

Application of Computational Methods for Kinase Small Molecule Drug Discovery and Design

DISSERTATION

zur Erlangung des akademischen Grades

Doctor rerum naturalium (Dr. rer. nat.)

der Naturwissenschaftlichen Fakultät I - Biowissenschaften

der Martin-Luther-Universität Halle-Wittenberg

von

Abdulkarim Najjar

geboren am 01. Jan. 1987

Gutachter:

1. Prof. Dr. Wolfgang Sippl
2. Prof. Dr. Mike Schutkowski
3. Prof. Dr. Johannes Kirchmair

Datum der Disputation: 12.12.2018

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere gratitude to my supervisor Prof. Dr. Wolfgang Sippl for the continuous support of my projects. His support to join his research group came in the very critical time of my career, which allowed me to overcome the challenges. Moreover, his guidance helped me to improve my knowledge and collect more experiences not only in science but also in the real life.

Besides my supervisor, I would like to thank our collaborators Prof. Dr. Manfred Jung, Prof. Dr. Andreas Hilgeroth, and their research groups for their contributions to our projects.

Collective and individual acknowledgements are also owed to my colleagues, especially to Dr. Dina Robaa, for many valuable encouragements and discussions. A special thank goes to PD. Dr. Matthias Schmidt for his constructive comments on the projects. His advice and experience have been tremendously helpful to the progress of my projects.

My sincere thanks also go to Prof. Dr. Johannes Kirchmair, who provided me with an opportunity to join his team as an intern and led me working on diverse exciting projects. He offered me valuable experiences during the internship allowing me to expand my knowledge.

Lastly, I would like to thank my family and friends for all their love and encouragement. For my parents who raised me with a love of science and supported me in all my pursuits. For the presence of my brothers and sisters. And most of all for my loving, supportive, encouraging, and patient wife Abir whose faithful support during my Ph.D. is so appreciated. For Sam, my son, who joined and added meaning to my life. Thank you all.

“If we knew what it was we were doing, it would not be called research, would it?”

Albert Einstein

ABSTRACT

Molecular modeling and computer-aided drug design (CADD) have been extensively used as innovative tools in medicinal chemistry to accelerate drug discovery. Hit identifications, hit-to-lead optimizations, ligand binding analysis and lead optimizations are applied in the early stages of the drug discovery pipeline. Moreover, Computational Drug Design is a widely used approach in pharmaceutical and biochemical industry.

The current work started by a combination of *in silico* and *in vitro* screening in order to identify novel protein kinase C-related kinase 1 (PRK1) and membrane-associated tyrosine/threonine (PKMYT1) inhibitors. Several approaches of structure- and ligand-based virtual screening were carried out to screen the kinase inhibitor libraries from GlaxoSmithKline (GSK dataset I and II), and Selleckchem. The employed computational approaches combined homology modeling, ensemble docking, binding free energy calculations, molecular dynamics simulations, protein-ligand interaction analysis, and applications of quantitative structure–activity relationship (QSAR) models including 2D and/or 3D molecular descriptors. As a consequence, several hits were identified as PRK1 inhibitors active in the low-nanomolar range. The identified inhibitors represent valuable tools to confirm the biological roles of PRK1, and to be leads for lead optimization. Meanwhile, few hits were found to inhibit PKMYT1 in the nanomolar range, and several chemical scaffolds were used as starting point to develop novel PKMYT1 inhibitors. Subsequently, novel PKMYT1 inhibitors were synthesized and biologically evaluated. The derived biological results provided the basis for further chemical optimizations and for further analysis of PKMYT1 as target for cancer therapy.

In further projects, molecular docking and structural analysis were implemented to explore the potential binding mode of novel epidermal growth factor receptor (EGFR), insulin-like growth factor 1 receptor (IGF-1R), platelet-derived growth factor receptor beta-type (PDGFR- β), and vascular endothelial growth factor receptor 2 (VEDFR2) kinase inhibitors and to rationalize the observed biological activities. This led to the development of dual kinase inhibitors active in the low-nanomolar range that were additionally tested for their cancer cell growth inhibition.

In conclusion, the applied molecular modeling methods showed a proven ability to guide, and rationalize the development of novel kinase inhibitors, which have to be further evaluated on their potential as candidates for cancer treatment.

Keywords: kinase, PKMYT1, PRK1, EGFR, IGF-1R, PDGFR- β , VEDFR2, kinase inhibitors, drug discovery, computer-aided drug design, molecular docking, binding free energy calculations, molecular dynamics simulation, homology modeling, QSAR, virtual screening, anticancer.

KURZFASSUNG

Molekulare Modellierung und computergestütztes Wirkstoffdesign haben sich in der medizinischen Chemie als innovative und essenzielle Werkzeuge zur Beschleunigung der Wirkstoffforschung und -entwicklung bewährt. Hit Identifikationen, „Hit-to-Lead“-Optimierung, Vorhersagen von Bindungsmodi und Optimierungen von Leitstrukturen werden in der frühen Phase der Arzneimittelentwicklung eingesetzt. Darüber hinaus werden computergestützte Wirkstoffdesign-Verfahren oft in der Pharmaindustrie eingesetzt.

Die aktuelle Arbeit begann mit einer Kombination aus *in silico* und *in vitro* Screening, um neue PRK1 und PKMYT1 Inhibitoren zu identifizieren. Mehrere Ansätze des virtuellen Screenings auf Struktur- und Ligandenbasis wurden durchgeführt, um fokussierte kommerzielle chemische Bibliotheken (GlaxoSmithKline; GSK-Datensatz I und II) und Selleckchem nach potentiellen Inhibitoren zu durchsuchen. Die computergestützten Ansätze kombinieren Homologie-Modellierung, Ensemble-Docking, Berechnungen der freien Bindungsenergie, Molekulardynamik-Simulationen, „Fingerabdrücke“ von Protein-Ligand-Wechselwirkungen und Anwendungen quantitativer Struktur-Aktivitäts-Beziehungs-Modelle (QSAR) einschließlich 2D / 3D-Moleküldeskriptoren zur Vorhersage der biologischen Aktivitäten. Folglich wurden einige Hits identifiziert, die PRK1 im niedrigen nanomolaren Bereich inhibieren konnten. Die identifizierten Inhibitoren stellen wertvolle Werkzeuge dar, um die biologische Rolle von PRK1 zu erforschen und dienen als Leitstrukturen für weitere medizinalchemischen Optimierungen. Weiterhin, wurden einige Hits identifiziert, die PKMYT1 im nanomolaren Bereich hemmen konnten. Verschiedene chemische Grundgerüste wurden daraufhin verwendet, um neuartige PKMYT1 Inhibitoren zu entwickeln. Somit wurde eine Reihe neuer PKMYT1-Inhibitoren synthetisiert und biologisch bewertet. Die erhaltenen Ergebnisse liefern die Grundlage für eine weitere chemische Optimierung der PKMYT1-Inhibitoren und für eine weitere Analyse von PKMYT1 als Ziel für die Krebstherapie.

Weiterhin wurde Docking eingesetzt, um die möglichen Bindungsmodi neuer EGFR-, IGF-1R-, PDGFR- β , VEDFR2-Kinase-Inhibitoren zu untersuchen und die beobachteten biologischen Aktivitäten zu erklären. Dies ermöglichte die Entwicklung neuer Dual-Kinase-Inhibitoren, die eine inhibitorische Aktivität im niedrigen nanomolaren Bereich zeigten und eine Krebszellenwachstumshemmung aufwiesen.

Schließlich zeigten Molekulare Modellierung und computergestütztes Wirkstoffdesign eine nachgewiesene Fähigkeit, die Entwicklung neuer Kinase-Inhibitoren zu steuern und zu rationalisieren. Diese Kinase-Inhibitoren könnten möglicherweise Kandidaten für die Behandlung von verschiedenen Tumorarten sein.

Schlagwörter: Kinase, PKMYT1, PRK1, EGFR, IGF-1R, PDGFR- β , VEDFR2, Kinaseinhibitoren, Wirkstoffforschung, computergestütztes Wirkstoffdesign, molekulares Docking, freie Bindungsenergie, Moleküldynamiksimulation, Homologiemodellierung, QSAR, virtuelles Screening, Tumorthapeutika.

CONTENTS

ACKNOWLEDGMENTS	III
ABSTRACT	V
KURZFASSUNG	VI
LIST OF FIGURES	X
LIST OF ABBREVIATIONS	XI
INTRODUCTION	1
1.1. Druggable genome.....	2
1.2. Kinases as drug targets.....	3
1.2.1. Kinases.....	4
1.2.2. Kinase 3D structure.....	7
1.2.3. Kinase inhibition.....	11
1.2.4. Classification of small molecule kinase inhibitors.....	13
1.2.5. Lipid kinase inhibitors.....	14
1.3. Limitations, challenges, and future perspectives in kinase drug discovery.....	15
AIM OF THE WORK	16
METHODS AND MATERIALS	18
2.1. Computer-aided drug design (CADD).....	19
2.1.1. Fragment-based drug design (FBDD).....	19
2.1.2. Virtual screening.....	20
2.2. Homology modeling.....	21
2.3. Molecular docking.....	21
2.4. Molecular Dynamics (MD) simulations.....	22
2.5. BFE calculations.....	23
2.6. <i>In vitro</i> testing.....	25
RESULTS AND DISCUSSIONS	27
3.1. Regulation of G2/M Transition by Inhibition of WEE1 and PKMYT1 Kinases.....	28

3.2. Identification of PKMYT1 Inhibitors by Screening the GSK Published Protein Kinase Inhibitor Set I and II	29
3.3. Computer-aided Design, Synthesis and Biological Characterization of Novel Inhibitors for PKMYT1 Kinase.....	30
3.4. Identification of Highly Potent Protein KinaseC-Related Kinase1 Inhibitors by Virtual Screening, Binding Free Energy Rescoring, and invitro Testing.	31
3.5. Application of Computer Modeling to Drug Discovery: Case Study of PRK1 Kinase Inhibitors as Potential Drugs in Prostate Cancer Treatment.	32
3.6. Discovery of Dually Acting Small-molecule Inhibitors of Cancer-Resistance Relevant Receptor Tyrosine Kinases EGFR and IGF-1R.	33
3.7. Discovery of Novel Dual Inhibitors of Receptor Tyrosine Kinases EGFR and IGF-1R.	34
3.8. Discovery of Novel Substituted Benzo-anellated 4-benzylamino Pyrrolopyrimidines as Dual EGFR and VEGFR2 Inhibitors.	35
3.9. Discovery of Novel Dual Inhibitors of Receptor Tyrosine Kinases EGFR and PDGFR-beta Related to Anticancer Drug Resistance.	36
SUMMARY	37
4.1. Structural analysis of WEE family kinase	37
4.2. Hit identification and lead optimization for PKMYT1	38
4.3. Virtual screening of commercial data sets on PRK1	41
4.4. Discovery of novel growth factors receptors inhibitors	43
CONCLUSIONS	46
REFERENCES	48
APPENDIX	56
5.1A. FDA-approved protein kinase inhibitors	56
5.2A. List of transferases: Phosphorous-containing groups (kinase) (EC 2.7.-.-).....	60
5.3A. Full-text publications	62

LIST OF FIGURES

Figure 1. Drug-target families	2
Figure 2. Bar chart represents number of drugs modulating four privileged families	3
Figure 3. Human kinome classification	6
Figure 4. Overall structure of the active kinase domain	8
Figure 5. Catalytic site of the kinase domain	10
Figure 6. Activation states of the kinase domain.....	11
Figure 7. Categories of established and clinical trial agents.	12
Figure 8. Scheme of the ATP-binding pocket of the kinase domain.	14
Figure 9. Superposition of the ATP-bindings pocket of WEE1 and PKMYT1	38
Figure 10. Scheme of the process to identify PKMYT1 inhibitors.	39
Figure 11. Determined binding modes of the identified PKMYT1 inhibitors	40
Figure 12. Binding mode of CAO2M within the PKMYT1 binding pocket.....	41
Figure 13. Identified compounds from GSK PKIS set active on PRK1.	42
Figure 14. Predicted binding mode for three of the most active PRK1 inhibitors	43
Figure 15. Putative binding modes of the most potent inhibitor (13b) in EGFR	44
Figure 16. Chemical structures of the identified dual-inhibitors of the growth factors receptors.	45

