, MM-GBSA and QM/MM-GBSA calculations after a short minimization of protein-ligand complex using an implicit solvent model. The results demonstrate that using only one snapshot derived after the minimization is sufficient to provide a reasonable prediction of the binding free energy, but only if the correct starting conformation is used. Therefore, in the absence of crystal structures, docking to ensemble of homology models can increase the chance to obtain the correct pose for the ligands.

We observed that a significant correlation between calculated relative binding free energies derived from the MM-PBSA calculations or predicted $\mathrm{pIC_{50}}$ by developed QSAR models and experimental $\mathrm{IC_{50}}$ values for PRK1 inhibitors can be obtained ($r^2 = 0.61$ using MM-PBSA approach and $r^2 = 0.78$ using QSAR_model_3 on the training set). The results of the validation on a test dataset, containing diverse PRK1 inhibitors with reported $\mathrm{K_d}$ values, confirmed the good predictive ability of the final QSAR model ($r^2 = 0.68$, $q^2 = 0.59$, $r_0^2 = 0.67$ using QSAR_model_3, one outlier) Furthermore, the approach was able to distinguish potent PRK1 inhibitors ($\mathrm{IC_{50}}$<100 nM) from moderate/weak binders ($\mathrm{IC_{50}}$>100 nM). Nevertheless, despite the good results in the predicting the biological activities of known PRK1 inhibitors, the above-described rescoring methodology has the disadvantage of weaker discriminating power between active and inactive compounds if compared to the GlideSP score. Thus, the suggested virtual screening strategy for the search of potential PRK1 inhibitors should comprise a combination of ensemble docking methodology with subsequent rescoring of the preselected best-scored compounds from the previous step. Moreover, the compiled in house library of 28 PRK1 inhibitors and 300 true decoys together with the developed approach for predicting the biological activity could be especially useful in the further optimization step of PRK1 kinase inhibitors.

Furthermore, the rescoring methodology developed in this work was validated on two external datasets (see Section 3.6), which resulted in the identification of a number of highly potent PRK1 inhibitors, proving the validity of the model.

4.1.3 Comparison of Different VS Approaches to Search for PRK1 Inhibitors

In summary, we could see that different virtual screening approaches such as pharmacophore- or similarity-based methods as well as docking followed by the rescoring using BFE calculations were successfully applied for a search of new PRK1 inhibitors. The resulting hit rates were found to be in a range of 29 to 100 % (see Table see Table 4.1). These numbers are obviously higher than the hit rate of *in vitro* screening using kinase-focused Biomol library (~3 %, see Table 4.1). In comparison, the average hit rate of high-throughput screening (HTS) may vary from 0.1 to 2 % depending on the assay conditions and the target.[205] These observations clearly show that the performance of virtual screening, which utilizes knowledge about the structure of a target or its inhibitors, exceeds the performance of random screening, thus, improving the time and cost efficiency of lead discovery process.

Table 4.1: The overview of the screening performance using different approaches for a search of PRK1 inhibitors. The overall hit rate (the number of compounds that bind to PRK1, "№ of active", divided by the number of experimentally tested molecules, "№ of tested") and the potency of identified inhibitors ("Activity range", measured as IC_{50}) are shown.

Method	**Ligand/ Database**	**№ of tested**	**№ of active**	**Hit rate**	**Activity range, IC_{50}**
Focused library screening	Biomol Kinase Inhibitor Set	287	9	3 %	0.8 nM ~ 2 μM
Similarity-based VS (MOLPRINT2D fingerprints)	HA-1077, H-7	4	3	75 %	29 ~ 70 μM
Similarity-based VS (MACCS fingerprints)	CP-690550	18	9	50 %	1 ~ 56 μM
Pharmacophore-based VS	PD-0166285	7	2	29 %	3 ~ 5 μM
Ensemble docking +BFE rescoring	Selleck Kinase Inhibitor Set	5	5	100 %	11 ~ 270 nM
Ensemble docking +BFE rescoring	GSK Kinase Inhibitor Set	25	8	32 %	40 nM ~ 5 μM

Even though such approaches as similarity- or pharmacophore-based VS followed by docking were successful in the identification of the molecules which bind to PRK1, they were not able to discriminate compounds by their activities - the hits found by these screenings were micromolar PRK1 inhibitors. The choice of compounds for *in vitro* testing in this case was based on a high docking score and visual inspection of ligand-receptor interactions, thus, there may be a few reasons why these approaches failed to identify highly potent PRK1 inhibitors:

- Scoring. The primary task of the docking score is the accurate pose prediction and enrichment of active compounds among the top-ranked compounds in a database; docking score is not very effective in predicting the biological activity of compounds, especially for structurally diverse molecules.
- Protein flexibility. If a wrong conformation of the protein binding pocket is used in the docking experiment, the program may fail to recognize the correct ligand binding mode. Most of the docking methodologies are sensitive for structural distortions of the binding pocket and require a high-quality structure of the receptor, which, ideally, must be co-crystallized with a ligand similar to docked compounds.
- Pose prediction. The correct pose of the inhibitor may be not identified in case if the active ligand conformation was not sampled or it was not ranked as the best one.

In contrast to the initial similarity- or pharmacophore-based VS, the docking combined with BFE rescoring was able to identify novel nanomolar PRK1 inhibitors, as it was shown on example of Selleck and GSK database screening (Sections 3.6.1 and 3.6.2). However, some false positives were present among the selected hits (see the results for GSK kinase inhibitor set, Table 3.10). As it was discussed previously (see Section 3.5.4.1), the rescoring using minimization and BFE calculations may, first, allow the protein to adjust for bigger ligands and, second, introduce larger errors when the energy of the receptor is taken into account. As we could see from the results for GSK kinase inhibitor set, the compounds selected for *in vitro* test had wide range of predicted pIC_{50} values ($4.84 \sim 7.02$, see Table 3.10) while most of false-positive hits had predicted pIC_{50} less than 6. In this case the selection of compounds was guided also by their availability.

Since the GSK PKIS set is available free of charge, we decided to test 25 compounds out of 367, which showed the best activity predictions including poorly predicted hits (predicted pIC_{50} of $4 \sim 5$). However, when only a small number of compounds can be tested, it is recommended to use higher threshold for the predicted pIC_{50}.

Following these observations, next recommendations can be concluded in order to improve the VS performance:

- Such approaches as similarity-, pharmacophore- or docking-based VS can be effectively used for the identification of PRK1 inhibitors, nevertheless, their scores usually do not reflect the biological activity of the compounds. In order to prioritize hits for further testing and to improve the screening success, ligand- and structure-based VS methodologies should be combined with post-processing approaches, e.g. rescoring using BFE calculations. The combination of different methodologies may help to compensate the disadvantages of each individual method.
- If the highly active hits are of interest, it is recommended to select compounds with predicted pIC_{50} (QSAR_model_3) above 6.

4.2 Analysis of PRK1 Crystal Structures

The recently released crystal structures of PRK1 (apo- and holo-, see Section 3.7) were compared to the homology models developed in this work as well as to the corresponding docking poses for ligands Ro-318220, lestaurtinib and tofacitinib. The results show no significant difference between the experimentally-determined and predicted PRK1 structures, especially in the region of the ATP-binding pocket ($RMSD_{\text{pocket Calpha}} = 0.57$ Å ~ 1.54 Å). Another important finding underlines the role of protein flexibility for ligand binding, showing that the loops of PRK1 outlining the ATP-binding pocket (e.g. C-tail and G-loop) may adopt different conformations and such conformational difference should be taken into account during the docking experiments, especially when structurally-diverse compounds are investigated. The ligand poses generated by the docking program overall are in agreement with poses determined crystallographically, exhibiting minor differences in conformations of rotatable bonds, which can be explained by the ligand flexibility or absence of the water molecules in the model. In general, the current findings demonstrate the usefulness of PRK1 homology models, which were shown to be effectively applied in virtual screenings.

4.3 Design of New PRK1 Inhibitors

During the last years a number of PRK1 inhibitors were identified using both *in vitro* and *in silico* screening techniques. These compounds represent different chemical classes and display various potency and selectivity on PRK1. The knowledge about the structure and activity of known PRK1 inhibitors as well as the analysis of their common and distinct features can be further utilized for the design of novel compounds which aim to inhibit PRK1 potently and selectively.

Below, the general conclusions and recommendations for the design of novel PRK1 inhibitors are discussed using tofacitinib (CP-690550) as an example. Tofacitinib represents one of the most potent PRK1 inhibitors identified. Being type I inhibitor, which often suffers from cross-reactivity, it has an exceptional selectivity profile (see Figure 4.1) and it represents an attractive lead for further SAR studies and optimization for PRK1. The ATP-binding pocket of PRK1 was compared to JAK and possible strategies for chemical modifications of tofacitinib were suggested.

1. The known PRK1 inhibitors are ATP-competitive and target active form of the kinase (type I inhibitors). The crystal structures of PRK1 as well as of related kinases PKC, PKA are solved in their active conformation. Furthermore, it is not known whether PRK1 is able to adopt inactive conformation, which seriously hampers the development of type II inhibitors. Moreover, the discovery of type I inhibitors tofacitinib and BIBW-2992 proved that it is possible to reach high selectivity levels comparable to most of type II inhibitors.[19]

2. All known PRK1 inhibitors interact with the hinge region - different hinge binding motifs can be studied in SAR on tofacitinib. The additional hydrogen, halogen or hydro-

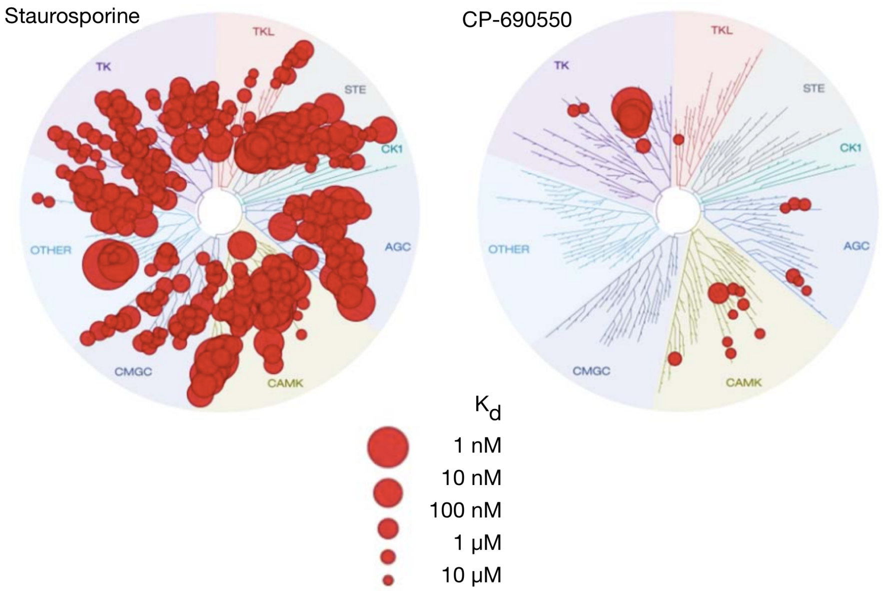


Figure 4.1: Kinome interaction maps for PRK1 inhibitors staurosporine and CP-690550 (tofacitinib) tested in a comprehensive analysis of kinase inhibitor selectivity by Davis at al.[19] The red circles indicate the affinity of a compound to a kinase, the bigger circle means higher affinity.

phobic interactions with the hinge region or surrounding residues might be introduced to tofacitinib by derivatizing of pyrrolopyrimidine scaffold with small substituents (chloro, methyl or amino groups), see positions 1 and 2 in Figure 4.2a. Additionally, changing of the pyrrolopyrimidine ring to a pyrrolopyridine might influence the PRK1/JAK selectivity ratio, since the pyrrolopyrimidine nitrogen of tofacitinib (see position 3 in Figure 4.2a) is known to interact with a conserved water molecule which is present in the JAK and TYK2 crystal structures (PDB codes: 3EYG, 2FUP, 3LXK and 3LXN) but not in PRK1 (PDB code: 4OTI), which has a Phe910 residue of the C-tail at this position, see Figure 4.2b.