LIST OF ABBREVIATIONS

2D	Two-dimensional
3D	Three-dimensional
AGC	Protein kinase A, G and C
A-loop	Activation loop
AM1	Austin model 1
Amber10:EHT	Amber10: extended huckel theory
APE-motif	Alanine-proline-glutamic
ATP	Adenosine-5-triphosphate
BFE	Binding free energy
BLAST	Basic local alignment search tool
CADD	Computer-aided drug design
CAMK	Calmodulin/Calcium regulated kinases
CK1	Casein kinases
CMGC	CDK, MAPK, GSK3 and CLK
CML	Chronic myeloid leukemia
CNS	Central nervous system
DFG-motif	Aspartic-phenylalanine-glycine motif
DMSO	Dimethyl sulfoxide
DOPE	Discrete optimized protein energy
EGFR	Epidermal growth factor receptor
FBDD	Fragment-based drug design
FDA	Food and drug administration
GAFF	General AMBER force field
GB	Generalized Born
GHI domain	Helixes G, H, and I
GPCRs	G-protein-coupled receptors
GSK	GlaxoSmithKline.
HM	Homology modeling
HTS	High throughput Screening
IGF-1R	Insulin-like growth factor 1 receptor
LE	Ligand efficiency
LLE	Ligand lipophilicity efficiency
MACCS	Molecular ACCess system
MD	Molecular dynamics

MM	Molecular mechanics
MM/QM-GB/SA	Quantum mechanics/molecular mechanics Poisson-Boltzmann (Generalized Born) surface area
MM-GB/SA	Molecular mechanics generalized Born surface area
MM-PB(GB)/SA	Molecular mechanics Poisson-Boltzmann (generalized Born) surface area
MOE	Molecular operation environment system
NMR	Nuclear magnetic resonance spectroscopy
NRot	Number of rotatable bonds
PB	Poisson Boltzmann
PDB	Protein Data Bank
PDGFR	Platelet-derived growth factor receptors
PKMYT1	Protein kinase, membrane associated tyrosine/threonine 1
P-loop, G-loop	Glycine-rich loop
PM3	Parameterized model number 3
PRK1	Protein kinase C-related kinase 1
QM	Quantum mechanics
QM/MM	Quantum mechanics/molecular mechanics
QSAR	Quantitative structure–activity relationship
RCSB PDB	Research collaboratory for structural bioinformatics Protein Data Bank
RGC	Receptor guanylate cyclase
SA	Surface area
SASA	Solvent-accessible surface area
siRNAs	Small interfering RNAs
STE	Sterile kinases
TK	Tyrosine kinase
TKL	Tyrosine kinase-like
VEGFR2	Vascular endothelial growth factor receptor 2
VS	Virtual screening
Wee1	Wee1 protein kinase
Y/HRD-motif	Tyrosine/histidine-arginine-aspartic-motif
α C	C-helix

INTRODUCTION

The introduction gives a general overview of the druggable genome and highlights the importance of the human kinome as drug target. The overall structure of the conserved kinase domain besides the ATP-binding pocket as an important pocket for designing kinase inhibitors is described. In addition, the role of conformational changes in the binding pocket is discussed. Small molecules as drug candidates for the kinase inhibition are discussed revealing several classes of recently approved kinase inhibitors. Finally, a general evaluation of the limitation and challenging face of the kinase drug discovery is introduced in this chapter.

1.1. Druggable genome

Druggable genome is a term firstly presented in 2002 by Hopkins and Groom [1]. The concept of druggable genome refers to a portion of the human genome, which expresses proteins able to bind drug-like molecules. Previously, an analysis of 483 targets revealed that there could be 5,000-10,000 potential targets which were estimated based on the number of disease-related genes, without focusing on the properties of the drugs that define those targets [2]. However, a druggable target must show specific characteristics mainly to be directly involved in a disorder, locally expressed in tissues or organs and functionally and structurally characterized [3, 4].

Hopkins et al. systematically mined the human genome for putative drug targets amenable to drug-like high affinity interactions ($K_i < 10 \mu\text{M}$) with small molecules that are compliant with the Lipinski's rules [5]. As a result, they obtained a list of about 400 non-redundant proteins fulfilling these requirements [1]. Moreover, they found a large percentage of these targets clustered into target families, such as G-protein-coupled receptors (GPCRs), serine, threonine and tyrosine kinases, serine, cysteine, aspartic acid and metalloproteases, ion channels and nuclear hormone receptors. Thus, approximately 3,050 putative protein targets were theoretically extracted by the systematic extrapolation within the corresponding families (**Fig. 1**).

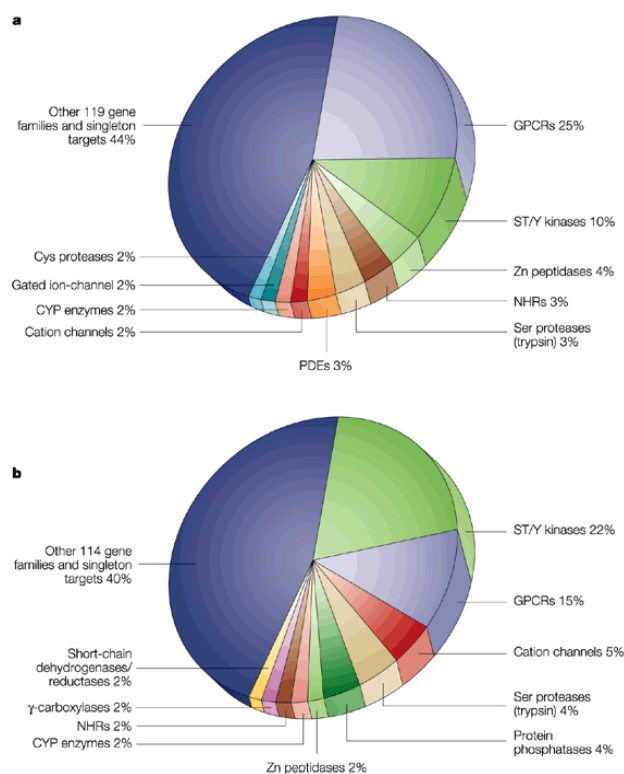


Figure 1. Drug-target families. Gene-family distribution of (a) the targets of current rule of five-complaint experimental and marketed drugs, and (b) the druggable genome. Adapted from Hopkins et al. [1].

A particular attention has been recently given to proteases, kinases and G-protein-coupled receptors (GPCRs) [6]. A recent analysis of the comprehensive map of molecular drug targets was performed in 2017 by Santos et al. covering 893 human and pathogen-derived biomolecules through which 1,578 US FDA-approved drugs act [7]. This analysis at the target family level revealed a higher number of recently approved drugs that modulate kinases compared to the number of recently approved drugs targeting either nuclear receptors or ion channels (**Fig. 2**). From 2011 to 2015, twenty kinase inhibitors were approved by the FDA (**Table 1A**, see **Appendix**). However, due to the broad poly-pharmacology typical of small-molecule kinase inhibitors, there is a large difference between the drug and the target fractions for kinases [7, 8].

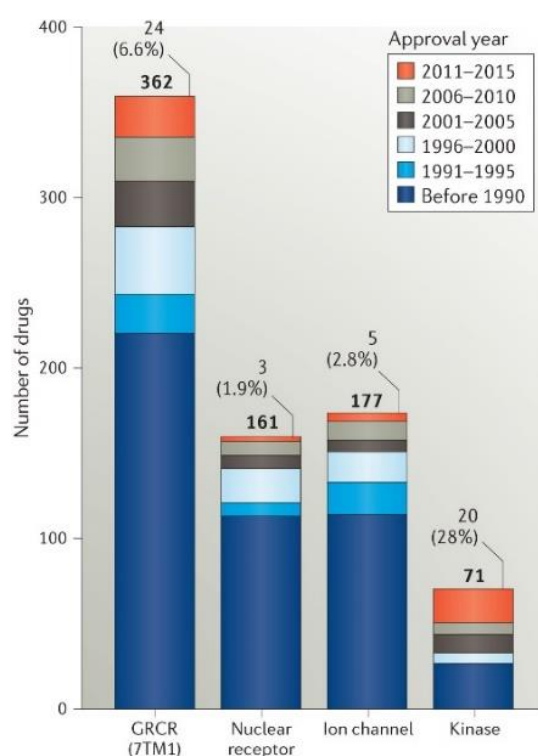


Figure 2. Bar chart represents number of drugs modulating four privileged target families, distributed per year of first approval. The total approved drug number is displayed on top of each bar (bold), together with the number and percentage of drugs approved from 2011 to 2015. Adapted from Santos et al. [7].

1.2. Kinases as drug targets

As can be seen in Fig1, the majority of targets in the druggable genome are related to kinases (**Fig. 1**). Moreover, Rask-Andersen et al. reported in 2014 that the kinase family is the largest category of the potentially novel clinical trial drug targets [9]. In total, 145 unique kinases were identified as the targets of established and clinical trial agents. Interestingly, approved pharmaceuticals do not target 93 of these potentially novel kinases, hence, making them the largest target-family in the clinical trials [9].

Kinases perform key regulatory roles in signal transduction modules that regulate apoptosis, cell cycle progression, cytoskeletal rearrangement, differentiation, development, the immune response, nervous system function, and transcription [10]. Kinases are considered one of the most promising drug targets for the treatment of several diseases [11], which are associated with the dysfunction of several kinases. The vast majority of kinases have been investigated for the treatment of cancer [12-14]. Furthermore, kinases have also been targeted in the treatment of inflammatory diseases [15, 16], central nervous system (CNS) disorders (such as Alzheimer's disease and Parkinson's disease) [17-19], cardiovascular diseases [20], and complications of diabetes [21]. Thus, the regulation of kinase activity is an important therapeutic strategy for the treatment of the above mentioned diseases.

1.2.1. Kinases

The kinase family is one of the largest family including 900 genes encoded within the human genome [10, 22]. In total, 539 kinases genes have been identified so far [23]. Protein kinases are enzymes that catalyze the transfer of a phosphate group from γ -phosphate of adenosine-5'-triphosphate (ATP) to protein substrates. The attached phosphate binds covalently to serine, threonine, or tyrosine (**Table 2A**, see **Appendix**). Based on the resource of the phosphorylated OH group, kinases are mainly classified as protein-serine/threonine kinases or protein-tyrosine kinases [24]. The protein-serine/threonine kinases are the largest sub-family within the kinase family. The majority of the protein-tyrosine kinases are tyrosine-kinase-receptors which are composed of an extracellular, a transmembrane, and an intracellular domain. The remaining protein-tyrosine kinases are non-receptors occurring intracellularly [24]. However, a small group of kinases, which is called dual specificity kinases e.g. MEK1 and MEK2, catalyzes the phosphorylation of both threonine and tyrosine on known substrates [24].

In 2002, Manning et al. identified 518 protein kinases and classified them in two major categories: typical kinases including 478 kinase and atypical kinases containing 40 kinase [10]. Their classification extended the original kinase classification introduced by Hanks and Hunter in 1995 [25], expanding it from five broad kinase groups, 44 families, and 51 sub-families to a total of nine groups, 90 families and 145 sub-families [10]. Kinases were initially classified based on sequence comparison of their catalytic domain, the knowledge of the sequence similarity and the domain structure outside of their catalytic domains, their known biological functions, and a similar classification of the yeast, worm, and fly kinomes [26]. **Figure 3** shows the human kinome and its classification. The nine kinase groups are generally as follows [10]:

- AGC** Protein kinase A, G and C (63 human kinases).
- CMGC** CDK, MAPK, GSK3 and CLK (61 human kinases).
- CAMK** Calmodulin/Calcium regulated kinases CAMK (74 human kinases).

CK1	Casein Kinases 1 (12 human kinases).
STE	Homologs of the yeast sterile kinases STE7, STE11 and STE20 (47 human kinases).
TK	Tyrosine kinase (90 human kinases).
TKL	Tyrosine kinase-like (43 human kinases).
RGC	Receptor Guanylate Cyclase (5 human kinases).
Other	Containing conserved sequences that have not yet been classified as domains and whose functions are unknown (83 human kinases).

The 40 atypical kinases show no sequence similarity to the typical kinases but they were reported to have biochemical kinase activity plus similar structural fold to the typical kinases. The atypical kinases were classified into 13 families in the human kinome [10].

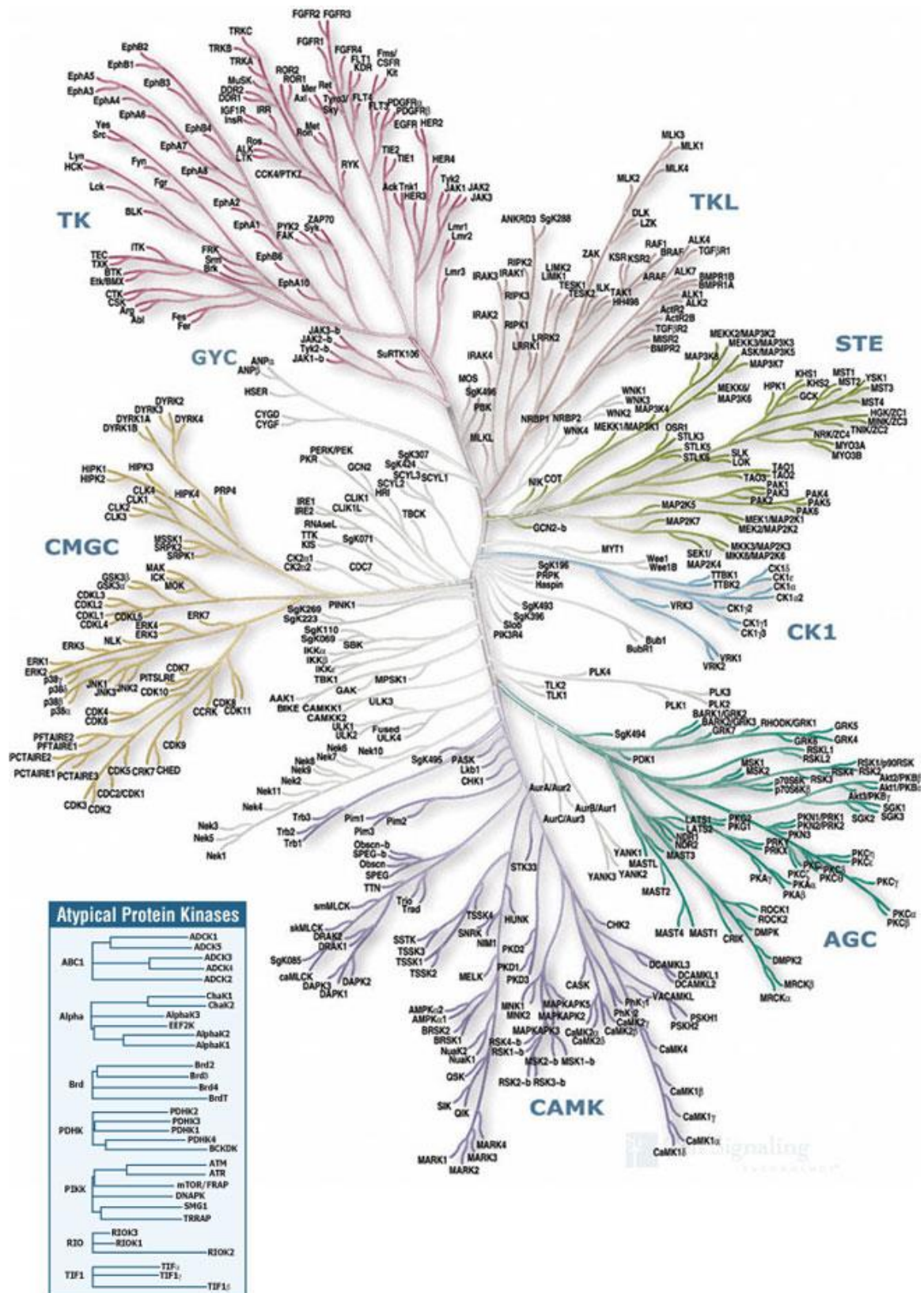


Figure 3. Human kinome classification. Calmodulin/Calcium-regulated kinases (CAMK) A, G and C protein kinases (AGC), Casein kinases 1 (CK1), Sterile kinases (STE), Tyrosine kinases (TK), Tyrosine kinases-like (TKL) and CDK, MAPK, GSK3 and CLK kinases (CMGC), Receptor Guanylate Cyclase (RGC) includes GYC kinases. Adapted from Manning et al. [10].

1.2.2. Kinase 3D structure

Despite the diversity in the primary amino acid sequences of human kinases, they share a great degree of similarity in their 3D structures, mainly in the catalytic domain where the ATP-binding pocket is located [27]. The overall structure of the kinase domain (ca. 300 residues) is conserved with 10 key residues mediating the core function of the catalytic domain [10, 28, 29]. Meanwhile, the remaining structural elements serve as either regulatory or targeting modules [30, 31]. The kinase domain consists of a small, dominantly β -stranded N-lobe and a larger α -helical C-lobe, which are connected by a short hinge region (**Figure 4**). The ATP binding pocket is located between the N- and C- terminal lobes of the kinase domain, where the adenine of the bound ATP is surrounded by hydrophilic residues and forms hydrogen bonds with the hinge region residues [32-34]. The N-terminal lobe is composed of five-stranded β -sheets (β 1- β 5) and a single α -helix, called C-helix (α C) (**Figure 4**). A flexible Glycine-rich (GxGxxG) loop, known as P-loop or G-loop, lies between the β 1 and β 2 sheets of the N-terminal lobe and contains a key hydrophobic residue at its tip, which contributes to the coordination of the phosphates of ATP (**Figure 4**) [33-35]. The P-loop folds over the γ -phosphate of ATP and adopts different conformations depending on the catalytic state [36].

The N-terminus of the C-helix interfaces with the activation loop, which is also called activation segment or A-loop (**Figure 4**) [37]. The activation loop adopts either an open, for the active state of the kinase, or various closed conformations indicating the inactive state with no access to the substrate binding pocket [34, 35]. The A-loop provides the platform, together with the helical subdomain of the C-lobe, for binding and positioning the hydroxyl group residue in the substrate [34, 38]. For an efficient catalysis, a conserved salt bridge, between the active site lysine (of the AXK-motif in the β 3-strand) and the glutamate from the C-helix (α C-in) is formed thereby assisting the correct positioning of the N-terminus of the C-helix. The presence of the salt bridge is required for the formation of the active state. Meanwhile, rotating the N-terminus of the C-helix in a suboptimal position for catalysis (α C-out) results in an inactive state of the kinase [29, 33, 35, 39]. The first residue in the N-terminus of the hinge region, which is located deep in the ATP binding pocket and called the gatekeeper, controls the access to a back-pocket located behind it (**Figure 5**) [33, 39-41].

The helices are dominant in the C-terminal lobe of the kinase domain (**Figure 4**). The catalytic loop starts with β 6 strand followed by β 7 strand, which contains the well-conserved catalytic machinery tyrosine/histidine-arginine-aspartate (Y/HRD-motif). The last two short β -strands (β 8 and β 9) are included in the activation loop (A-loop), which is stuck to the aspartate-phenylalanine-glycine motif (DFG-motif). The HRD-motif and the catalytic loop are surrounding the actual site of the phosphate transfer (**Figure 5**).

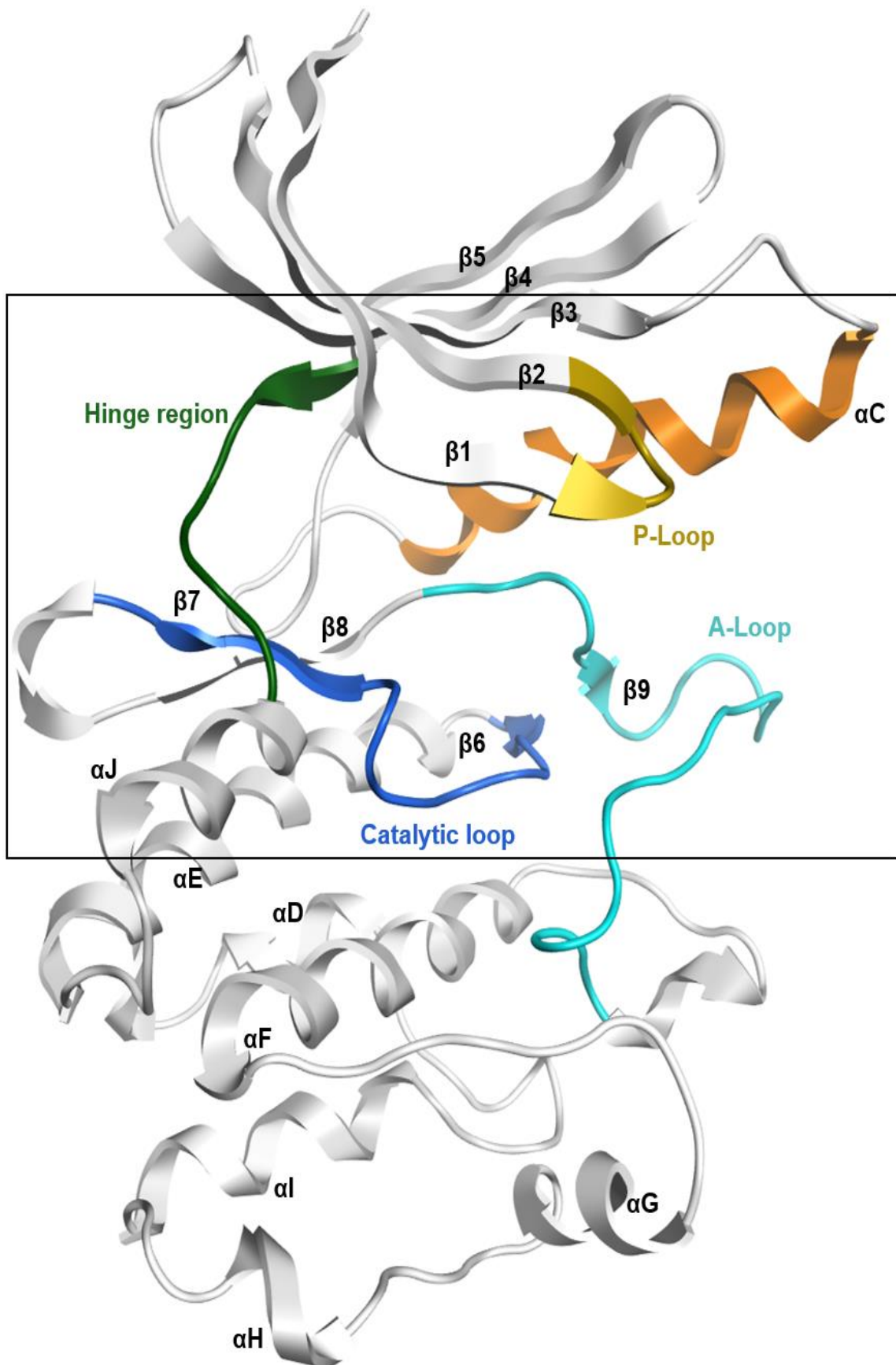


Figure 4. Overall structure of active kinase domain (PKMYT1, PDB ID: 5VCZ [42]). Black box marks the catalytic site of the kinase domain.

The conserved aspartate residue of the Y/HRD-motif is responsible for the correct orientation of the phosphate-OH acceptor group of the substrate. Meanwhile, the conserved Y/H of the Y/HRD-motif serves as a central scaffold, which binds with the aspartate side chain of the Y/HRD-motif and undergoes hydrophobic interactions with the phenylalanine of the DFG-motif (**Figure 5**). The Mg^{2+} positioning loop including the DFG-motif precedes the $\beta 9$ strand, which forms an antiparallel β -sheet with the $\beta 6$ strand of the catalytic loop (**Figure 4**). The Mg^{2+} positioning loop consists of the first five residues of the A-loop and binds Mg^{2+} which in turn coordinates the α -, β -, and γ -phosphates of ATP. In most kinases, the A-loop starts with the DFG-motif and ends with Ala-Pro-Glu (APE-motif) [24]. The phenylalanine of the DFG-motif has an impact on the appropriate position of the aspartate residue of the DFG-motif and the ability for the accommodation of the C-helix, which facilitates the formation of the Lys-Glu salt bridge (**Figure 5**). Hence, the DFG-motif controls the access to the active site [43]. Moreover, the DFG-motif reveals distinct conformations depending on the activation states of the kinase domain [40, 44]. In the inactive state, which displays a DFG-out conformation, the Asp of the DFG-motif is directed outwards the active site and the Phe of the DFG-motif is directed inside (**Figure 6**). In the active state, the Asp of the DFG-motif is directed into the active site and the Phe is outward (**Figure 6**). The remaining helices of the C-lobe including helices G, H, and I (GHI domain, (**Figure 4**) are unique to the different kinases and serve as a platform for the binding of the substrate and regulatory proteins [33]. Therefore, the catalysis needs the collaboration of the following three motifs (**Figure 5**) [25, 33, 35]:

- The active site Lys of the AXK-motif forms a salt bridge with the conserved Glu of the C-helix which interacts with the α and β phosphates of ATP to stabilize and orient the ATP.
- The Asp within the Y/HRD-motif is considered a catalytic residue and serves as a base acceptor for the proton transfer from the phosphate-OH acceptor group of the substrate, thereby, facilitating the nucleophilic attack of the hydroxyl group on the γ -phosphate group of ATP.
- The Asp of the DFG-motif binds to the Mg^{2+} ion that coordinates the β and γ phosphates of the ATP in the binding pocket preparing the ATP for the phosphate transfer.