3. Most of the potent PRK1 inhibitors make additional hydrogen bonds to polar residues in the ATP sugar binding region (e.g. Asp708, Asp750). The differences at these positions can be further exploited in order to improve the PRK1/JAK selectivity (these residues are mutated to Cys909 and Arg953 in JAK3, see Figure 4.2b). The

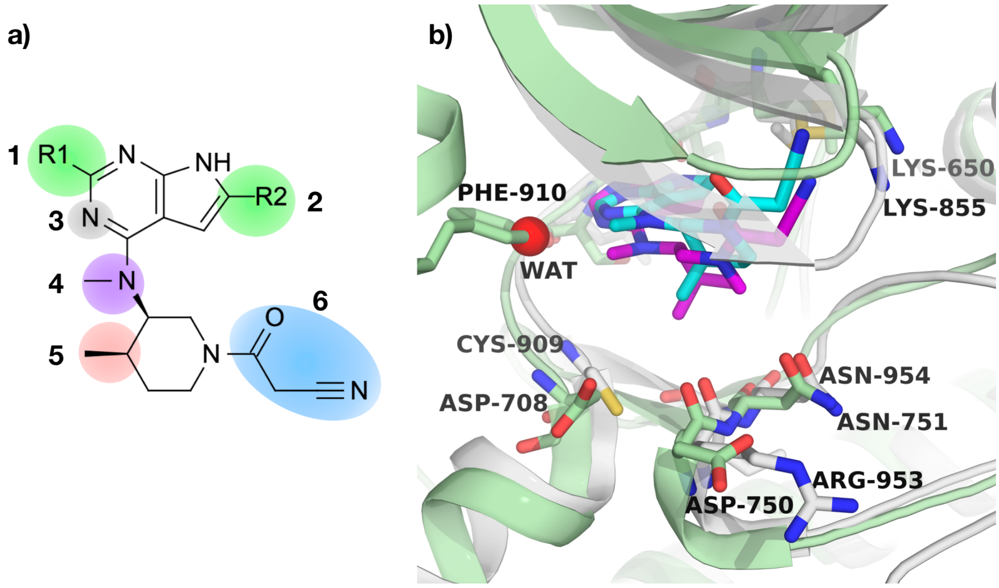


Figure 4.2: Structure and binding mode of PRK1 inhibitor tofacitinib (CP-690550): **a)** the chemical structure of tofacitinib and the positions for the suggested modifications to this compound; **b)** the comparison of ATP-binding pockets of JAK3 kinase (PDB code: 3LXK, light grey protein and magenta ligand, the conserved water interacting with tofacitinib is depicted as a red sphere) and PRK1 (PDB code: 4OTI, pale green protein and cyan ligand) in complex with tofacitinib.

hydrogen-bond donating groups can be introduced to tofacitinib in order to target the residues Asp708 or Asp750 by modifying compound at positions 4 and 5 (amino-methyl group and methyl of piperidine, see Figure 4.2a).

4. The hydrophobic pockets I and II (see Figure 1.3) are often explored to modulate selectivity of kinase inhibitors. The derivatization of the tofacitinib pyrrolopyrimidine (or, alternatively, pyrrolopyridine) scaffold with various hydrophobic groups including chloro, bromo, fluoro substituents can be done in order to target these pockets.

5. Another distinct feature of PRK1, as well as other AGC kinases, is the presence of the C-terminal tail which can insert Phe904 into the ATP-binding pocket, while many other kinases including JAKs are solvent-accessible at this region (see Figure 4.2b). Like it was shown for inhibitors such as H-7 and HA-1077, the isoquinoline scaffold makes aromatic interaction with Phe904. The pyrrolopyrimidine scaffold and the amino-methyl group of tofacitinib are also able to form van der Waals contacts to the side chain of

Phe904, see crystal structure 4OTI.[184] Different C-tail conformations and possibility of interactions with Phe904 should be considered during the design of potent and selective PRK1 inhibitors.

6. The glycine-rich loop (G-loop) is known to adopt different conformations and to play a role in selectivity modulation.[206] The alpha cyano acetamide moiety of tofacitinib fits well into the small pocket formed by G-loop and the catalytic Lys650 and makes numerous contacts with G-loop residues, which might be one of the reasons for its high kinome selectivity. The replacement of the cyano acetamide moiety of tofacitinib (see position 6 in Figure 4.2a) to other chemical groups with similar features might help to understand the structure-activity relationships better and to design new compounds with improved selectivity to PRK1.

In summary, tofacitinib as well as other PRK1 inhibitors identified in the current work can be further explored in structure-activity relationships studies and serve as a starting point for lead optimization in order to improve potency and selectivity for PRK1.

Bibliography

(1) Cheetham, G. M. T. Novel protein kinases and molecular mechanisms of autoinhibition. *Current opinion in structural biology* **2004**, *14*, 700–705.

(2) Kondapalli, L.; Soltani, K.; Lacouture, M. E. The promise of molecular targeted therapies: protein kinase inhibitors in the treatment of cutaneous malignancies. *Journal of the American Academy of Dermatology* **2005**, *53*, 291–302.

(3) Cohen, P. Protein kinases–the major drug targets of the twenty-first century? *Nature Reviews. Drug Discovery* **2002**, *1*, 309–315.

(4) Gassel, M.; Breitenlechner, C. B.; König, N.; Huber, R.; Engh, R. A.; Bossemeyer, D. The protein kinase C inhibitor bisindolyl maleimide 2 binds with reversed orientations to different conformations of protein kinase A. *The Journal of biological chemistry* **2004**, *279*, 23679–23690.

(5) Manning, G; Whyte, D. B.; Martinez, R; Hunter, T; Sudarsanam, S The protein kinase complement of the human genome. *Science* **2002**, *298*, 1912–1934.

(6) FDA, http://www.fda.gov, 2014.

(7) Daub, H.; Specht, K.; Ullrich, A. Strategies to overcome resistance to targeted protein kinase inhibitors. *Nature Reviews Drug Discovery* **2004**, *3*, 1001–1010.

(8) Morphy, R.; Kay, C.; Rankovic, Z. From magic bullets to designed multiple ligands. *Drug discovery today* **2004**, *9*, 641–651.

(9) Broekman, F.; Giovannetti, E.; Peters, G. J. Tyrosine kinase inhibitors: Multi-targeted or single-targeted? *World Journal of Clinical Oncology* **2011**, *2*, 80–93.

(10) Tamaoki, T; Nomoto, H; Takahashi, I; Kato, Y; Morimoto, M; Tomita, F Staurosporine, a potent inhibitor of phospholipid/Ca++dependent protein kinase. *Biochemical and biophysical research communications* **1986**, *135*, 397–402.

(11) Rüegg, U. T.; Burgess, G. M. Staurosporine, K-252 and UCN-01: potent but nonspecific inhibitors of protein kinases. *Trends in pharmacological sciences* **1989**, *10*, 218–220.

(12) Knight, Z. A.; Shokat, K. M. Features of selective kinase inhibitors. *Chemistry & biology* **2005**, *12*, 621–637.

(13) Liao, J. J. Molecular recognition of protein kinase binding pockets for design of potent and selective kinase inhibitors. *Journal of medicinal chemistry* **2007**, *50*, 409–424.

(14) Xing, L. et al. Structural bioinformatics-based prediction of exceptional selectivity of p38 MAP kinase inhibitor PH-797804. *Biochemistry* **2009**, *48*, 6402–6411.

(15) Cherry, M; Williams, D. H. Recent kinase and kinase inhibitor X-ray structures: mechanisms of inhibition and selectivity insights. *Current medicinal chemistry* **2004**, *11*, 663–673.

(16) Guimarães, C. R. W.; Rai, B. K.; Munchhof, M. J.; Liu, S.; Wang, J.; Bhattacharya, S. K.; Buckbinder, L. Understanding the impact of the P-loop conformation on kinase selectivity. *Journal of chemical information and modeling* **2011**, *51*, 1199–1204.

(17) Flanagan, M. E. et al. Discovery of CP-690,550: a potent and selective Janus kinase (JAK) inhibitor for the treatment of autoimmune diseases and organ transplant rejection. *Journal of medicinal chemistry* **2010**, *53*, 8468–8484.

(18) Vijayakrishnan, L.; Venkataramanan, R; Gulati, P. Treating inflammation with the Janus kinase inhibitor CP-690,550. *Trends in pharmacological sciences* **2011**, *32*, 25–34.

(19) Davis, M. I.; Hunt, J. P.; Herrgard, S.; Ciceri, P.; Wodicka, L. M.; Pallares, G.; Hocker, M.; Treiber, D. K.; Zarrinkar, P. P. Comprehensive analysis of kinase inhibitor selectivity. *Nature biotechnology* **2011**, *29*, 1046–1051.

(20) Martz, K. E.; Dorn, A.; Baur, B.; Schattel, V.; Goettert, M. I.; Mayer-Wrangowski, S. C.; Rauh, D.; Laufer, S. A. Targeting the hinge glycine flip and the activation loop: novel approach to potent p38α inhibitors. *Journal of medicinal chemistry* **2012**, *55*, 7862–7874.

(21) O'Brien, S. G.; Deininger, M. W. N. Imatinib in patients with newly diagnosed chronic-phase chronic myeloid leukemia. *Seminars in hematology* **2003**, *40*, 26–30.