The re-orientation of the C-helix (C-helix-in) brings the conserved Glu residue into the active site. Then, a salt bridge between Glu and the active site Lys of the AXK-motif is formed (**Figure 5**). Consequently, the kinase is activated. The activation is also associated with the altering in the DFG-motif conformation, where the Asp of the DFG-motif moves from out (DFG-out) into the binding site (DFG-in). The catalytic loop Y/HRD-motif is stable during the changing between the active and inactive states [37]. The fully open conformation of the A-loop can be stabilized either via a salt bridge with the side chain of the HRD-motif and one or more of phosphorylated residues in the A-loop or interactions with accessory regulatory proteins [34, 41, 45]. The P-loop assists in coordinating the phosphates of ATP [35].

In summary, three main regulatory elements of the kinase domain are observed including the N-terminal lobe of the C-helix (α C-in, active; α C-out, inactive), the DFG conformations (Asp: DFG-in, active; DFG-out, inactive), and the A-loop conformations (open, active; closed, inactive) (Figure 6).

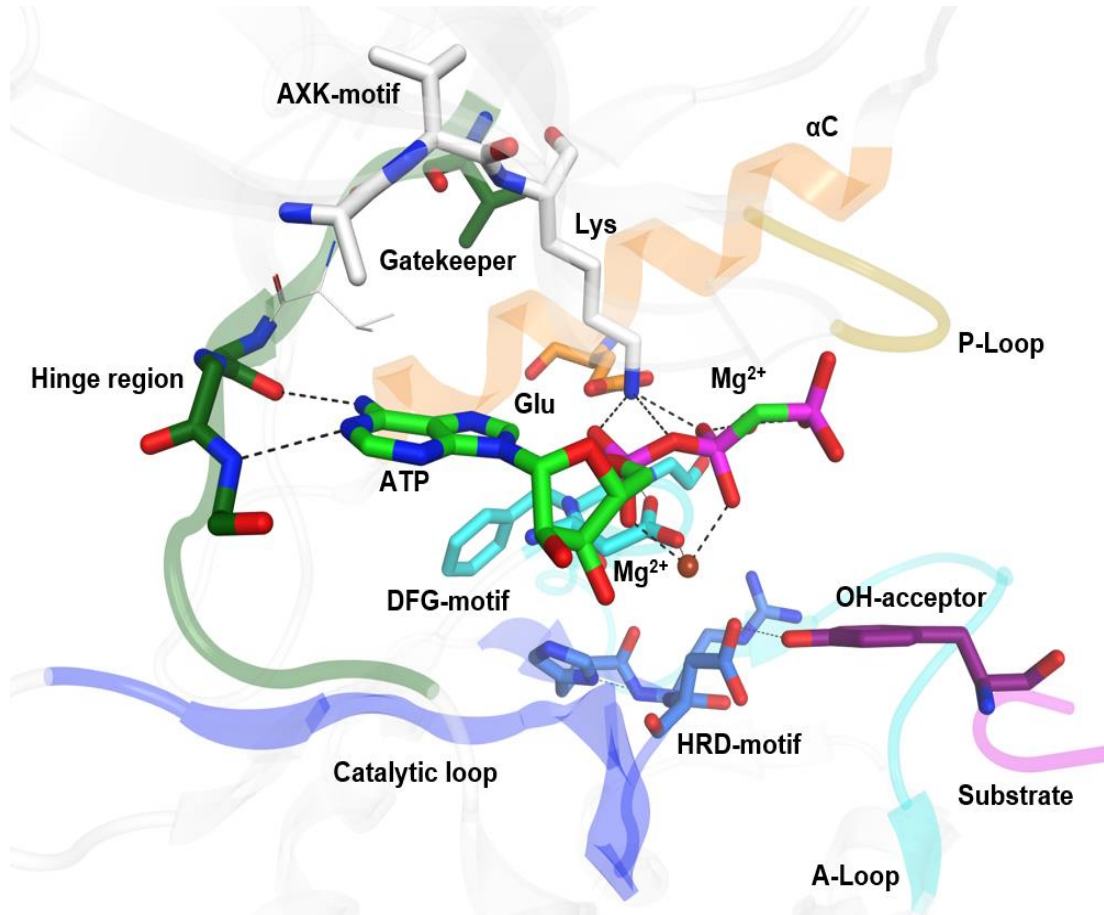


Figure 5. Catalytic site of the kinase domain (FGFR2, PDB ID: 2PVF [46]). Key residues in the hinge region (gatekeeper and residues forming conserved hydrogen bonds) are colored dark green. The conserved catalytic Glu in the C-helix colored gold. DFG-motif is colored cyan as well as the A-loop. HRD-motif of the catalytic loop is colored blue. The active site Lys and the AXK-motif are colored white. ATP residue is shown in green. The substrate and the acceptor OH group of the tyrosine are colored violet. Hydrogen bonds and metal coordination are displayed as black dashed lines. The brown sphere indicate Mg^{2+} ion.

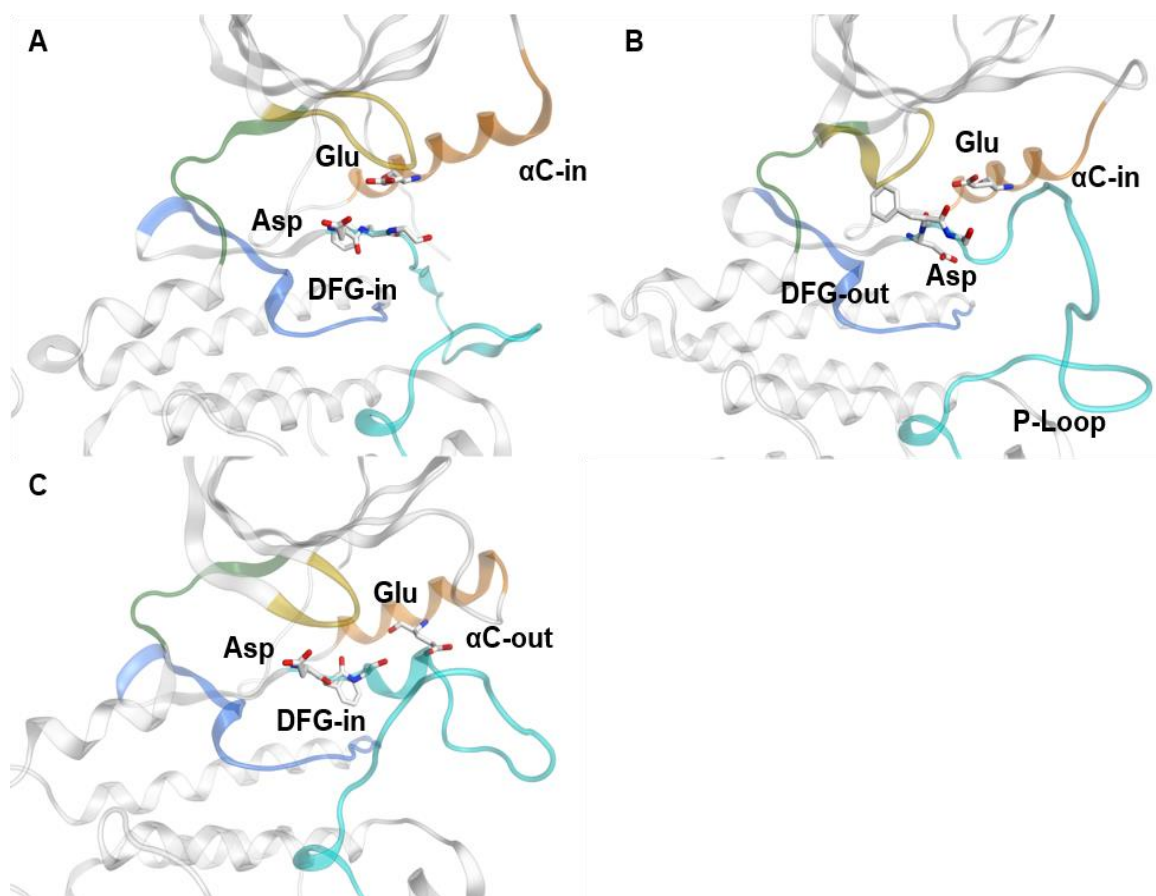


Figure 6. Activation states of the kinase domain. A: Active state of EGFR (PDB ID: 1M17 [47]), which reveals DFG-in, α C-in, and open-A-loop conformations. B: Inactive state of Abl (PDB ID: 3OXZ [48]) displays DFG-out, α C-in, and closed-A-loop conformations. C: Inactive EGFR (PDB ID: 4HJO [49]) shows DFG-in, α C-out, and closed-A-loop conformations.

1.2.3. Kinase inhibition

The largest category of established and clinical trial agents are small molecules. They comprise ~86% of FDA-approved therapeutic agents and about 63% of clinical trial agents that target the human druggable targets [9]. The remaining categories are namely monoclonal antibodies and recombinant proteins, synthetic peptides, proteins purified from biological sources, vaccines, adenoviral gene delivery, and small interfering RNAs (siRNAs) (**Fig. 7**) [9].

Santos et al. analyzed the FDA-approved kinase inhibitors till 2015 (**Figure 2**). In their work, they recognized that most of the approved kinase inhibitors were within the period of 2011 to 2015 [7]; the majority of them are small molecules with molecular weight less than 500 Daltons. A deeper analysis of the FDA-approved kinase inhibitors was performed by Wu et al [50]. They reported 15 new small molecule kinase inhibitors that were approved by FDA for the time between January 2012 and February 2015 [50] in addition to a large number of other compounds currently being evaluated in clinical and preclinical trails. So far, small molecule kinase inhibitors have been identified for 20% of the human kinome. By April 2018, 41 small molecule kinase inhibitors were

approved by FDA a (**Table 1A** adapted from www.brimr.org/PKI/PKIs.htm, Fabbro et al. and Roskoski, see **Appendix**) [24, 37]. So, additional small molecule kinase inhibitors have been approved by FDA from January 2015 to April 2018, (**Table 1A**, see **Appendix**). Thus, the development of small molecule kinase inhibitors can be considered as one of the most extensively pursued areas of drug discovery.

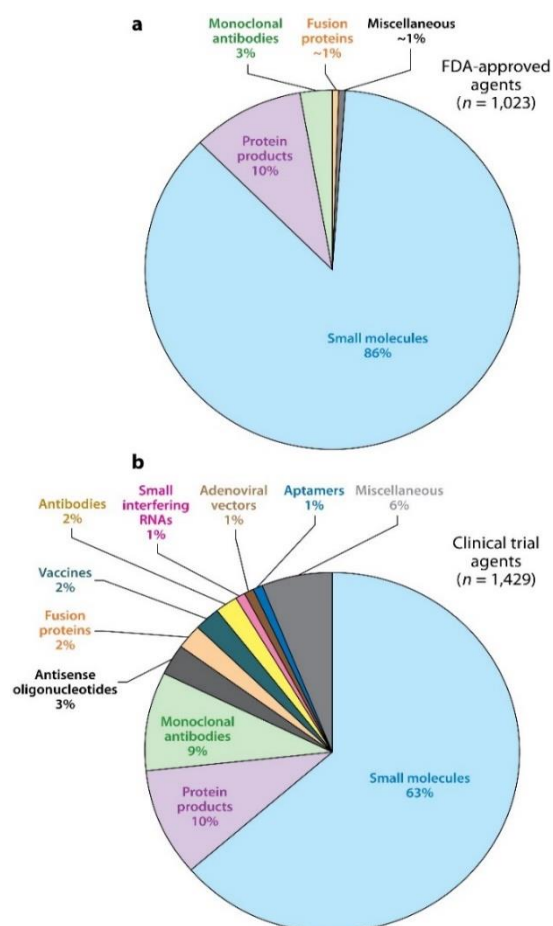


Figure 7. Categories of established (a) and clinical trial agents (b). Adapted from Rask-Andersen et al [9].

Historically, kinase drug discovery started in 1995 when fasudil (HA-1077) was approved in Japan as an ATP-competitive inhibitor of ROCK1/2 kinase for the treatment of cerebral vasospasm [51, 52]. The second approved kinase inhibitor was sirolimus (Rapamune) in 1999, which is an allosteric inhibitor of mTOR kinase; used in a combination with cyclosporine for the prevention of the organ rejection in patients receiving renal transplants [53, 54]. The first FDA-approved kinase inhibitor was imatinib (CGP57148, STI571, Glivec, Gleevec), a compound designed and developed by Novartis in 2001. Imatinib binds to the inactive state of the Abelson (Abl) kinase in the context of the Bcr-Abl translocation found in chronic myeloid leukemia (CML) [55, 56]. The kinase drug discovery was running slowly in the first decade of the 21st century with almost one new approval per year on average.

1.2.4. Classification of small molecule kinase inhibitors

Generally, several classifications of the kinase inhibitors have been described based on the activation state of the protein kinase. Dar and Shokat defined three classes of small molecule kinase inhibitors, which they labeled type I, II, and III [57]. Kinase inhibitors that bind to the active state (DFG-in), were considered as type I. Type II refers to inhibitors binding to the inactive state (DFG-out), while type III indicates non-ATP competitive inhibitors or allosteric inhibitors binding outside the ATP-binding site [58]. An additional class, called type I_{1/2}, was introduced by Zuccotto et al. to describe the kinase inhibitors that bind to the kinase state which display DFG-in and α C-out conformations [59]. Garvin and Saiah further divided allosteric inhibitors into two classes (III and IV), where type III inhibitors bind within the cleft between the small and large lobes adjacent to the ATP-binding pocket, whereas type IV describes inhibitors outside the cleft and the phosphor-acceptor region [60]. In 2012, class V was presented to label bivalent molecules that extend in two regions of the kinase domain while small molecules that form covalent adducts with the kinase domain were labeled as type VI [61]. A recent and relevant classification of the small molecules kinase inhibitors was described by Roskoski [24]. He divided the type I_{1/2} and type II into A and B subtypes. A summary of the most advanced classification of small molecule kinase inhibitors is shown in **Table 1** below [24].

Table 1. Classification of small molecule protein kinase inhibitors [24, 57, 59-61].

Properties	Type I	type I_{1/2} (A/B)	Type II (A/B)	Type III	Type IV	Type V	Type VI
	Binds in the ATP-pocket of the active state	Binds in the ATP-pocket of the inactive state	Binds in the ATP-pocket of the inactive state	Allosteric inhibitors bound next to the ATP-site	Allosteric inhibitors not bound next to the ATP-site	Bivalent inhibitor spanning two regions	Covalent inhibitors
Extends into back cleft	No	(A)Yes/ (B) No	(A)Yes/ (B) No	Yes	No	Variable	Variable
DFG	In	In	Out	Variable	Variable	Variable	Variable
A-loop	Out	Variable	Variable	Variable	Variable	Variable	Variable
αC	In	Variable	Variable	Out	Variable	Variable	Variable
ATP-competitive	Yes	Yes	Yes	No	No	Variable	No
Reversible	Yes	Yes	Yes	Yes	Yes	Yes	Usually not

Generally, the FDA-approved inhibitors form hydrogen bonds with one or more hinge residues except the allosteric ones. Most of the inhibitors form hydrophilic contacts with residues of the solvent-accessible area such as β 2, β 3, and β 7 strands (**Figure 4** and **8**). Moreover, many inhibitors are involved in hydrophobic interactions with the aliphatic side chain of the active site lysine. About half of the inhibitors interact with one or more of the C-helix residues and the gatekeeper residue (**Figure 8**). In addition, the gatekeeper may also be involved in a hydrogen bond or hydrophobic interactions depending on its nature [24]. Two-thirds of the inhibitors interact with one or more residues of the C-helix and β 4 in the back-pocket.

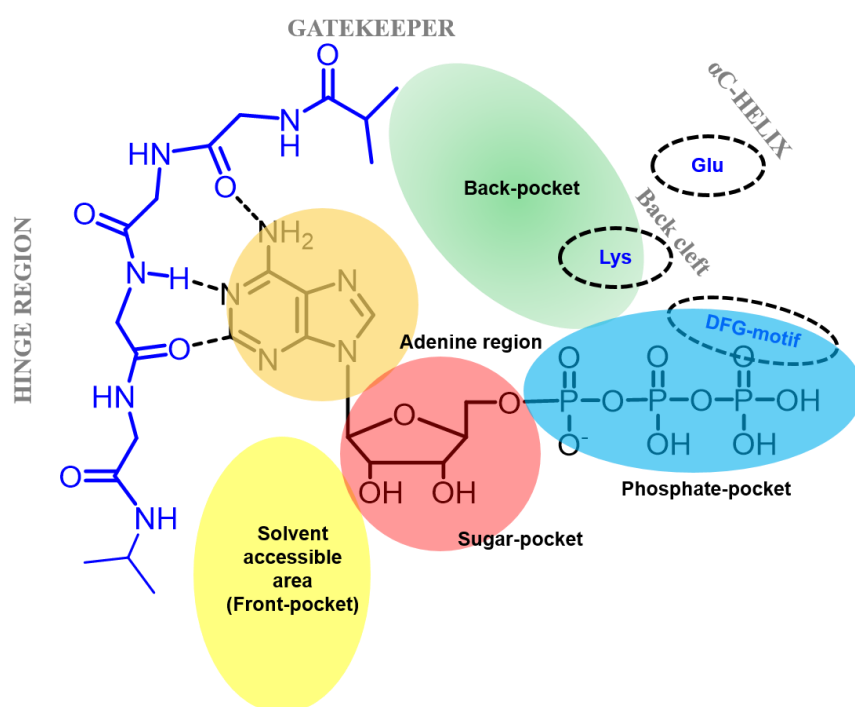


Figure 8. Scheme of the ATP-binding pocket of the kinase domain. Colored spheres illustrate: Solvent-accessible area, which includes the front-pocket (yellow), ribose-binding pocket (red), phosphate region (blue), adenine region (white brown), back cleft as a part of the back-pocket (green).

1.2.5. Lipid kinase inhibitors

In July 2014, the first lipid kinase inhibitors, idelalisib (Zydelig, Gilead Sciences), was approved for the treatment of chronic lymphocytic leukemia [62]. Idelalisib is a PI3K δ inhibitor and reveals a type II binding, where it adopts a propeller binding mode with the DFG-motif conformations [63]. Interestingly, it shows the conserved interactions found in small molecule kinase inhibitors including hydrogen bonds with hinge residues and hydrophobic interactions with the back-pocket [64]. More than 20 other PI3K inhibitors are in development for the treatment of cancer and

inflammatory diseases e.g. buparlisib (Phase III, Novartis) as a selective-isoform inhibitor of PI3K α , dual inhibitors of PI3K and mTOR, and pan-class I PI3K inhibitors [65].

1.3. Limitations, challenges, and future perspectives in kinase drug discovery

Despite the progress in the kinase drug discovery in the 21th century, several significant challenges for developing kinase inhibitors should be addressed.

Firstly, only a small subset of the human kinome has so far been studied. Furthermore, most of the discovered kinase inhibitors are limited to a select group of the human kinome. Also, the vast majority of the FDA-approved kinase inhibitors (18 out of 27) only target three groups of tyrosine kinases, namely Bcr-Abl, ErbBs, and VEGFRs [50]. Therefore, there is a need to develop chemical tools to uncover the functions of the so far less studied kinases.

Secondly, although kinase signaling cascades regulate diverse cellular activities and are related to several diseases including inflammatory diseases, CNS disorders, cardiovascular disease, diabetes, and cancer, most of the currently available kinase inhibitors (e.g. 26 of 28) are mainly used in cancer treatments [50].

Thirdly, only a small subset of the chemotypes has been investigated within the approved inhibitors. This is due to the fact that many of the kinase inhibitors were developed based on previously approved molecules. Thus, novel tools are needed to explore pharmacophores so as to diversify the scaffolds of kinase inhibitors.

Fourthly, the selectivity of the kinase inhibitors is a big challenge. Most of the inhibitors are reversible and bind within the ATP-binding pocket, where the similarity is very high among the kinases. It is, therefore, sometimes a necessary requirement to design selective kinase inhibitors in order to avoid off-targets interactions. However at the same time, dual- and/or multi-targets kinase inhibitors might be more favourable for cancer treatments than highly selective inhibitors. Thus, a favourable selectivity profile of the kinase inhibitors is required to balance efficacy and toxicity.

AIM OF THE WORK

Membrane-associated inhibitory kinase PKMYT1 belongs to Wee1-kinase family and regulates the cell cycle at G2/M transition. G2 checkpoint abrogation through the inhibition of PKMYT1 immediately causes apoptotic or non-apoptotic cell death. Inhibition of PKMYT1 by small molecules is therefore proposed to be a promising option for further analysis of PKMYT1 as a target for cancer treatments.

Protein kinase C related kinase (PRK1) is a serine/threonine enzyme and is a regulator of the Androgen Receptors (AR). Recently it was shown that PRK1 is a promising therapeutic target for prostate cancer therapy.

The aim of this work (**Publications 1-5**) was to analyze ligand binding of small molecule inhibitors to PKMYT1 and PRK1 and to design novel specific inhibitors using structure/ligand-based drug design methods. In addition, novel computer-based approaches for the prediction of kinase activity were evaluated on larger kinase data sets. Due to the absence of crystal structures for PRK1 at the beginning of the project, homology modeling approaches were applied to generate 3D structures of the enzyme. Virtual screening methods were employed on the PRK1 homology models and the available PKMYT1 crystal structures to search for novel lead compounds in commercial databases. Molecular dynamics based binding free energy methods were used to optimize or prioritize these lead compounds. In order to develop novel inhibitors, the identified inhibitors were used to derive a quantitative structure-activity relationships (QSAR) to guide the development of novel PKMYT1 and PRK1 inhibitors. The availability of small molecule inhibitors of PRK1 and PKMYT1 that are selective and bioavailable will represent a major breakthrough in this field for further analysis of PKMYT1 and PRK1 as targets for cancer treatments.

Developing kinase inhibitors for receptors of the growth factors EGFR, IGF-1R, VEGFR2, and PDGFR1 (**Publications 6-9**). Cancer initiation requires the accumulation of a number of genetic mutations, which transform the normal cells into highly malignant cells. The expansion and progression into metastatic carcinomas depends on a multi-step process regulated by several growth factors. The growth factors bind to transmembrane receptors harboring kinase activity such as EGFR, IGF-1R, VEGFR2, and PDGFR- β to stimulate specific intracellular signaling pathways. On the other hand, overexpression of these receptors may enable cancer cells to become hyper-responsive to the growth factors. Therefore, inhibition of the receptors of the growth factors offers opportunities for the cancer therapy.

Due to occurring resistance against the known inhibitors of EGFR, IGF-1R, VEGFR2, and PDGFR- β , the search for novel lead compounds with new scaffolds is necessary to reverse the over-activity of the growth factors receptors. Part of this research project aimed to develop novel pyrrolopyrimidine derivatives in order to inhibit specific growth factors receptors, which might be

useful in cancer therapy. Therefore, molecular modeling and analysis of the ligand binding was investigated to provide a potential rationalization for the detected *in vitro* activity and to further guide the development of novel inhibitors. Moreover, to design selective inhibitors for EGFR, IGF-1R, VEGFR2, and PDGFR- β , structural analysis of their ATP-binding pockets was performed in order to study their similarity and their diversity. For lead optimization, the relevant pharmacophore features within the ATP-binding pocket were considered to enhance the activity of the novel compounds. The availability of novel inhibitors for growth factors kinase receptors might reduce the resistance development in cancer cells.

METHODS AND MATERIALS

A brief description of the computational methods used in the current work is presented in this chapter. These methods combine binding mode predictions, fragment-based approaches, virtual screening, homology modeling, molecular docking, molecular dynamics simulation, and binding free energy calculations. In addition, the implemented in vitro assays are introduced.

2.1. Computer-aided drug design (CADD)

Computational methods have been extensively used as innovative tools to accelerate drug discovery. Hits identifications, binding modes predictions and lead optimizations are considered the main applications of computational approaches. The recognition of the biologically active conformer for a molecule plays an important role in the structure-based drug design. Rational drug design relies on the accurate characterization of the bioactive conformation of a given molecule within the protein binding pocket. The observed binding modes serve as atomic-level descriptions of how the ligands interact with the binding pocket residues [66]. The binding modes also provide a way to map the experimental biological activities to individual structural features, which is useful for designing novel drug candidates [67, 68]. Therefore, several computational and experimental tools from biophysical methods e.g. X-ray crystallography and NMR have been established to reveal the binding modes of the ligands within the protein structures. Molecular docking is considered as an important process to predict the ligand binding mode(s) besides its ability in hits identifications and lead optimizations [69, 70]. Numerous studies have evaluated the docking algorithms to assess their ability to re-produce the experimental binding conformations of the ligands [70-77]. Although most of the docking algorithms perform well for ligand binding modes generations and prediction, the scoring functions often fail to predict the corresponding affinity [70-77]. Molecular visualization of the ligand binding modes in the current work has been performed using several programs including Molecular Operation Environment System (MOE) (Chemical Computing Group Inc., Montreal, QC, Canada), Maestro (Schrödinger, LLC, New York, NY), PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC), and Chimera [78].