(22) Schindler, T; Bornmann, W; Pellicena, P; Miller, W. T.; Clarkson, B; Kuriyan, J Structural mechanism for STI-571 inhibition of abelson tyrosine kinase. *Science (New York, N.Y.)* **2000**, *289*, 1938–1942.

(23) Simard, J. R.; Klüter, S.; Grütter, C.; Getlik, M.; Rabiller, M.; Rode, H. B.; Rauh, D. A new screening assay for allosteric inhibitors of cSrc. *Nature chemical biology* **2009**, *5*, 394–396.

(24) Ohren, J. F. et al. Structures of human MAP kinase kinase 1 (MEK1) and MEK2 describe novel noncompetitive kinase inhibition. *Nature structural & molecular biology* **2004**, *11*, 1192–1197.

(25) Mukai, H; Ono, Y A novel protein kinase with leucine zipper-like sequences: its catalytic domain is highly homologous to that of protein kinase C. *Biochemical and biophysical research communications* **1994**, *199*, 897–904.

(26) Mukai, H. The structure and function of PKN, a protein kinase having a catalytic domain homologous to that of PKC. *Journal of biochemistry* **2003**, *133*, 17–27.

(27) Lim, W. G.; Tan, B. J.; Zhu, Y.; Zhou, S.; Armstrong, J. S.; Li, Q. T.; Dong, Q.; Chan, E.; Smith, D.; Verma, C.; Tan, S.; Duan, W. The very C-terminus of PRK1/PKN is essential for its activation by RhoA and downstream signaling. *Cellular Signalling* **2006**, *18*, 1473–1481.

(28) Amano, M; Mukai, H; Ono, Y; Chihara, K; Matsui, T; Hamajima, Y; Okawa, K; Iwamatsu, A; Kaibuchi, K Identification of a putative target for Rho as the serine-threonine kinase protein kinase N. *Science (New York, N.Y.)* **1996**, *271*, 648–650.

(29) Watanabe, G.; Saito, Y.; Madaule, P.; Ishizaki, T.; Fujisawa, K.; Morii, N.; Mukai, H.; Ono, Y.; Kakizuka, A.; Narumiya, S. Protein kinase N (PKN) and PKN-related protein rhophilin as targets of small GTPase Rho. *Science (New York, N.Y.)* **1996**, *271*, 645–648.

(30) Maesaki, R; Ihara, K; Shimizu, T; Kuroda, S; Kaibuchi, K; Hakoshima, T The structural basis of Rho effector recognition revealed by the crystal structure of human RhoA complexed with the effector domain of PKN/PRK1. *Molecular cell* **1999**, *4*, 793–803.

(31) Hutchinson, C. L.; Lowe, P. N.; McLaughlin, S. H.; Mott, H. R.; Owen, D. Mutational analysis reveals a single binding interface between RhoA and its effector, PRK1. *Biochemistry* **2011**, *50*, 2860–2869.

(32) Metzger, E.; Müller, J. M.; Ferrari, S.; Buettner, R.; Schüle, R. A novel inducible transactivation domain in the androgen receptor: implications for PRK in prostate cancer. *The EMBO journal* **2003**, *22*, 270–280.

(33) Metzger, E.; Imhof, A.; Patel, D.; Kahl, P.; Hoffmeyer, K.; Friedrichs, N.; Müller, J. M.; Greschik, H.; Kirfel, J.; Ji, S.; Kunowska, N.; Beisenherz-Huss, C.; Günther, T.; Buettner, R.; Schüle, R. Phosphorylation of histone H3T6 by PKCβI controls demethylation at histone H3K4. *Nature* **2010**, *464*, 792–796.

(34) Turner, E. C.; Kavanagh, D. J.; Mulvaney, E. P.; McLean, C.; Wikström, K.; Reid, H. M.; Kinsella, B. T. Identification of an interaction between the TPalpha and TPbeta isoforms of the human thromboxane A2 receptor with protein kinase C-related kinase (PRK) 1: implications for prostate cancer. *The Journal of Biological Chemistry* **2011**, *286*, 15440–15457.

(35) Mukai, H.; Miyahara, M.; Sunakawa, H.; Shibata, H.; Toshimori, M.; Kitagawa, M.; Shimakawa, M.; Takanaga, H.; Ono, Y. Translocation of PKN from the cytosol to the nucleus induced by stresses. *Proceedings of the National Academy of Sciences of the United States of America* **1996**, *93*, 10195–10199.

(36) Palmer, R. H.; Dekker, L. V.; Woscholski, R.; Le Good, J. A.; Gigg, R.; Parker, P. J. Activation of PRK1 by phosphatidylinositol 4,5-bisphosphate and phosphatidylinositol 3,4,5-trisphosphate. A comparison with protein kinase C isotypes. *The Journal of Biological Chemistry* **1995**, *270*, 22412–22416.

(37) Yoshinaga, C.; Mukai, H.; Toshimori, M.; Miyamoto, M.; Ono, Y. Mutational analysis of the regulatory mechanism of PKN: the regulatory region of PKN contains an arachidonic acid-sensitive autoinhibitory domain. *Journal of Biochemistry* **1999**, *126*, 475–484.

(38) Mellor, H.; Parker, P. J. The extended protein kinase C superfamily. *The Biochemical Journal* **1998**, *332 (Pt 2)*, 281–292.

(39) Harrison, B. C.; Huynh, K.; Lundgaard, G. L.; Helmke, S. M.; Perryman, M. B.; McKinsey, T. A. Protein kinase C-related kinase targets nuclear localization signals in a subset of class IIa histone deacetylases. *FEBS letters* **2010**, *584*, 1103–1110.

(40) Galgano, M. T.; Conaway, M.; Spencer, A. M.; Paschal, B. M.; Frierson, H. F. PRK1 distribution in normal tissues and carcinomas: overexpression and activation in ovarian serous carcinoma. *Human Pathology* **2009**, *40*, 1434–1440.

(41) Yasui, T.; Sakakibara-Yada, K.; Nishimura, T.; Morita, K.; Tada, S.; Mosialos, G.; Kieff, E.; Kikutani, H. Protein kinase N1, a cell inhibitor of Akt kinase, has a central role in quality control of germinal center formation. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, 21022–21027.

(42) Metzger, E.; Yin, N.; Wissmann, M.; Kunowska, N.; Fischer, K.; Friedrichs, N.; Patnaik, D.; Higgins, J. M. G.; Potier, N.; Scheidtmann, K.; Buettner, R.; Schüle, R. Phosphorylation of histone H3 at threonine 11 establishes a novel chromatin mark for transcriptional regulation. *Nature cell biology* **2008**, *10*, 53–60.

(43) Lonergan, P. E.; Tindall, D. J. Androgen receptor signaling in prostate cancer development and progression. *Journal of Carcinogenesis* **2011**, *10*, 20.

(44) Heinlein, C. A.; Chang, C. Androgen receptor in prostate cancer. *Endocrine Reviews* **2004**, *25*, 276–308.

(45) Jemal, A.; Siegel, R.; Xu, J.; Ward, E. Cancer statistics, 2010. *CA: a cancer journal for clinicians* **2010**, *60*, 277–300.

(46) Roy, A. K.; Tyagi, R. K.; Song, C. S.; Lavrovsky, Y; Ahn, S. C.; Oh, T. S.; Chatterjee, B Androgen receptor: structural domains and functional dynamics after ligand-receptor interaction. *Annals of the New York Academy of Sciences* **2001**, *949*, 44–57.

(47) Wissmann, M.; Yin, N.; Müller, J. M.; Greschik, H.; Fodor, B. D.; Jenuwein, T.; Vogler, C.; Schneider, R.; Günther, T.; Buettner, R.; Metzger, E.; Schüle, R. Cooperative demethylation by JMJD2C and LSD1 promotes androgen receptor-dependent gene expression. *Nature Cell Biology* **2007**, *9*, 347–353.

(48) Croce, L. D.; Shiekhattar, R. Thrilling transcription through threonine phosphorylation. *Nature Cell Biology* **2008**, *10*, 5–6.

(49) Lachmann, S.; Jevons, A.; De Rycker, M.; Casamassima, A.; Radtke, S.; Collazos, A.; Parker, P. J. Regulatory domain selectivity in the cell-type specific PKN-dependence of cell migration. *PloS One* **2011**, *6*, e21732.

(50) Jilg, C. A.; Ketscher, A.; Metzger, E.; Hummel, B.; Willmann, D.; Rüsseler, V.; Drendel, V.; Imhof, A.; Jung, M.; Franz, H.; Hölz, S.; Krönig, M.; Müller, J. M.; Schüle, R. PRK1/PKN1 controls migration and metastasis of androgen-independent prostate cancer cells. *Oncotarget* **2014**, *5*, 12646–12664.

(51) Leach, A. R.; Shoichet, B. K.; Peishoff, C. E. Prediction of protein-ligand interactions. Docking and scoring: successes and gaps. *Journal of medicinal chemistry* **2006**, *49*, 5851–5855.

(52) Yuriev, E.; Ramsland, P. A. Latest developments in molecular docking: 2010-2011 in review. *Journal of molecular recognition: JMR* **2013**, *26*, 215–239.

(53) Cross, J. B.; Thompson, D. C.; Rai, B. K.; Baber, J. C.; Fan, K. Y.; Hu, Y.; Humblet, C. Comparison of Several Molecular Docking Programs: Pose Prediction and Virtual Screening Accuracy. *Journal of Chemical Information and Modeling* **2009**, *49*, 1455–1474.

(54) Meng, X.; Zhang, H.; Mezei, M.; Cui, M. Molecular Docking: A powerful approach for structure-based drug discovery. *Current computer-aided drug design* **2011**, *7*, 146–157.

(55) Klebe, G. Virtual ligand screening: strategies, perspectives and limitations. *Drug discovery today* **2006**, *11*, 580–594.

(56) Kubinyi, H. In*Computer Applications in Pharmaceutical Research and Development*, 2006, pp 377–424.

(57) Matter, H.; Sotriffer, C. In*Methods and Principles in Medicinal Chemistry*, 2011, pp 319–358.

(58) Seifert, M. H. J.; Lang, M. Essential factors for successful virtual screening. *Mini reviews in medicinal chemistry* **2008**, *8*, 63–72.

(59) Warren, G. L.; Andrews, C. W.; Capelli, A.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. A critical assessment of docking programs and scoring functions. *Journal of medicinal chemistry* **2006**, *49*, 5912–5931.

(60) Yuriev, E.; Agostino, M.; Ramsland, P. A. Challenges and advances in computational docking: 2009 in review. *Journal of molecular recognition: JMR* **2011**, *24*, 149–164.

(61) Scior, T.; Bender, A.; Tresadern, G.; Medina-Franco, J. L.; Martínez-Mayorga, K.; Langer, T.; Cuanalo-Contreras, K.; Agrafiotis, D. K. Recognizing Pitfalls in Virtual Screening: A Critical Review. *Journal of Chemical Information and Modeling* **2012**, *52*, 867–881.

(62) Erickson, J. A.; Jalaie, M.; Robertson, D. H.; Lewis, R. A.; Vieth, M. Lessons in molecular recognition: the effects of ligand and protein flexibility on molecular docking accuracy. *Journal of medicinal chemistry* **2004**, *47*, 45–55.

(63) McCammon, J. A. Target flexibility in molecular recognition. *Biochimica et biophysica acta* **2005**, *1754*, 221–224.

(64) Bolstad, E. S. D.; Anderson, A. C. In pursuit of virtual lead optimization: the role of the receptor structure and ensembles in accurate docking. *Proteins* **2008**, *73*, 566–580.

(65) Lin, J.; Perryman, A. L.; Schames, J. R.; McCammon, J. A. The relaxed complex method: Accommodating receptor flexibility for drug design with an improved scoring scheme. *Biopolymers* **2003**, *68*, 47–62.

(66) Zhong, S.; Zhang, Y.; Xiu, Z. Rescoring ligand docking poses. *Current opinion in drug discovery & development* **2010**, *13*, 326–334.

(67) Wang, J; Morin, P; Wang, W; Kollman, P. A. Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. *Journal of the American Chemical Society* **2001**, *123*, 5221–5230.

(68) Gilson, M. K.; Zhou, H. Calculation of protein-ligand binding affinities. *Annual review of biophysics and biomolecular structure* **2007**, *36*, 21–42.

(69) Mobley, D. L.; Dill, K. A. Binding of small-molecule ligands to proteins: "what you see" is not always "what you get". *Structure* **2009**, *17*, 489–498.

(70) Michel, J.; Essex, J. W. Prediction of protein-ligand binding affinity by free energy simulations: assumptions, pitfalls and expectations. *Journal of computer-aided molecular design* **2010**, *24*, 639–658.

(71) Greenidge, P. A.; Kramer, C.; Mozziconacci, J.; Wolf, R. M. MM/GBSA binding energy prediction on the PDBbind data set: successes, failures, and directions for further improvement. *Journal of chemical information and modeling* **2013**, *53*, 201–209.

(72) Homeyer, N.; Gohlke, H. FEW: a workflow tool for free energy calculations of ligand binding. *Journal of computational chemistry* **2013**, *34*, 965–973.

(73) Parenti, M. D.; Rastelli, G. Advances and applications of binding affinity prediction methods in drug discovery. *Biotechnology advances* **2012**, *30*, 244–250.

(74) Rastelli, G.; Del Rio, A.; Degliesposti, G.; Sgobba, M. Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. *Journal of computational chemistry* **2010**, *31*, 797–810.

(75) Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *Journal of chemical information and modeling* **2011**, *51*, 69–82.

(76) Uciechowska, U.; Schemies, J.; Scharfe, M.; Lawson, M.; Wichapong, K.; Jung, M.; Sippl, W. Binding free energy calculations and biological testing of novel thiobarbiturates as inhibitors of the human NAD+ dependent histone deacetylase Sirt2. *MedChemComm* **2012**, *3*, 167.

(77) Wichapong, K.; Lawson, M.; Pianwanit, S.; Kokpol, S.; Sippl, W. Postprocessing of Protein-Ligand Docking Poses Using Linear Response MM-PB/SA: Application to Wee1 Kinase Inhibitors. *Journal of Chemical Information and Modeling* **2010**, *50*, 1574–1588.

(78) Ferrari, A. M.; Degliesposti, G.; Sgobba, M.; Rastelli, G. Validation of an automated procedure for the prediction of relative free energies of binding on a set of aldose reductase inhibitors. *Bioorganic & Medicinal Chemistry* **2007**, *15*, 7865–7877.

(79) Kuhn, B.; Gerber, P.; Schulz-Gasch, T.; Stahl, M. Validation and use of the MM-PBSA approach for drug discovery. *Journal of medicinal chemistry* **2005**, *48*, 4040–4048.

(80) Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the performance of the molecular mechanics/Poisson Boltzmann surface area and molecular mechanics/generalized Born surface area methods. II. The accuracy of ranking poses generated from docking. *Journal of Computational Chemistry* **2011**, *32*, 866–877.

(81) Lindström, A.; Edvinsson, L.; Johansson, A.; Andersson, C. D.; Andersson, I. E.; Raubacher, F.; Linusson, A. Postprocessing of docked protein-ligand complexes using implicit solvation models. *Journal of chemical information and modeling* **2011**, *51*, 267–282.

(82) Lyne, P. D.; Lamb, M. L.; Saeh, J. C. Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. *Journal of medicinal chemistry* **2006**, *49*, 4805–4808.

(83) Rastelli, G.; Degliesposti, G.; Del Rio, A.; Sgobba, M. Binding estimation after refinement, a new automated procedure for the refinement and rescoring of docked ligands in virtual screening. *Chemical biology & drug design* **2009**, *73*, 283–286.

(84) Sgobba, M.; Caporuscio, F.; Anighoro, A.; Portioli, C.; Rastelli, G. Application of a post-docking procedure based on MM-PBSA and MM-GBSA on single and multiple protein conformations. *European journal of medicinal chemistry* **2012**, *58*, 431–440.

(85) Case, D. et al. Amber 11, 2010.

(86) Gleeson, M. P.; Gleeson, D. QM/MM as a tool in fragment based drug discovery. A cross-docking, rescoring study of kinase inhibitors. *Journal of chemical information and modeling* **2009**, *49*, 1437–1448.

(87) Hayik, S. A.; Dunbrack Roland, J.; Merz Kenneth M, J. A Mixed QM/MM Scoring Function to Predict Protein-Ligand Binding Affinity. *Journal of chemical theory and computation* **2010**, *6*, 3079–3091.

(88) Illingworth, C. J. R.; Morris, G. M.; Parkes, K. E. B.; Snell, C. R.; Reynolds, C. A. Assessing the role of polarization in docking. *The journal of physical chemistry. A* **2008**, *112*, 12157–12163.

(89) Cho, A. E.; Guallar, V.; Berne, B. J.; Friesner, R. Importance of accurate charges in molecular docking: quantum mechanical/molecular mechanical (QM/MM) approach. *Journal of computational chemistry* **2005**, *26*, 915–931.

(90) Garcia-Viloca, M.; Truhlar, D. G.; Gao, J. Importance of substrate and cofactor polarization in the active site of dihydrofolate reductase. *Journal of molecular biology* **2003**, *327*, 549–560.

(91) Ji, C. G.; Zhang, J. Z. H. Protein polarization is critical to stabilizing AF-2 and helix-2' domains in ligand binding to PPAR-gamma. *Journal of the American Chemical Society* **2008**, *130*, 17129–17133.

(92) Raha, K.; Peters, M. B.; Wang, B.; Yu, N.; Wollacott, A. M.; Westerhoff, L. M.; Merz Kenneth M, J. The role of quantum mechanics in structure-based drug design. *Drug discovery today* **2007**, *12*, 725–731.

(93) Walker, R. C.; Crowley, M. F.; Case, D. A. The implementation of a fast and accurate QM/MM potential method in Amber. *Journal of Computational Chemistry* **2008**, *29*, 1019–1031.

(94) Del Rio, A.; Baldi, B. F.; Rastelli, G. Activity prediction and structural insights of extracellular signal-regulated kinase 2 inhibitors with molecular dynamics simulations. *Chemical biology & drug design* **2009**, *74*, 630–635.

(95) Johnson, M.; Maggiora, G., *Concepts and applications of molecular similarity*; John Wiley and Sons: 1990.

(96) Bender, A.; Mussa, H.; Glen, R.; Reiling, S. Molecular Similarity Searching Using Atom Environments, Information-Based Feature Selection, and a Naïve Bayesian Classifier. *Journal of Chemical Information and Modeling* **2004**, *44*, 170–178.

(97) Bender, A.; Mussa, H.; Glen, R.; Reiling, S. Similarity Searching of Chemical Databases Using Atom Environment Descriptors (MOLPRINT 2D): Evaluation of Performance. *Journal of Chemical Information and Modeling* **2004**, *44*, 1708–1718.

(98) Willett, P. Similarity-based virtual screening using 2D fingerprints. *Drug discovery today* **2006**, *11*, 1046–1053.

(99) MOE (Molecular Operating Environment), v2012.10, 2012.

(100) Expasy, http://web.expasy.org/docs/relnotes/relstat.html, 2014.

(101) ProteinDataBank, http://www.pdb.org/pdb/statistics/holdings.do, 2014.

(102) Martí-Renom, M. A.; Stuart, A. C.; Fiser, A; Sánchez, R; Melo, F; Sali, A Comparative protein structure modeling of genes and genomes. *Annual review of biophysics and biomolecular structure* **2000**, *29*, 291–325.

(103) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic local alignment search tool. *Journal of Molecular Biology* **1990**, *215*, 403–410.

(104) NCBI BLAST, http://blast.be-md.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_-TYPE=BlastSearch&LINK_LOC=blasthome, 2010.

(105) Sali, A; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *Journal of molecular biology* **1993**, *234*, 779–815.

(106) Shen, M.; Sali, A. Statistical potential for assessment and prediction of protein structures. *Protein science: a publication of the Protein Society* **2006**, *15*, 2507–2524.

(107) Maestro, version 9.3, 2012.

(108) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *Journal of the American Chemical Society* **1996**, *118*, 11225–11236.

(109) Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. Prediction of Absolute Solvation Free Energies using Molecular Dynamics Free Energy Perturbation and the OPLS Force Field. *Journal of Chemical Theory and Computation* **2010**, *6*, 1509–1519.

(110) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* **1993**, *26*, 283–291.

(111) Morris, A. L.; MacArthur, M. W.; Hutchinson, E. G.; Thornton, J. M. Stereochemical quality of protein structure coordinates. *Proteins* **1992**, *12*, 345–364.

(112) Case, D. et al. Amber 12, 2012.

(113) Kellenberger, E.; Rodrigo, J.; Muller, P.; Rognan, D. Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins* **2004**, *57*, 225–242.

(114) Perola, E.; Walters, W. P.; Charifson, P. S. A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance. *Proteins* **2004**, *56*, 235–249.

(115) Kontoyianni, M.; McClellan, L. M.; Sokol, G. S. Evaluation of docking performance: comparative data on docking algorithms. *Journal of medicinal chemistry* **2004**, *47*, 558–565.