2.1.1. Fragment-based drug design (FBDD)

One of the promising approaches in CADD is fragment-based Drug Design (FBDD) which was established in last decade as an approach for hit identification and lead generations [79-81]. FBDD shows a high ability to sample the chemical space compared to other techniques [82]. Fragments are defined by Astex using the Jorgensen's "Rule of Three" as small molecules for which the molecular weight is <300, the number of hydrogen bond donors is ≤ 3 , the number of hydrogen bond acceptors is ≤ 3 , and the cLogP is ≤ 3 [83]. Once the fragment hits are identified, the process of hit to lead generation is performed by growing, linking and merging approaches. In addition, fragment growing approach provides useful tools for hit and lead optimization, which aims to build additional interactions with the adjacent pockets [84, 85]. Two metrics ligand efficiency (LE) and ligand lipophilicity efficiency (LLE) are available to aid in the fragment hit selection and optimization. Whereas the LE metric is used to rank fragments and to monitor progress of the optimization, LLE is a metric used to monitor the lipophilicity in respect to the *in vitro* potency of a molecule [86-88].

During the optimization of novel kinase inhibitors, several chemical scaffolds, which were extracted from the identified hits, were used as starting points. The fragment growing approach, which allows the coverage of a larger chemical space, was applied to design novel kinase inhibitors. The ideal values of LLE are located in the range of 5 to 7 or greater [80, 89, 90]. The following equations were used to estimate the LE and LLE:

$$LE = 1.4(-\log K_i)/HAC, \quad HAC \text{ is heavy atoms count}$$

$LLE = pK_i - cLogP$, $cLogP$ was calculated using MOE (Chemical Computing Group Inc., Montreal, QC, Canada) or Maestro (Schrödinger, LLC, New York, NY) [91-93].

2.1.2. Virtual screening

Virtual Screening (VS) is considered as a complementary approach to experimental or physical high throughput screening HTS [94]. It is a computational tool for a rapid *in silico* screening of large chemical databases in order to identify bioactive compounds prior to *in vitro* testing. VS approaches are divided into two categories: ligand-based and structure-based VS. Both approaches were used in the current work.

Ligand-based VS depends on the availability of known active compounds. In most cases, it is employed when the structure of the drug target is unknown or there is not enough structural information about the target protein. Ligand-based VS relies on the fact that similar molecules are prone to exhibit a similar biological activity [95, 96]. Due to the lack of structural information of the targeted kinase e.g. PRK1 kinase, a ligand-based approach was employed to search for molecules similar to known actives based on the 2D chemical fingerprints or 3D shape-based. The similarity search was performed using either MOLEPRINT 2D [97, 98] or MACCS key fingerprints implemented in MOE. The quantification of the similarity was given by the Tanimoto coefficient [99].

Structure-based VS, on the other hand, depends mainly on the availability of the 3D structure of the target protein, chemical database, and a docking algorithm [100, 101]. Structure-based VS makes use of docking algorithm to place and score molecules within the protein binding site. This allows to predict the binding modes and evaluate each molecule separately [102, 103]. In the current work, the 3D structures of the targeted kinase e.g. PKMYT1 and EGFR were obtained either from the Protein Data Bank (RSCB PDB; www.rcsb.org) [104] or from homology modeling e.g. PRK1, (see Section 2.3). Molecular docking into the ATP-binding pocket was performed using GOLD v5.2 [105] and Glide (Schrödinger, LLC, New York, NY) [91-93] as described in section 2.4.

2.2. Homology modeling

Homology modeling (HM) or comparative modeling is a methodology which allows modeling 3D structures for unknown protein structures. HM has been used to derive the 3D protein structures of non-resolved protein kinases e.g. PRK1. The HM process depends on two fundamental points: First of all, knowing the primary amino acid sequences of the target protein. Secondly, similar sequences display similar 3D structures [106]. Generally, HM includes [107]:

1. Identification and recognition of the closest homolog, which is used as a template.
2. Alignment of the target sequence and the template sequence.
3. Building models from the target sequence based on a 3D structure of the template.
4. Validation of the models.

The primary sequence of the target kinases was retrieved from UniProt Knowledgebase (www.uniprot.org) [108]. Then, a basic local alignment search tool (BLAST) [109] search using the Protein Data Bank (RCSB PDB; www.rcsb.org) [104] or Swiss-Model [110], was used as a query to identify the best templates, which were other family members of the human kinome. The templates were selected based on the sequence similarity, the X-ray resolution, as well as a consideration of several conformations of the kinase domain. Once the template was established, the alignment of the sequences was performed using either MOE (Chemical Computing Group Inc., Montreal, QC, Canada) or MODELLER [111, 112]. The alignment performance was evaluated based on the key residues, for instance DFG-motif, gatekeeper, AXK-motif, and HRD-motif in the kinase domain, which are conserved among kinases. Finally, the homology models were built in MODELLER [111, 112]. Generally, five homology models were generated for each query. The selection of the best model relied on an internal score in MODELLER, which is called Discrete Optimized Protein Energy (DOPE) [113] given for each model. The quality of the homology models were subsequently evaluated by means of Ramachandran plot using PROCHECK to analyze the stereochemical quality [114].

2.3. Molecular docking

Preparation of the inhibitors databases

Ligands were prepared using LigPrep module in Schrödinger (LigPrep, Schrödinger, LLC, New York, NY). The preparation process involved 3D protonation at pH=7.4 using Epik module [115, 116]. This was followed by an energy minimization step using the integrated Optimized Potentials for Liquid Simulations (OPLS_2005) force field [117]. Lastly, the generation of ten low energy conformations per ligand was executed. Further preparation was done using the conformer generation module of OMEGA in OpenEye Scientific Software [118]. A maximum of twenty conformers for each ligand molecule were generated.

Preparation of protein structures

Structures of respective proteins were obtained either by downloading them from the PDB website (RSCB PDB; www.rcsb.org) [104] or from homology modeling approach. The structures were resolved in an active state displaying a DFG-in conformation either as an apo-form, or a holo-form. Preparation module in MOE (Chemical Computing Group Inc., Montreal, QC, Canada) or Protein Preparation Wizard (Schrödinger, LLC, New York, NY) were used for preparing the kinase domains. The process consisted of 3D protonation, followed by energy minimization using either Amber10: Extended Huckel Theory (EHT) force field implemented in MOE [119, 120] with a tethering of 0.5 Å and a gradient of 0.1 Kcal.mol⁻¹. Å⁻² for all the atoms during the minimization. Meanwhile, the integrated Optimized Potentials for Liquid Simulations (OPLS_2005) force field [117] with a tethering of 0.3 Å using protein preparation wizard in Maestro (Schrödinger, LLC, New York, NY) was used for the protein preparation.

The prepared ligands were docked into the respective protein structures using GOLD v5.2 [105] and Glide (Schrödinger, LLC, New York, NY) [91-93]. The co-crystallized ligands of the holo-form structures were defined as a center of the binding site. Meanwhile a residue of the hinge region was used to define the center of the binding site in the apo-form structures. In addition, the diameter of the box was enlarged from the default setting to cover the whole binding pocket. At least one hydrogen bond with the hinge region residue was used as a constraint. Several docking scores, which were supplied in Glide and GOLD, were applied for docking the ligands. A Linux-based cluster was used in docking. Generally, several poses were generated for each docked ligand. Other docking options were kept as default.

2.4. Molecular Dynamics (MD) simulations

Molecular dynamics (MD) simulations were performed to study the conformational stability of the generated docking complexes and to calculate the binding free energy. MD simulations were carried out using AMBER 12 [121] and employing AMBER03 force field for the protein residues [122-124]. Antechamber module was used to generate parameters for atom types and atomic charges (AM1-BCC [125]) for each ligand, in general AMBER force field GAFF was adopted [126]. The preparation of the protein-ligand complexes was performed using LEaP module in AMBER, leading to the generation of parameter/topology files and coordinate files for the protein-ligand complexes. Counter ions were then added to neutralize the system followed by solvation of the system. The TIP3BOX water model was used to solvate the system (protein-ligand complex and ions) in a truncated octahedron box with edges of 10 Å [127]. Several steps were performed to prepare the system for MD simulations; starting with two consecutive steps of energy minimization. The first minimization was carried out for water and ion molecules while keeping the ligand-protein complex restrained to their initial coordinates with a force constant of 500 Kcal.mol⁻¹. Å⁻². The

procedure helped to reduce unreal van der Waals interactions with the surrounding solvent molecules and rearrangement of the inserted water and ions molecules. In this step, 2000 iterations (beginning with 1000 steepest descent and followed by 1000 conjugate gradient) were applied. The second minimization step was applied for the whole system through 10000 iterations (first 5000 steepest descent and then 5000 conjugate gradients) without any restraint. After the energy minimization of the system, the next step was heating. In this process, the temperature was gradually elevated for 100 ps from 0 to 300 K and over a timescale 2 fs to avoid problems with the hot solvent cold solute. Langevin dynamics was set for temperature control using a collision frequency of 1 ps^{-1} during the temperature equilibration [128]. The cutoff force constant of the restrained system was set at $10 \text{ Kcal.mol}^{-1} \cdot \text{\AA}^{-2}$. The output coordinates after the heating were inserted in pressure equilibration routine, in a third step is called density, which was used to set the constant pressure periodic boundary. The constant pressure periodic boundary was converted from 1 bar to 2 bar during 100 ps with a timescale of 2 fs at 300 K and applying the same cutoff force constant at $10 \text{ Kcal.mol}^{-1} \cdot \text{\AA}^{-2}$. The density step was followed by either a 100 ps or 100 ns MD simulation depending on the purpose of the MD step. A cutoff of 10 \AA for non-bonded atoms was applied using Particle Mesh Ewald methods [129]. All simulation were ran using SHAKE to constrain hydrogen bonds [130].

2.5. BFE calculations

The binding free energy calculations were performed for re-scoring docking solutions. AMBERTools12 [121] and MOE (Chemical Computing Group Inc., Montreal, QC, Canada) were used to calculate the binding free energies. The calculation was initially carried out using one snapshot after a short energy minimization of the whole system in explicit water. Later, ten snapshots (frames) of ligand-protein complexes from the MD simulation were considered for averaging the binding free energy. Different solvation theories were applied using either Generalized Born (GB) [131, 132] or Poisson Boltzmann models (PB) [133]. The molecular mechanics (MM) and quantum mechanics/molecular mechanics (QM/MM) hybrid were considered too [134, 135]. Consequently, the binding free energy calculations of the models MM-PB(GB)/SA and MM/QM-GB/SA were computed using MMPBSA.py [136] module in AMBERTools12 [137-142]. In addition, the model MM-GB/SA, implemented in MOE, was used to estimate the binding free energy.

The affinity of a ligand for a protein can be calculated by differences between the free energy of the complex and the sum of free energies of the ligand and the protein separately, eq. (1) [143]:

$$\Delta G_{\text{binding}} = G_{\text{protein/ligand}} - (G_{\text{protein}} + G_{\text{ligand}}) \quad (1)$$

The absolute free energy for each part can be estimated as a sum of the potential free energy or the so-called gas-phase free energy E_{MM} , the changing of the free energy due to the solvation $\Delta G_{solvation}$, and the entropic contribution $T\Delta S$, where T is the absolute temperature and S the entropy of the molecule, eq. (2) [144] :

$$G_{molecule} = E_{MM} + \Delta G_{solvation} - T\Delta S \quad (2)$$

E_{MM} : Potential free energy or gas-phase free energy, which is represented by the empirical forcefield, can be computed by molecular mechanics, as seen in eq. (3, 4) [143, 144]:

$$E_{MM} = E_{int} + E_{vdw} + E_{elec} \quad (3)$$

$$E_{int} = E_{bond} + E_{angle} + E_{tors} \quad (4)$$

Where E_{int} represents the internal free energy and counts the free energies of the bond stretching (E_{bond}), angle bending (E_{angle}) bond torsions/dihedrals (E_{tors}). The second term is van der Waals energy (E_{vdw}), which is used to consider the short and long ranged interactions between the atoms. Meanwhile, E_{elec} refers to the electrostatic energy, which counts the Coulomb interactions as a result of partial atomic charge interactions. SANDER module of AMBERTools12 [121] was used to estimate the values of the previously mentioned energy units. The solvation energy $\Delta G_{solvation}$ counts the impacts of the solvent on the molecules. The solvation energy consists of two terms; the electrostatic solvation energy (polar contribution, $\Delta G_{PB/GB}$) and the non-electrostatic solvation energy (non-polar contribution, ΔG_{SA}) [145]. Equation (5) is used to compute the solvation energy. Currently, there are two postulated theories to calculate the polar energy which are either the (PB) or (GB) equations using MMPBSA.py [136] module in AMBERTools12. The second term ΔG_{SA} is the effect of solvent on the surface area of the molecules and is estimated depending on the equation of solvent-accessible surface area SASA. Generally, it is referred to as molecular Surface Area free energy SA (ΔG_{SA}) [146], as seen in eq. (6), where γ is the parameter of the surface tension and b a parameterized value [145]:

$$\Delta G_{solvation} = \Delta G_{PB/GB} + \Delta G_{SA} \quad (5)$$

$$\Delta G_{SA} = \gamma \cdot SASA + b \quad (6)$$

In the current study, the solvation free energy was calculated by applying different GB models. These models are implemented in AMBERTools12 as GB^{HCT} ($igb = 1$), $GBOC1$ ($igb = 2$), GB^{OC2} ($igb = 5$), as the last release model ($igb = 8$) [119, 147-151]. The radii settings ($mbondi$ 0, 2, and 3) were adjusted according to the GB model. The radii setting have impact on the polar energy term so likewise the total solvation energy term.