(116) Xue, M.; Zheng, M.; Xiong, B.; Li, Y.; Jiang, H.; Shen, J. Knowledge-based scoring functions in drug design. 1. Developing a target-specific method for kinase-ligand interactions. *Journal of Chemical Information and Modeling* **2010**, *50*, 1378–1386.

(117) O'Boyle, N. M.; Liebeschuetz, J. W.; Cole, J. C. Testing assumptions and hypotheses for rescoring success in protein-ligand docking. *Journal of chemical information and modeling* **2009**, *49*, 1871–1878.

(118) Charifson, P. S.; Corkery, J. J.; Murcko, M. A.; Walters, W. P. Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. *Journal of medicinal chemistry* **1999**, *42*, 5100–5109.

(119) Glide, version 5.8, 2012.

(120) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of medicinal chemistry* **2004**, *47*, 1739–1749.

(121) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *Journal of medicinal chemistry* **2006**, *49*, 6177–6196.

(122) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *Journal of medicinal chemistry* **2004**, *47*, 1750–1759.

(123) Jones, G.; Willett, P.; Glen, R. C. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *Journal of Molecular Biology* **1995**, *245*, 43–53.

(124) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *Journal of Molecular Biology* **1997**, *267*, 727–748.

(125) Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. Improved protein-ligand docking using GOLD. *Proteins* **2003**, *52*, 609–623.

(126) Meier, R.; Pippel, M.; Brandt, F.; Sippl, W.; Baldauf, C. ParaDockS: a framework for molecular docking with population-based metaheuristics. *Journal of Chemical Information and Modeling* **2010**, *50*, 879–889.

(127) Scharfe, M.; Pippel, M.; Sippl, W. ParaDockS - an open-source framework for molecular docking: implementation of target-class-specific scoring methods. *Journal of Cheminformatics* **2013**, *5*, P11.

(128) LigPrep, version 2.5, 2012.

(129) Wermuth, C. Glossary of terms used in medicinal chemistry (iupac recommendations 1997). *Annu. Rep. Med. Chem.* **1998**, *33*, 385–395.

(130) Wolber, G.; Langer, T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *Journal of chemical information and modeling* **2005**, *45*, 160–169.

(131) Steindl, T. M.; Schuster, D.; Laggner, C.; Langer, T. Parallel screening: a novel concept in pharmacophore modeling and virtual screening. *Journal of chemical information and modeling* **2006**, *46*, 2146–2157.

(132) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65*, 712–725.

(133) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *Journal of computational chemistry* **2004**, *25*, 1157–1174.

(134) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of molecular graphics & modelling* **2006**, *25*, 247–260.

(135) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983**, *79*, 926.

(136) Pastor, R. W.; Brooks, B. R.; Szabo, A. An analysis of the accuracy of Langevin and molecular dynamics algorithms. *Molecular Physics* **1988**, *65*, 1409–1419.

(137) Ryckaert, J.; Ciccotti, G.; Berendsen, H. J. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics* **1977**, *23*, 327–341.

(138) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N.log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* **1993**, *98*, 10089.

(139) Kollman, P. A.; Massova, I; Reyes, C; Kuhn, B; Huo, S; Chong, L; Lee, M; Lee, T; Duan, Y; Wang, W; Donini, O; Cieplak, P; Srinivasan, J; Case, D. A.; Cheatham T E, r. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Accounts of chemical research* **2000**, *33*, 889–897.

(140) Massova, I.; Kollman, P. A. Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. *Perspectives in Drug Discovery and Design* **2000**, *18*, 113.

(141) Connolly, M. L. Analytical molecular surface calculation. *Journal of Applied Crystallography* **1983**, *16*, 548–558.

(142) Weis, A.; Katebzadeh, K.; Söderhjelm, P.; Nilsson, I.; Ryde, U. Ligand affinities predicted with the MM/PBSA method: dependence on the simulation method and the force field. *Journal of medicinal chemistry* **2006**, *49*, 6596–6606.

(143) Brown, S. P.; Muchmore, S. W. Rapid Estimation of Relative Protein-Ligand Binding Affinities Using a High-Throughput Version of MM-PBSA. *Journal of Chemical Information and Modeling* **2007**, *47*, 1493–1503.

(144) Köhler, J.; Erlenkamp, G.; Eberlin, A.; Rumpf, T.; Slynko, I.; Metzger, E.; Schüle, R.; Sippl, W.; Jung, M. Lestaurtinib inhibits histone phosphorylation and androgen-dependent gene expression in prostate cancer cells. *PloS one* **2012**, *7*, e34973.

(145) Shang, Y.; Nguyen, H.; Wickstrom, L.; Okur, A.; Simmerling, C. Improving the description of salt bridge strength and geometry in a Generalized Born model. *Journal of Molecular Graphics & Modelling* **2011**, *29*, 676–684.

(146) Tan, C.; Yang, L.; Luo, R. How well does Poisson-Boltzmann implicit solvent agree with explicit solvent? A quantitative analysis. *The Journal of Physical Chemistry. B* **2006**, *110*, 18680–18687.

(147) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *Journal of computational chemistry* **2002**, *23*, 1623–1641.

(148) Onufriev, A.; Bashford, D.; Case, D. A. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins* **2004**, *55*, 383–394.

(149) Rocha, G. B.; Freire, R. O.; Simas, A. M.; Stewart, J. J. P. RM1: a reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I. *Journal of computational chemistry* **2006**, *27*, 1101–1111.

(150) Morrice, N. A.; Gabrielli, B; Kemp, B. E.; Wettenhall, R. E. A cardiolipin-activated protein kinase from rat liver structurally distinct from the protein kinases C. *The Journal of biological chemistry* **1994**, *269*, 20040–20046.

(151) Mukai, H.; Ono, Y. Purification and kinase assay of PKN. *Methods in enzymology* **2006**, *406*, 234–250.

(152) Amano, M; Fukata, Y; Kaibuchi, K Regulation and functions of Rho-associated kinase. *Experimental cell research* **2000**, *261*, 44–51.

(153) Amano, M; Chihara, K; Nakamura, N; Kaneko, T; Matsuura, Y; Kaibuchi, K The COOH terminus of Rho-kinase negatively regulates rho-kinase activity. *The Journal of biological chemistry* **1999**, *274*, 32418–32424.

(154) Xiang, Z. Advances in homology protein structure modeling. *Current Protein & Peptide Science* **2006**, *7*, 217–227.

(155) Kannan, N.; Haste, N.; Taylor, S. S.; Neuwald, A. F. The hallmark of AGC kinase functional divergence is its C-terminal tail, a cis-acting regulatory module. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104*, 1272–1277.

(156) Biondi, R. M.; Komander, D.; Thomas, C. C.; Lizcano, J. M.; Deak, M.; Alessi, D. R.; van Aalten, D. M. F. High resolution crystal structure of the human PDK1 catalytic domain defines the regulatory phosphopeptide docking site. *The EMBO journal* **2002**, *21*, 4219–4228.

(157) Balendran, A.; Biondi, R. M.; Cheung, P. C.; Casamayor, A.; Deak, M.; Alessi, D. R. A 3-phosphoinositide-dependent protein kinase-1 (PDK1) docking site is required for the phosphorylation of protein kinase Czeta (PKCzeta) and PKC-related kinase 2 by PDK1. *The Journal of Biological Chemistry* **2000**, *275*, 20806–20813.

(158) Prade, L.; Engh, R. A.; Girod, A.; Kinzel, V.; Huber, R.; Bossemeyer, D. Staurosporine-induced conformational changes of cAMP-dependent protein kinase catalytic subunit explain inhibitory potential. *Structure* **1997**, *5*, 1627–1637.

(159) García-Sosa, A. T.; Hetényi, C.; Maran, U. Drug efficiency indices for improvement of molecular docking scoring functions. *Journal of Computational Chemistry* **2010**, *31*, 174–184.

(160) Grinter, S. Z.; Zou, X. Challenges, applications, and recent advances of protein-ligand docking in structure-based drug design. *Molecules (Basel, Switzerland)* **2014**, *19*, 10150–10176.

(161) Shashi Nayana, M. R.; Sekhar, Y. N.; Siva Kumari, N.; Mahmood, S. K.; Ravikumar, M. CoMFA and docking studies on triazolopyridine oxazole derivatives as p38 MAP kinase inhibitors. *European Journal of Medicinal Chemistry* **2008**, *43*, 1261–1269.

(162) Cournia, Z.; Leng, L.; Gandavadi, S.; Du, X.; Bucala, R.; Jorgensen, W. L. Discovery of human macrophage migration inhibitory factor (MIF)-CD74 antagonists via virtual screening. *Journal of Medicinal Chemistry* **2009**, *52*, 416–424.

(163) Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. ZINC: A Free Tool to Discover Chemistry for Biology. *Journal of Chemical Information and Modeling* **2012**, *52*, 1757–1768.

(164) McGovern, S. L.; Shoichet, B. K. Kinase inhibitors: not just for kinases anymore. *Journal of Medicinal Chemistry* **2003**, *46*, 1478–1483.

(165) Karaman, M. W. et al. A quantitative analysis of kinase inhibitor selectivity. *Nature Biotechnology* **2008**, *26*, 127–132.

(166) Tasler, S.; Müller, O.; Wieber, T.; Herz, T.; Pegoraro, S.; Saeb, W.; Lang, M.; Krauss, R.; Totzke, F.; Zirrgiebel, U.; Ehlert, J. E.; Kubbutat, M. H. G.; Schächtele, C. Substituted 2-arylbenzothiazoles as kinase inhibitors: hit-to-lead optimization. *Bioorganic & Medicinal Chemistry* **2009**, *17*, 6728–6737.

(167) Apsel, B.; Blair, J. A.; Gonzalez, B.; Nazif, T. M.; Feldman, M. E.; Aizenstein, B.; Hoffman, R.; Williams, R. L.; Shokat, K. M.; Knight, Z. A. Targeted polypharmacology: discovery of dual inhibitors of tyrosine and phosphoinositide kinases. *Nature Chemical Biology* **2008**, *4*, 691–699.

(168) Selness, S. R. et al. Discovery of N-substituted pyridinones as potent and selective inhibitors of p38 kinase. *Bioorganic & Medicinal Chemistry Letters* **2009**, *19*, 5851–5856.

(169) Qiao, L.; Choi, S.; Case, A.; Gainer, T. G.; Seyb, K.; Glicksman, M. A.; Lo, D. C.; Stein, R. L.; Cuny, G. D. Structure-activity relationship study of EphB3 receptor tyrosine kinase inhibitors. *Bioorganic & Medicinal Chemistry Letters* **2009**, *19*, 6122–6126.

(170) Okuzumi, T.; Fiedler, D.; Zhang, C.; Gray, D. C.; Aizenstein, B.; Hoffman, R.; Shokat, K. M. Inhibitor hijacking of Akt activation. *Nature Chemical Biology* **2009**, *5*, 484–493.

(171) Degliesposti, G.; Portioli, C.; Parenti, M. D.; Rastelli, G. BEAR, a novel virtual screening methodology for drug discovery. *Journal of biomolecular screening* **2011**, *16*, 129–133.

(172) Graves, A. P.; Shivakumar, D. M.; Boyce, S. E.; Jacobson, M. P.; Case, D. A.; Shoichet, B. K. Rescoring Docking Hit Lists for Model Cavity Sites: Predictions and Experimental Testing. *Journal of Molecular Biology* **2008**, *377*, 914–934.

(173) Pickett, S. D.; Sternberg, M. J. Empirical scale of side-chain conformational entropy in protein folding. *Journal of molecular biology* **1993**, *231*, 825–839.

(174) Luo, R.; Gilson, M. K. Synthetic Adenine Receptors: Direct Calculation of Binding Affinity and Entropy. *Journal of the American Chemical Society* **2000**, *122*, 2934–2937.

(175) Golbraikh, A.; Tropsha, A. Beware of q2! *Journal of molecular graphics & modelling* **2002**, *20*, 269–276.

(176) Golbraikh, A.; Tropsha, A. Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection. *Journal of computer-aided molecular design* **2002**, *16*, 357–369.

(177) Golbraikh, A.; Shen, M.; Xiao, Z.; Xiao, Y.; Lee, K.; Tropsha, A. Rational selection of training and test sets for the development of validated QSAR models. *Journal of computer-aided molecular design* **2003**, *17*, 241–253.

(178) SelleckChemicals, http://www.selleckchem.com/screening/kinase-inhibitor-library.html, 2014.

(179) Yung-Chi, C.; Prusoff, W. H. Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochemical Pharmacology* **1973**, *22*, 3099–3108.

(180) *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*; Vogel, H. G., Maas, J., Gebauer, A., Eds.; Springer Berlin Heidelberg: 2011.

(181) Dranchak, P.; MacArthur, R.; Guha, R.; Zuercher, W. J.; Drewry, D. H.; Auld, D. S.; Inglese, J. Profile of the GSK Published Protein Kinase Inhibitor Set Across ATP-Dependent and-Independent Luciferases: Implications for Reporter-Gene Assays. *PLoS ONE* **2013**, *8*, e57888.

(182) ChEMBL The European Bioinformatics Institute, https://www.ebi.ac.uk/chembl/, 2014.

(183) Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington, J. P. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Research* **2012**, *40*, D1100–1107.

(184) Chamberlain, P.; Delker, S.; Pagarigan, B.; Mahmoudi, A.; Jackson, P.; Abbasian, M.; Muir, J.; Raheja, N.; Cathers, B. Crystal structures of PRK1 in complex with the clinical compounds lestaurtinib and tofacitinib reveal ligand induced conformational changes. *PloS One* **2014**, *9*, e103638.

(185) Bursulaya, B. D.; Totrov, M.; Abagyan, R.; Brooks, C. L. Comparative study of several algorithms for flexible ligand docking. *Journal of Computer-Aided Molecular Design* **2003**, *17*, 755–763.

(186) Davis, P. D.; Hill, C. H.; Keech, E.; Lawton, G.; Nixon, J. S.; Sedgwick, A. D.; Wadsworth, J.; Westmacott, D.; Wilkinson, S. E. Potent selective inhibitors of protein kinase C. *FEBS letters* **1989**, *259*, 61–63.

(187) Alessi, D. R. The protein kinase C inhibitors Ro 318220 and GF 109203X are equally potent inhibitors of MAPKAP kinase-1beta (Rsk-2) and p70 S6 kinase. *FEBS letters* **1997**, *402*, 121–123.

(188) Deak, M.; Clifton, A. D.; Lucocq, L. M.; Alessi, D. R. Mitogen- and stress-activated protein kinase-1 (MSK1) is directly activated by MAPK and SAPK2/p38, and may mediate activation of CREB. *The EMBO journal* **1998**, *17*, 4426–4441.

(189) Hers, I.; Tavaré, J. M.; Denton, R. M. The protein kinase C inhibitors bisindolylmaleimide I (GF 109203x) and IX (Ro 31-8220) are potent inhibitors of glycogen synthase kinase-3 activity. *FEBS letters* **1999**, *460*, 433–436.

(190) Davies, S. P.; Reddy, H.; Caivano, M.; Cohen, P. Specificity and mechanism of action of some commonly used protein kinase inhibitors. *The Biochemical Journal* **2000**, *351*, 95–105.

(191) Miknyoczki, S. J.; Dionne, C. A.; Klein-Szanto, A. J.; Ruggeri, B. A. The novel Trk receptor tyrosine kinase inhibitor CEP-701 (KT-5555) exhibits antitumor efficacy against human pancreatic carcinoma (Panc1) xenograft growth and in vivo invasiveness. *Annals of the New York Academy of Sciences* **1999**, *880*, 252–262.

(192) Weisel, K. C.; Yildirim, S.; Schweikle, E.; Kanz, L.; Möhle, R. Effect of FLT3 inhibition on normal hematopoietic progenitor cells. *Annals of the New York Academy of Sciences* **2007**, *1106*, 190–196.

(193) Hexner, E. O.; Serdikoff, C.; Jan, M.; Swider, C. R.; Robinson, C.; Yang, S.; Angeles, T.; Emerson, S. G.; Carroll, M.; Ruggeri, B.; Dobrzanski, P. Lestaurtinib (CEP701) is a JAK2 inhibitor that suppresses JAK2/STAT5 signaling and the proliferation of primary erythroid cells from patients with myeloproliferative disorders. *Blood* **2008**, *111*, 5663–5671.

(194) Smith, B. D.; Levis, M.; Beran, M.; Giles, F.; Kantarjian, H.; Berg, K.; Murphy, K. M.; Dauses, T.; Allebach, J.; Small, D. Single-agent CEP-701, a novel FLT3 inhibitor, shows biologic and clinical activity in patients with relapsed or refractory acute myeloid leukemia. *Blood* **2004**, *103*, 3669–3676.

(195) Knapper, S.; Burnett, A. K.; Littlewood, T.; Kell, W. J.; Agrawal, S.; Chopra, R.; Clark, R.; Levis, M. J.; Small, D. A phase 2 trial of the FLT3 inhibitor lestaurtinib (CEP701) as first-line treatment for older patients with acute myeloid leukemia not considered fit for intensive chemotherapy. *Blood* **2006**, *108*, 3262–3270.

(196) Levis, M. et al. Results from a randomized trial of salvage chemotherapy followed by lestaurtinib for patients with FLT3 mutant AML in first relapse. *Blood* **2011**, *117*, 3294–3301.

(197) Collins, C.; Carducci, M. A.; Eisenberger, M. A.; Isaacs, J. T.; Partin, A. W.; Pili, R.; Sinibaldi, V. J.; Walczak, J. S.; Denmeade, S. R. Preclinical and clinical studies with the multi-kinase inhibitor CEP-701 as treatment for prostate cancer demonstrate the inadequacy of PSA response as a primary endpoint. *Cancer Biology & Therapy* **2007**, *6*, 1360–1367.

(198) Chrencik, J. E. et al. Structural and thermodynamic characterization of the TYK2 and JAK3 kinase domains in complex with CP-690550 and CMP-6. *Journal of Molecular Biology* **2010**, *400*, 413–433.

(199) Fleischmann, R.; Kremer, J.; Cush, J.; Schulze-Koops, H.; Connell, C. A.; Bradley, J. D.; Gruben, D.; Wallenstein, G. V.; Zwillich, S. H.; Kanik, K. S.; Investigators, O. S. Placebo-controlled trial of tofacitinib monotherapy in rheumatoid arthritis. *The New England Journal of Medicine* **2012**, *367*, 495–507.

(200) Van Vollenhoven, R. F.; Fleischmann, R.; Cohen, S.; Lee, E. B.; García Meijide, J. A.; Wagner, S.; Forejtova, S.; Zwillich, S. H.; Gruben, D.; Koncz, T.; Wallenstein, G. V.; Krishnaswami, S.; Bradley, J. D.; Wilkinson, B.; Investigators, O. S. Tofacitinib or adalimumab versus placebo in rheumatoid arthritis. *The New England Journal of Medicine* **2012**, *367*, 508–519.

(201) Cada, D. J.; Demaris, K.; Levien, T. L.; Baker, D. E. Tofacitinib. *Hospital Pharmacy* **2013**, *48*, 413–424.

(202) Boy, M. G.; Wang, C.; Wilkinson, B. E.; Chow, V. F.; Clucas, A. T.; Krueger, J. G.; Gaweco, A. S.; Zwillich, S. H.; Changelian, P. S.; Chan, G. Double-blind, placebo-controlled, dose-escalation study to evaluate the pharmacologic effect of CP-690,550 in patients with psoriasis. *The Journal of Investigative Dermatology* **2009**, *129*, 2299–2302.

(203) Sandborn, W. J.; Ghosh, S.; Panes, J.; Vranic, I.; Wang, W.; Niezychowski, W.; Investigators, S. A. A phase 2 study of tofacitinib, an oral Janus kinase inhibitor, in patients with Crohn's disease. *Clinical Gastroenterology and Hepatology: The Official Clinical Practice Journal of the American Gastroenterological Association* **2014**, *12*, 1485–1493.e2.

(204) Busque, S. et al. Calcineurin-inhibitor-free immunosuppression based on the JAK inhibitor CP-690,550: a pilot study in de novo kidney allograft recipients. *American Journal of Transplantation: Official Journal of the American Society of Transplantation and the American Society of Transplant Surgeons* **2009**, *9*, 1936–1945.

(205) *Lead-seeking approaches*; Hayward, M. M., Bikker, J. A., Eds.; 5; Springer: 2010.

(206) Patel, R. Y.; Doerksen, R. J. Protein kinase-inhibitor database: structural variability of and inhibitor interactions with the protein kinase P-loop. *Journal of Proteome Research* **2010**, *9*, 4433–4442.