A change of the entropy refers to the loss of translational, rotational and conformational degrees of freedom caused by ligand binding [138, 145, 152]. In most cases, the entropic contribution ($-T\Delta S$) does not improve the correlation and the accuracy of the binding free energy calculation [138]. However, the entropy contribution can be estimated using the normal mode analysis [153], which is time-consuming and computationally expensive. Therefore, it was neglected. However, there is an approach to approximate the entropy contribution based on the total number of rotatable bonds in the ligand (NRot) [135]. This approximation is a useful tool to estimate the total binding free energy [154-157], eq. (7):

$$T\Delta S_{\text{NRot}} = \text{NRot} \times 1.0 \text{ Kcal/mol} \quad (7)$$

The entropic contribution was considered in some parts of the current study. Nevertheless, in most cases only the change of the enthalpy (ΔH) was considered to find the correlation with biological data and it will indicate to the binding free energy, eq. (8).

$$\Delta H = E_{\text{MM}} + \Delta G_{\text{solvation}} \quad (8)$$

The QM region of the QM/MM hybrid calculations was firstly applied for the ligand only and later for the ligand and selected residues from the hinge region, the P-loop, the hydrophobic pocket and the DFG-motif, surrounding the ligand. Two QM models were used: Parameterized Model number 3 (PM3) and Austin Model 1 (AM1) [121].

2.6. *In vitro* testing

The biological testing in the current work included three kinds of assays, binding, activity, and cellular assays. The results from the binding assay were confirmed by the activity and the cellular assays.

Binding assay: Screening was carried out using different fluorescence polarization based binding assays (e.g. LanthaScreenTM) with specific concentrations of the target kinase, antibody, and kinase tracer. The screening was performed using several concentrations of the compounds. In addition, positive and negative (DMSO 1%) controls were used for comparison. Compounds with a tracer displacement above a certain percentage were further used to determine the IC_{50} and K_i values.

Activity assay: The inhibitory activities of the hit compounds against kinase were determined using various fluorescence polarization based kinase assay e.g. FlashPlatesTM. Similar to the binding assay, positive and negative (DMSO 1%) controls were used. Thus, the IC_{50} and K_i values were determined.

Cellular assay: Two kinds of cellular assays (functional cellular assay and cellular growth inhibition assay) were conducted in the current work. The functional cellular assay was applied to test the inhibitory performance of the identified inhibitors over several series of cancer cell lines that were treated for 24h. Then the cells were lysed and the amount of the phosphorylated substrates was quantified, isolated and prepared for western blotting using a specific antibody. As a result, the correlation between the cellular assays and the data of either the binding or activity assay was calculated. On the other hand, cellular growth inhibition assay [158, 159] was carried out to measure the growth inhibitory activity at one given concentration of the most promising compounds for 24h. Accordingly, the percentage of the growth inhibition at the given screening concentration using the sulforhodamine fluorescence assay was determined.

RESULTS AND DISCUSSIONS

This section presents abstracts from nine publications relevant to this project. The full texts of the publications are given in the Appendix. The presented manuscripts discuss the structure-based design and development of inhibitors for PKMYT1, PRK1, EGFR, IGF-1R, PDGRF- β , and VEGFR2.

3.1. Regulation of G2/M Transition by Inhibition of WEE1 and PKMYT1 Kinases

Schmidt, M., Rohe, A., Platzer, C., Najjar, A., Erdmann, F., & Sippl, W.

Molecules, 22, 2045, 2017.

<https://doi.org/10.3390/molecules22122045>

Abstract: In the cell cycle, there are two checkpoint arrests that allow cells to repair damaged DNA in order to maintain genomic integrity. Many cancer cells have defective G1 checkpoint mechanisms, thus depending on the G2 checkpoint far more than normal cells. G2 checkpoint abrogation is therefore a promising concept to preferably damage cancerous cells over normal cells. The main factor influencing the decision to enter mitosis is a complex composed of Cdk1 and cyclin B. Cdk1/CycB is regulated by various feedback mechanisms, in particular inhibitory phosphorylations at Thr14 and Tyr15 of Cdk1. In fact, Cdk1/CycB activity is restricted by the balance between WEE family kinases and Cdc25 phosphatases. The WEE kinase family consists of three proteins: WEE1, PKMYT1, and the less important WEE1B. WEE1 exclusively mediates phosphorylation at Tyr15, whereas PKMYT1 is dual-specific for Tyr15 as well as Thr14. Inhibition by a small molecule inhibitor is therefore proposed to be a promising option since WEE kinases bind Cdk1, altering equilibria and thus affecting G2/M transition.

3.2. Identification of PKMYT1 Inhibitors by Screening the GSK Published Protein Kinase Inhibitor Set I and II

Platzer, C., Najjar, A., Rohe, A., Erdmann, F., Sippl, W., & Schmidt, M.

Bioorganic & Medicinal Chemistry, 2018

<https://doi.org/10.1016/j.bmc.2018.06.027>

Abstract: As a member of the Wee-kinase family protein kinase PKMYT1 is involved in G2/M checkpoint regulation of the cell cycle. Recently, a peptide microarray approach led to the identification of a small peptide; EFS247-259 as substrate of PKMYT1, which allowed for subsequent development of an activity assay. The developed activity assay was used to characterize the PKMYT1 catalyzed phosphorylation of EFS247-259. For the first time kinetic parameters for PKMYT1, namely K_m , K_m , ATP and v_{max} were determined. The optimized assay was used to screen the published protein kinase inhibitor sets (PKIS I and II), two sets of small molecule ATP-competitive kinase inhibitors reported by GlaxoSmithKline. We identified ten inhibitors, providing different scaffolds. The inhibitors were further characterized by using binding assay, activity and functional assay. In addition, docking studies were carried out in order to rationalize the observed biological activities. The derived results provide the basis for further chemical optimization of PKMYT1 inhibitors and for further analysis of PKMYT1 as target for anti-cancer therapy.

3.3. Computer-aided Design, Synthesis and Biological Characterization of Novel Inhibitors for PKMYT1 Kinase

Najjar, A., Platzer, C., Luft, A., Assmann, C., ElGhazawy, N., Erdmann, F., Sippl, W., & Schmidt, M.

European Journal of Medicinal Chemistry, 2019

<https://doi.org/10.1016/j.ejmech.2018.10.050>

Abstract: In the current work, we applied computational methods to analyze the membrane-associated inhibitory kinase PKMYT1 and small molecule inhibitors. PKMYT1 regulates the cell cycle at G2/M transition and phosphorylates Thr14 and Tyr15 in the Cdk1-cyclin B complex. A combination of *in silico* and *in vitro* screening was applied to identify novel PKMYT1 inhibitors. The computational approach combined structural analysis, molecular docking, binding free energy calculations, and quantitative structure–activity relationship (QSAR) models. In addition, a computational fragment growing approach was applied to a set of previously identified diaminopyrimidines. Based on the derived computational models several derivatives were synthesized and tested *in vitro* on PKMYT1. Novel inhibitors active in the sub-micro molar range were identified which provide the basis for further characterization of PKMYT1 as putative target for cancer therapy.

3.4. Identification of Highly Potent Protein KinaseC-Related Kinase1 Inhibitors by Virtual Screening, Binding Free Energy Rescoring, and invitro Testing

Slynko, I., Schmidtkunz, K., Rumpf, T., Klaeger, S., Heinzlmeir, S., Najjar, A., Metzger, E., Kuester, B., Schuele, R., Jung, M., Sippl, W.

ChemMedChem, 11(18), 2084-2094, 2016.

<https://doi.org/10.1002/cmdc.201600284>

Abstract: Despite the considerable interest in protein kinase C-related kinase 1 (PRK1) as a target in cancer research, there is still a lack of PRK1 inhibitors with suitable selectivity profiles and physicochemical properties. To identify new PRK1 inhibitors we applied a virtual screening approach, which combines ensemble docking, minimization of the protein–ligand complex, binding free energy calculations, and application of quantitative structure–activity relationship (QSAR) models for predicting in vitro activity. The developed approach was then applied in a prospective manner to screen available libraries of kinase inhibitors from Selleck and GlaxoSmithKline (GSK). Compounds that showed favorable prediction were then tested in vitro for PRK1 inhibition. Some of the hits were found to inhibit PRK1 in the low-nanomolar range. Three in vitro hits were additionally tested in a mass-spectrometry-based cellular kinase profiling assay to examine selectivity. Our findings show that nanomolar and drug-like inhibitors can be identified by the virtual screening approach presented herein. The identified inhibitors are valuable tools for gaining a better understanding of PRK1 inhibition, and the identified hits can serve as starting points for further chemical optimization.

3.5. Application of Computer Modeling to Drug Discovery: Case Study of PRK1 Kinase Inhibitors as Potential Drugs in Prostate Cancer Treatment

Najjar, A., Ntie-Kang, F., & Sippl, W.

In Unique Aspects of Anti-cancer Drug Development, *INTECHOPEN*, 2017.

<https://doi.org/10.5772/intechopen.68910>

Abstract: Computer modeling of natural products (NPs) and NP scaffolds is increasingly gaining importance in drug discovery, particularly in hit/lead discovery programs and at the lead optimization stage. Even though industry had lost interest in the implication of NPs in hit/lead searches, recent reports still show that computer modeling could be a useful assert for the identification of starting scaffolds from nature, which could be further exploited by synthetic modifications. In this chapter, the focus is on some useful tools for computer modeling aimed at the discovery of anticancer drugs from NP scaffolds. We also focus on some recent developments toward the identification of potential anticancer agents by the application of computer modeling. The chapter will lay emphasis on natural sources of anticancer compounds, present some useful databases and computational tools for anticancer drug discovery, and show some recent case studies of the application of computational modeling in anticancer drug discovery, as well as some success stories in virtual screening applications in anticancer drug discovery, highlighting some useful results on the application of on lead discovery (including promising NP scaffolds) against an interesting anticancer drug target, the protein kinase C-related kinase (PRK1).

3.6. Discovery of Dually Acting Small-molecule Inhibitors of Cancer-Resistance Relevant Receptor Tyrosine Kinases EGFR and IGF-1R

Hempel, C., Najjar, A., Totzke, F., Schaechtele, C., Sippl, W., Ritter, C., & Hilgeroth, A.

MedChemComm, 7(11), 2159-2166, 2016.

<https://doi.org/10.1039/c6md00329j>

Abstract: Novel benzo-anellated furo- and pyrrolo[2,3-b]pyridines with a 4-benzylamine substitution have been evaluated as inhibitors of the epidermal growth factor receptor (EGFR). Substituent effects on the determined protein kinase affinity have been discussed based on varied benzylamine residues at the differently substituted molecular scaffolds. Docking studies were carried out in order to explore the potential binding modes of the novel inhibitors. The observed activity data encouraged the measurement of the inhibition of the insulin-like growth factor receptor (IGF-1R), which is known to play an important role in the cancer resistance development against EGFR inhibitors via receptor heterodimerization with IGF-1R. We identified novel dual inhibitors of both kinases and report their first cancer cell growth inhibition data.

3.7. Discovery of Novel Dual Inhibitors of Receptor Tyrosine Kinases EGFR and IGF-1R

Hempel, C., Totzke, F., Schaechtele, C., Najjar, A., Sippl, W., Ritter, C., & Hilgeroth, A.

J. Enzyme Inhib. Med. Chem, 32(1), 271-276, 2017.

<https://doi.org/10.1080/14756366.2016.1247062>

Abstract: Novel 4-benzylamino benzo-anellated pyrrolo[2,3-b]pyridines have been synthesized with varied substitution patterns both at the molecular scaffold of the benzo-anellated ring and at the 4-benzylamino residue. With a structural similarity to substituted thieno[2,3-d]pyrimidines as epidermal growth factor receptor (EGFR) inhibitors, we characterized the inhibition of EGFR for our novel compounds. As receptor heterodimerization gained certain interest as mechanism of cancer cells to become resistant against novel protein kinase inhibitors, we additionally measured the inhibition of insulin-like growth factor receptor IGF-1R which is a prominent receptor for such heterodimerization with EGFR. Structure–activity relationships are discussed for both kinase inhibitions depending on the varied substitution patterns. We discovered novel dual inhibitors of both receptor tyrosine kinases with interest for further studies to reduce inhibitor resistance developments in cancer treatment.