APPENDIX A

Supplementary Information

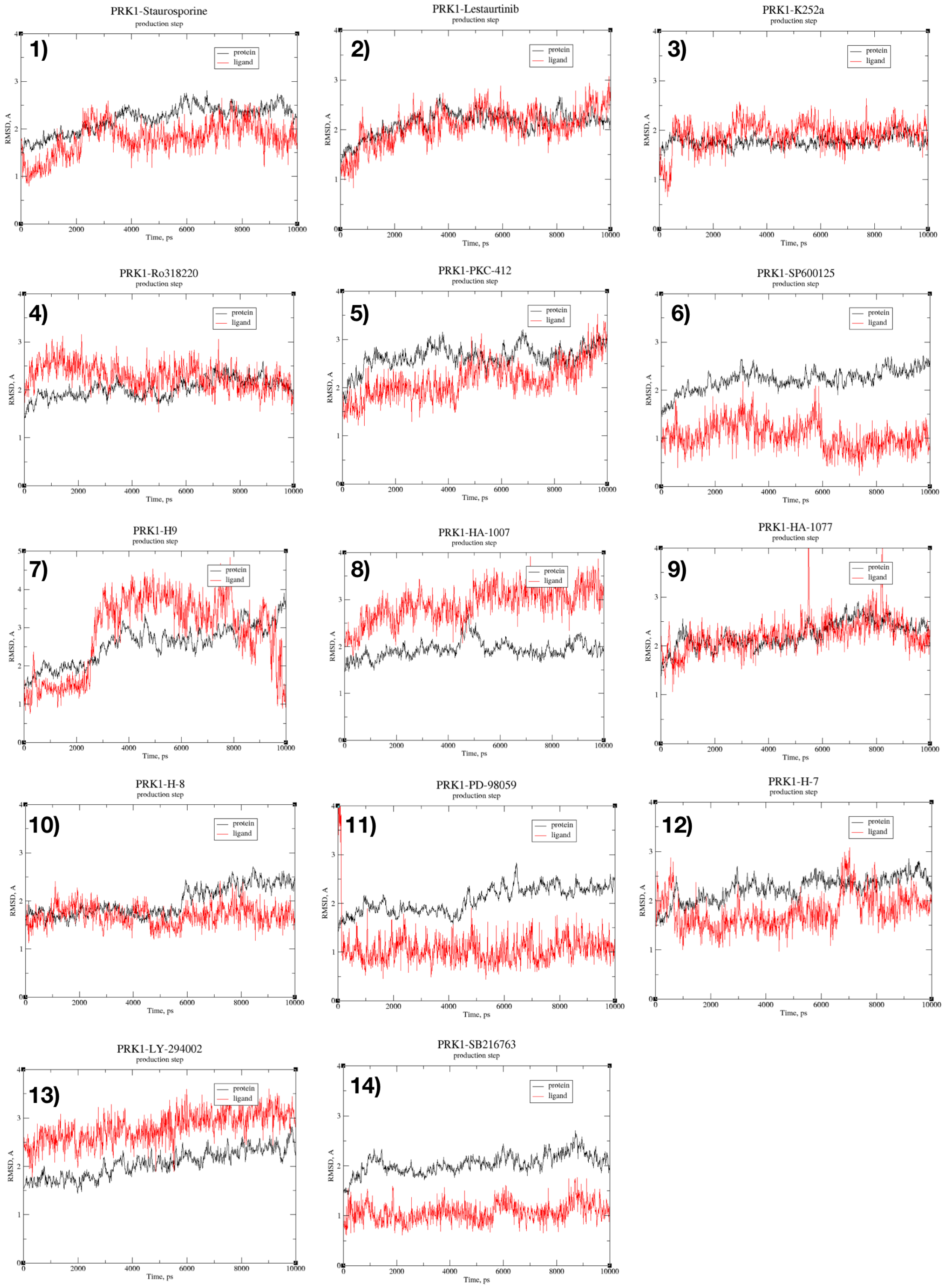


Figure A.1: Root mean square deviation plots (RMSD, Å, y-axis) for 14 Biomol compounds for which 10 ns MD simulation (Time, ps, x-axis) was performed (RMSD for ligand is shown in red and for protein in black).

Table A.1: Compounds from ChEMBL database, which were tested on PRK1 before the year 2011 and have reported activity. Duplicates were removed, e.g. when the same ChEMBL compound was tested in one assay, but has few different publication records; or when the same ChEMBL compound has few entries for residual activity.[166]

N	ChEMBL ID	Name	Standard type	Relation	Standard value	Standard units	Year
1	38380	Fasudil	IC_{50}	=	1700	nM	2003
2	388978	Staurosporine	K_d	=	1.3	nM	2008
3	603469	Lestaurtinib	K_d	=	5.3	nM	2009
4	608533	Midostaurin	K_d	=	9.3	nM	2008
5	1240703	CGP-52421	K_d	=	57	nM	2009
6	522892	Dovitinib	K_d	=	180	nM	2008
7	221959	Tofacitinib	K_d	=	200	nM	2008
8	91829	Ruboxistraurin	K_d	=	350	nM	2008
9	535	Sunitinib	K_d	=	710	nM	2008
10	572878	Tozasertib	K_d	=	1500	nM	2008
11	1721885		K_d	=	1900	nM	2008
12	428690	Alvocidib	K_d	=	3500	nM	2008
13	595373		RA[a]	=	68-97	%	2009
14	595143		RA	=	83-99	%	2009
15	604483		RA	=	91-99	%	2009
16	1234815		Inhibition	=	-9	%	2008
17	1241674		Inhibition	=	1	%	2008
18	1233882		Inhibition	=	2	%	2008
19	1230790		Inhibition	=	3	%	2008
20	1081312		Inhibition	=	5	%	2008
21	1233881		Inhibition	=	6	%	2008
22	1241578		Inhibition	=	6	%	2008
23	1077739		Inhibition	=	9	%	2009
24	521179	SR-3677	Inhibition	<	10	%	2008
25	1254209		Inhibition	=	18	%	2009
26	1170139		Inhibition	=	28	%	2010
27	1081678		Inhibition	<	50	%	2009
28	1078665		Inhibition	<	50	%	2009
29	1164180		Inhibition	<	50	%	2010
30	1163565		Inhibition	<	50	%	2010
31	1165499		Inhibition	<	50	%	2010
32	1164181		Inhibition	<	50	%	2010
33	1164265		Inhibition	<	50	%	2010
34	1082758		Inhibition	=	53	%	2010
35	1163566		Inhibition	=	82	%	2010
36	577784	BX-795	Inhibition	=	84	%	2009
37	1171523		Inhibition	=	85	%	2010
38	379300	A-443654	Inhibition	=	100	%	2009

[a] RA = residual activity

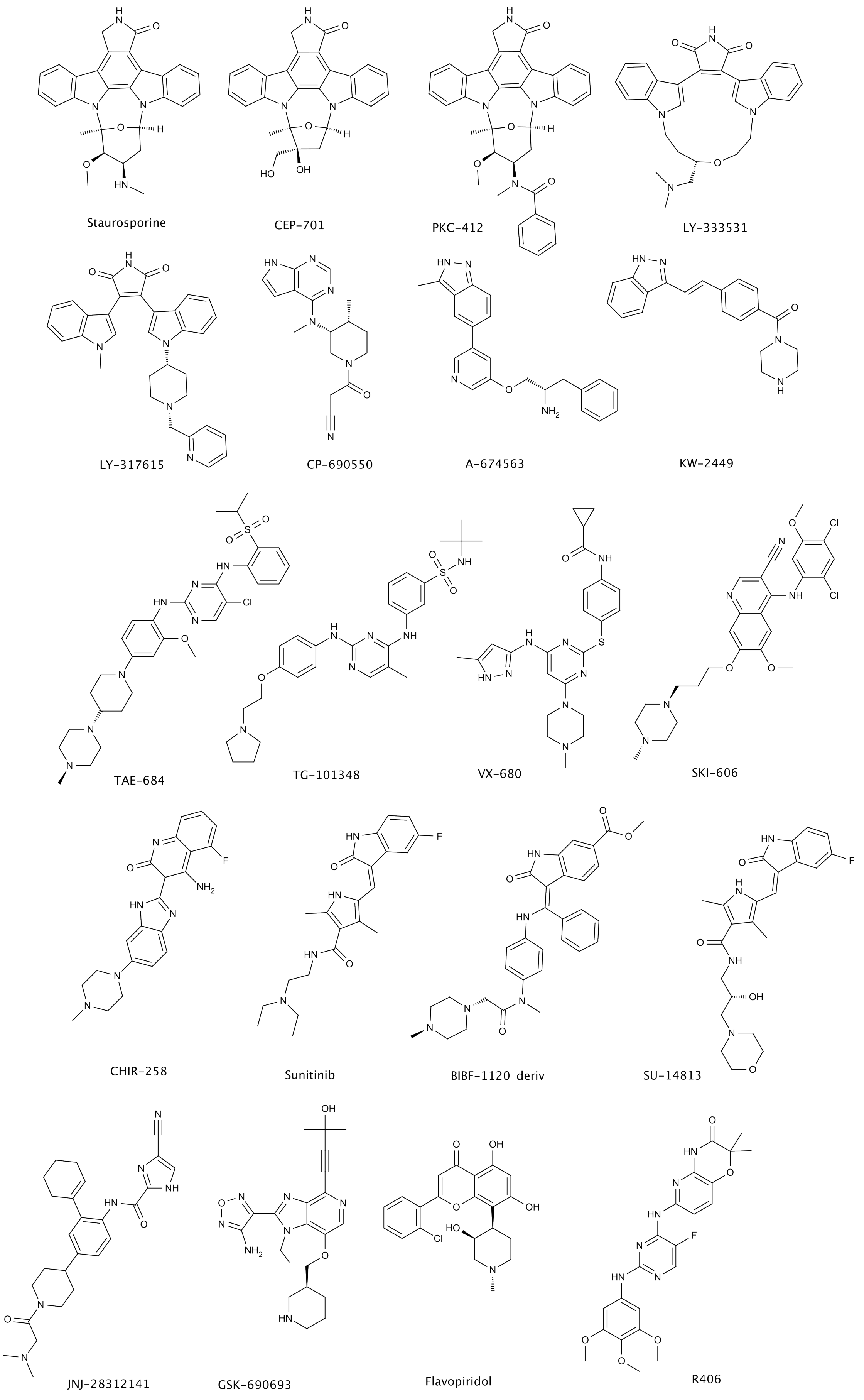

Staurosporine
CEP-701
PKC-412
LY-333531
LY-317615
CP-690550
A-674563
KW-2449
TAE-684
TG-101348
VX-680
SKI-606
CHIR-258
Sunitinib
BIBF-1120 deriv
SU-14813
JNJ-28312141
GSK-690693
Flavopiridol
R406

Figure A.2: PRK1 inhibitors identified by comprehensive assay set for protein kinases by Davis et al.[19] (test set, DS2).

Inna Slynko

Current Address
74 Avenue des Vosges
67000 Strasbourg
France

Contact Information
LinkedIn Profil
✉ inna.slynko@pharmazie.uni-halle.de
✉ slynko@unistra.fr
☎ +49 160 99651056 (mobile)

Date of birth: 25 October 1986
Birth Place: Cherkassy, Ukraine
Nationality: Ukrainian

Experience

06/2014 – 08/2014 **Bayer HealthCare AG** Wuppertal, DE

Computational Chemistry Department
- Internship. Supervisors: Andreas Göller, Jan Dreher
- Project: “Identifying target potential for chemistry driven libraries via optimization of chemical design”

Postgraduate Education

10/2010 – Present **Martin-Luther-University Halle (MLU)** Halle (Saale), DE

Medicinal Chemistry Department
- PhD student. Supervisor: Prof. Dr. habil. Wolfgang Sippl
- Research topic: “Structural analysis and computer based design of novel inhibitors for human protein-kinase-C-related enzyme PRK1”.
- From 2011 member of EUROPIN project (European Pharmacoinformatics Initiative - a joint PhD program between seven European institutions aiming to promote creativity, competitiveness of young scientists)
- From 2014 member of German Pharmaceutical Society

Higher Education

09/2007 – 06/2010 **National Technical University of Ukraine (NTUU “KPI”)** Kyiv, UA

Department of Organic Chemistry and Technology of Organic Substances
- Master’s Degree in Chemical Technology and Engineering
- Diploma work: “Synthesis and Antibacterial Activity Evaluation of Triclosan Derivatives as FAB Inhibitors”.

09/2008 – 08/2010 **Korea Institute of Science and Technology (KIST)** Seoul, KR

Center of Chemoinformatics
- Master of Science
- Double Diploma program between NTUU “KPI” and KIST

09/2003 – 07/2007 **National Technical University of Ukraine (NTUU “KPI”)** Kyiv, UA

Department of Organic Chemistry and Technology of Organic Substances
- Bachelor of Chemical Technology and Engineering
- Diploma project on benzylidenacetophenon production

Primary & Secondary Education

09/1993 – 06/2003 **Secondary School № 19** Cherkassy, UA

- Chemistry and biology class
- Member of Small Academy of Sciences of Ukraine in the city of Cherkassy (Chemistry Department, 2001-2003)

Research experience

06/2014 – 08/2014 **Internship,** Bayer HealthCare AG Wuppertal, DE

- Implementation of productive Pipeline Pilot workflow for analog searching, grouping, result visualization and reporting
- Setup of automated docking for important target family members

10/2010 – Present **PhD student,** MLU Halle (Saale), DE

Medicinal Chemistry Department, Prof. Dr. habil. Wolfgang Sippl

- Design of novel inhibitors that block the activity of Protein kinase C related kinase (PRK1) using computer-based approaches such as homology modeling, structure- and ligand-based virtual screening, binding free energy calculations
- Practical lessons in computer-aided drug design for graduate students (irregular, partial responsibility)

09/2008 –07/2009 **Research Assistant,** KIST Seoul, KR

Center of Chemoinformatics, Prof. Gyochang Keum

- Synthesis, determination of structure of triclosan derivatives as fatty acid biosynthesis inhibitors and evaluation of their antibacterial activity

09/2007 – 06/2008 **Lab Assistant,** Construction Department “Storm” Kyiv, UA

- Dealt mainly with fuel additives and biodiesel modification
- Distillation of petroleum products for composition analysis

09/2006 – 05/2007 **Lab Assistant,** Institute of Bioorganic Chemistry and Petroleum Chemistry Kyiv, UA

Department of Chemistry of Bioactive Nitrogencontaining Heterocyclic Bases, Dr. Sci. Prof. V.S. Brovarets

- Synthesis of heterocyclic compounds under supervision of PhD student

Additional Skills

Computer Skills

- OS: Linux, Windows, MacOS
- Applications: computational chemistry software (MOE, Modeller, Gold, Schroedinger, ParaDocks, Amber, LigandScout, MarvinSketch, SciFinder, OpenEye, Chimera, PyMol, Pipeline Pilot)
- basic bash scripting
- MS Office, LaTeX

Methodologies

- homology modeling
- virtual screening (structure- and ligand-based, e.g. docking, substructure and similarity search, pharmacophore-based VS)
- molecular dynamics
- binding free energy calculation
- QSAR model development

Languages

- English – fluent
- Ukrainian, Russian – native
- German – upper intermediate (B2-C1)
- Korean – elementary level

Personal Skills

Responsible and hard-working person, organized, open-minded, punctual, collaborative, strong communication and presentation skills, excellent analytical and problem-solving skills, able to work independently and as a part of a team, like to travel

Place, Date Sign

Halle, 13 Mar 2015

Publications, Abstracts and Conference Talks

Publications

- Mahmoud, K. A.; Krug M.; Wersig,T.; Slynko, I.; Schächtele, C.; Totzke, F.; Sippl, W. and Hilgeroth A. **Discovery of 4-anilino α-carbolines as novel Brk inhibitors.** *Bioorg. Med. Chem. Lett.* **2014,** *24 (8):* 1948–1951.
- Wichapong, K.; Rohe, A.; Platzer, C.; Slynko, I.; Erdmann, F.; Schmidt, M. and Sippl. W. **Application of Docking and QM/MM-GBSA Rescoring to Screen for Novel Myt1 Kinase Inhibitors.** *J. Chem. Inf. Mod.* **2014,** *54 (3):* 881–893.
- Mahmoud, K. A.; Wersig,T.; Slynko, I.; Totzke, F.; Schächtele, C.; Oelze, M.; Sippl, W.; Ritter, C. and Hilgeroth A. **Novel inhibitors of breast cancer relevant kinases Brk and HER2.** *Med. Chem. Comm.* **2014**, *5:* 659–664.
- Slynko, I.; Scharfe, M.; Rumpf, T.; Eib, J.; Metzger, E.; Schüle, R.; Jung, M.; Sippl, W. **Virtual Screening of PRK1 Inhibitors: Ensemble Docking, Rescoring Using Binding Free Energy Calculation and QSAR Model Development.** *J. Chem. Inf. Mod.* **2014,** *54 (1)*: *138–150*.
- Köhler, J.; Erlenkamp, G.; Eberlin, A.; Rumpf, T.; Slynko, I.; Metzger, E.; Schüle, R.; Sippl, W.; Jung, M. **Lestaurtinib inhibits histone phosphorylation and androgen-dependent gene expression in prostate cancer cells.** *PlosOne* **2012**, *7 (4)*: e34973.

Conference Talks / Presentations

- Slynko, I. **Virtual screening of PRK1 inhibitors: ensemble docking, rescoring using binding free energy calculation and QSAR model development.** DPhG-Doktorandentagung 2014. 10-12.03.2014, Bayer HealthCare, Wuppertal, Germany.
- Slynko, I. **Application of MM-PB(GB)SA and QM/MM-GBSA rescoring approaches for predicting biological activities of novel PRK1 kinase inhibitors.** 7th Summer School on Drug Design. 15-22.09.2013, Vienna, Austria.
- Slynko, I.; Erlenkamp, G.; Köhler, J.; Rumpf, T.; Metzger, E.; Schüle, R.; Jung, M.; Sippl, W. **Identification of PRK1 kinase inhibitors by virtual screening and binding free energy calculations.** 19th EuroQSAR Symposium. 26-30.08.2012, Vienna, Austria.
- Slynko, I.; Bhattarai, D.; Keum, G. C. **Synthesis and antibacterial activity evaluation of triclosan derivatives as FAB inhibitors.** III International Chemistry and Chemical Technology Conference for Students, Postgraduates and Young Scientists. 21-23.04.2010, Kiev, Ukraine.

Publications, Abstracts and Conference Talks

Poster Abstracts / Posters

- Slynko, I.; Rumpf, T.; Metzger, E.; Schüle, R.; Jung, M.; Sippl, W. **Targeting posttranslational histone modifications – search for novel PRK1 inhibitors.** Epigenetics: From Bench To Bedside COST Conference. 05-08.05.2014, Athens, Greece.
- Slynko, I.; Rumpf, T.; Metzger, E.; Schüle, R.; Jung, M.; Sippl, W. **Rescoring of docking poses using MM-GBSA and QM/MM-GBSA – Predicting biological activities of novel PRK1 kinase inhibitors.** Workshop "Innovative Approaches to Computational Drug Discovery". 01-04.10.2013, Lausanne, Switzerland.
- Slynko, I.; Rumpf, T.; Metzger, E.; Schüle, R.; Jung, M.; Sippl, W. **Combination of ligand- and structure-based approaches for searching novel PRK1 kinase inhibitors.** EuroCUP VI, OpenEye's 6th annual European science meeting. 15-17.05.2013, Santpoort, Netherlands.
- Slynko, I.; Erlenkamp, G.; Köhler, J.; Rumpf, T.; Metzger, E.; Schüle, R.; Jung, M.; Sippl, W. **Identification of PRK1 kinase inhibitors by virtual screening and binding free energy calculations.** 6th Summer School Medicinal Chemistry. 26-28.09.2012, Regensburg, Germany.
- Slynko, I.; Erlenkamp, G.; Köhler, J.; Rumpf, T.; Metzger, E.; Schüle, R.; Jung, M.; Sippl, W. **Identification of PRK1 kinase inhibitors by virtual screening and binding free energy calculations.** 19th EuroQSAR Symposium. 26-30.08.2012, Vienna, Austria. *Poster selected for the oral presentation.*
- Slynko, I.; Erlenkamp, G.; Köhler, J.; Rumpf, T.; Metzger, E.; Schüle, R.; Jung, M.; Sippl, W. **Homology modeling and docking studies of PRK1 kinase.** 25th Molecular Modeling Workshop. 04-06.04.2011, Erlangen, Germany. *Poster award.*
- Slynko, I.; Bhattarai, D.; Keum, G. C. **Synthesis and antibacterial activity evaluation of triclosan derivatives as FAB inhibitors.** III International Chemistry and Chemical Technology Conference for Students, Postgraduates and Young Scientists. 21-23.04.2010, Kiev, Ukraine.
- Slynko, I.; Bhattarai, D.; Pae, A. N.; Keum, G. C. **Synthesis of triclosan derivatives as antibacterial agents.** 105th Spring Meeting of the Korean Chemical Society. 29-30.04.2010, Incheon, South Korea.

Erklärung (Declaration of Authorship)

Hiermit erkläre ich, dass ich die vorliegende Dissertationsschrift selbständig und ohne fremde Hilfe angefertigt, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und die aus ihnen wörtlich oder inhaltlich entnommen Stellen als solche kenntlich gemacht habe. Die Arbeit wurde ausschließlich der Mathematisch-Naturwissenschaftlichen Fakultät der Martin-Luther-Universität Halle-Wittenberg vorgelegt und an keiner anderen Universität oder Hochschule weder im In- oder Ausland zur Erlangung des Doktorgrades eingereicht.

Halle (Saale), den 16.03.2015

Inna Slynko