3.8. Discovery of novel substituted benzo-anellated 4-benzylamino pyrrolopyrimidines as dual EGFR and VEGFR2 inhibitors

Fischer, T., Krueger, T., Najjar, A., Totzke, F., Schaechtele, C., Sippl, W., Ritter, C., Hilgeroth, A.

Bioorganic & Med. Chem. Lett, 27(12), 2708-2712, 2017.

<https://doi.org/10.1016/j.bmcl.2017.04.053>

Abstract: The quinazoline scaffold is the main part of many marketed EGFR inhibitors. Resistance developments against those inhibitors enforced the search for novel structural lead compounds. We developed novel benzo-anellated 4-benzylamine pyrrolopyrimidines with varied substitution patterns at both the molecular scaffold and the attached residue in the 4-position. The structure-dependent affinities towards EGFR are discussed and first nanomolar derivatives have been identified. Docking studies were carried out for EGFR in order to explore the potential binding mode of the novel inhibitors. As the receptor tyrosine kinase VEGFR2 recently gained an increasing interest as an upregulated signaling kinase in many solid tumors and in tumor metastasis we determined the affinity of our compounds to inhibit VEGFR2. So we identified novel dually acting EGFR and VEGFR2 inhibitors for which first anticancer screening data are reported. Those data indicate a stronger antiproliferative effect of a VEGFR2 inhibition compared to the EGFR inhibition.

3.9. Discovery of Novel Dual Inhibitors of Receptor Tyrosine Kinases EGFR and PDGFR-beta Related to Anticancer Drug Resistance

Fischer, T., Najjar, A., Totzke, F., Schaechtele, C., Sippl, W., Ritter, C., & Hilgeroth, A.

J Enzyme Inhib Med Chem, 33(1), 1-8, 2018.

<https://doi.org/10.1080/14756366.2017.1370583>

Abstract: With ongoing resistance problems against the marketed EGFR inhibitors having a quinazoline core scaffold there is a need for the development of novel inhibitors having a modified scaffold and, thus, expected lower EGFR resistance problems. An additional problem concerning EGFR inhibitor resistance is an observed heterodimerization of EGFR with PDGFR-b that neutralises the sole inhibitor activity towards EGFR. We developed novel pyrimido[4,5-b]indoles with varied substitution patterns at the 4-anilino residue to evaluate their EGFR and PDGFR-b inhibiting properties. We identified dual inhibitors of both EGFR and PDGFR-b in the nanomolar range which have been initially screened in cancer cell lines to prove a benefit of both EGFR and PDGFR-b inhibition.

SUMMARY

Kinases represent important drug targets of the current century as Philp Cohen titled his paper “Protein kinases - the major drug targets of the twenty-first century?” [160]. Kinases play key regulatory roles in most aspects of cell biology. The dysregulation of protein kinases occurs in a variety of diseases including cancer, diabetes, autoimmune, cardiovascular, inflammatory, and neurological disorders. Protein kinases represent one of the most important classes of drug targets. In recent times, efforts have been paid in academia and industry to determine the physiological and pathological functions of different kinases. Concomitantly, the discovery of kinase inhibitors as drugs has significantly accelerated within the last decade, leading to ~ 40 small molecule kinase inhibitors approved by the FDA. Rational drug design has succeeded in 2002 to design the first approved kinase inhibitor (imatinib) for the treatment of chronic myelogenous leukemia.

In the present work, we contributed to the kinase drug discovery by targeting several kinase members. We were able to use computational approaches in combination with medicinal chemistry to identify novel kinase inhibitors which inhibit the active form of the studied kinases. Interestingly, the identified inhibitors have been observed to inhibit the cell growth in some cancer cell lines. The following subsections provide summaries as well as highlights of the main achievements from these studies.

4.1. Structural analysis of WEE family kinase

PKMYT1 and WEE1 are members of the WEE family kinases and play critical roles in DNA-damage recovery. In the present work, a comparison of the WEE family kinase members was performed to recognize the structural diversity, to establish the structural basis of different inhibitors and to present biological assays used to identify WEE family kinase inhibitors. For structural analysis, the deposited crystal structures of PKMYT1 and WEE1 in PDB were retrieved and analyzed. The deposited X-ray structures of WEE1 are co-crystallized with series of pyrrolocarbazoles and pyrroloindoles derivatives whereas PKMYT1 was resolved in complex with known kinase inhibitors e.g. dasatinib, saracatinib, and MK-1775. The ATP-binding pocket reveals high similarity (sequence similarity: 76.2%, sequence identity: 61.9%, RMSD deviation of the backbone atoms between Wee1 and PKMYT1: 0.88 Å). Notably, a bulky gatekeeper residue in WEE1 (Asn376) was found to restrict the back pocket. In contrast, PKMYT1 has a small gatekeeper residue (Thr178), which allows an access to the back hydrophobic-pocket (**Figure 9**). The amino acids at the top of the P-loop (WEE1: Phe310, PKMYT1: Tyr121) can adopt different conformations based on the activation state of the kinase. The conserved cysteine of the hinge region (WEE1: Cys379, PKMYT1: Cys190) plays an important role in stabilizing the ATP and the inhibitors via the formation of hydrogen bonds (**Figure 9**).

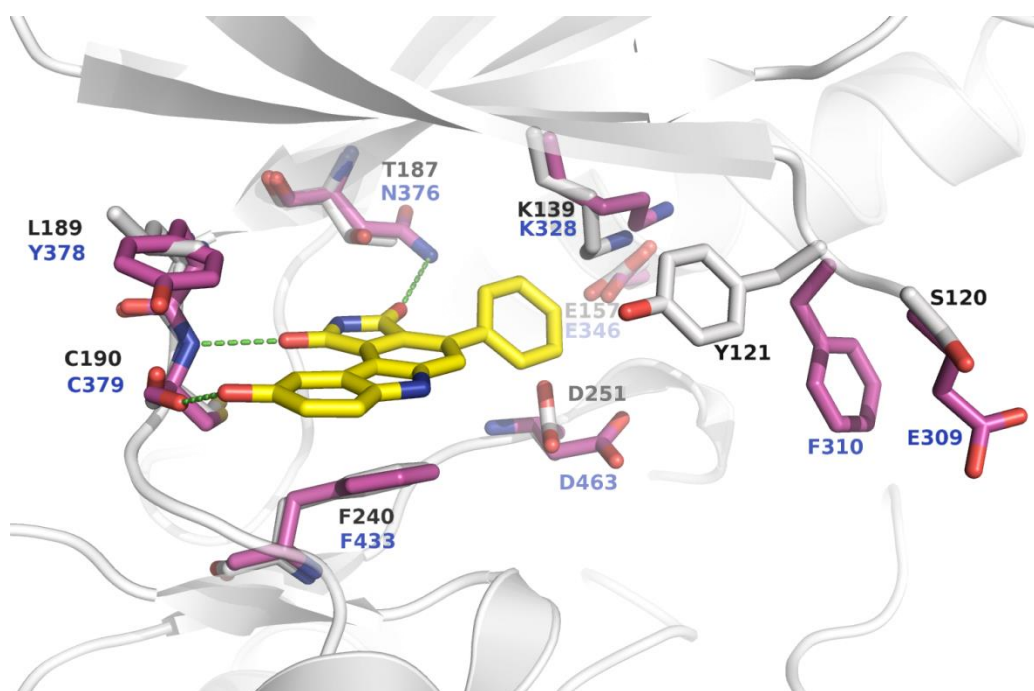


Figure 9. Superposition of the ATP-binding pocket of WEE1 (PDB ID: 1X8B, violet residues and blue label) and PKMYT1 (PDB ID: 3P1A, white residues and black label). The co-crystallized inhibitor PD0407824 of WEE1 is shown in yellow. Green dashed lines refer to hydrogen bonds.

Thus, the observed structural similarity and diversity of PKMYT1 and WEE1 might explain its substrate specificity to phosphorylate either Tyr15 of CDK1 by WEE1 or Tyr15 and Thr14 of CDK1 by PKMYT1. The structural analysis of PKMYT1 and WEE1 facilitated also the design of selective inhibitors in order to assess their potential for cancer treatment (**Publication 1**).

4.2. Hit identification and lead optimization for PKMYT1

In order to identify novel inhibitors for PKMYT1, focused libraries were screened. The kinase inhibitor data sets I and II from GSK (in total 800 known kinase inhibitors) were biologically screened using a fluorescence polarization based assay. **Figure 10** summarizes the working process of the study. An initial screening using a binding assay at 20 μ M was first performed to identify the hits, which showed a displacement percentage above a specific threshold. Then, the K_i and IC_{50} values were determined. This was followed by an activity assay as a means to measure the inhibitory activity of the identified hits by measuring K_i and IC_{50} values. Additionally, a cell line was used to validate the activity of the identified hits in a cellular assay. Hence, several hits were identified as nanomolar inhibitors, thus, providing new scaffolds, which were used to develop novel selective PKMYT1 inhibitors (**Figure 10**). In order to rationalize the binding mode of the identified inhibitors, cross-docking to all currently reported crystal structures of PKMYT1 was performed. The aim of this cross-docking procedure was to validate different docking algorithms (GOLD and

Glide), and to assess which protein structure is the best for further docking studies. The results revealed that docking in the X-ray structure of PKMYT1 cocrystallized with pelitinib (PDB ID: 5VCW) could successfully reproduce the binding mode of native as well as non-native ligands. Thus, it was used to predict the binding mode of the new hits. The predicted binding modes were further validated by comparing the generated poses to those of available PDB-structures with similar co-crystallized derivatives. Despite the chemical diversity of the new hits, they all form a hydrogen bond with the backbone of conserved hinge region residue Cys190 similar to the ATP (**Figure 11**). Additional interactions such as aromatic interactions (π - π) with the side chain of Phe240, targeting the hydrophobic back-pocket by a substituted phenyl or small chemical group, interactions with the front-pocket residues e.g. Leu116, Pro191 and Gln196, and polar interactions with Asp251 and Tyr121 were observed in most of the PKMYT1 inhibitors. The results obtained, offered an opportunity for structure-activity relationship study and guided further optimization of the identified hits, Thereby, paving the way for the design of novel and selective PKMYT1 inhibitors (**Publication 2**).

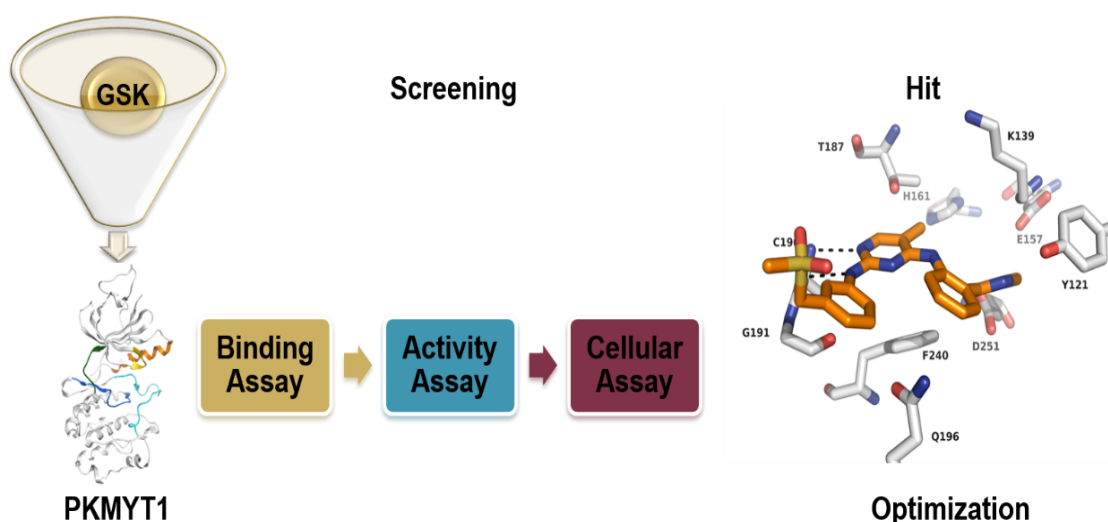


Figure 10. Scheme of the process to identify PKMYT1 inhibitors.

Structure-based optimization of a series of hits containing a diaminopyrimidine scaffold was carried out to improve the potency and selectivity against PKMYT1. Therefore, a computational model able to predict the affinity of the novel leads was developed. The current study confirmed the ability of the docking algorithm to predict the binding mode of the small molecules. However, the docking algorithm failed to rightly predict the affinity of the ligands, where a weak correlation was obtained between the computed and experimental data. We were able to show that binding free energy calculation (BFE), here using the MM-PB/SA solvation model, was better able to predict the affinity

of the ligands. Moreover, including 2D molecular descriptors based on the partial charges of the ligands improved the quality of the prediction slightly.

A QSAR model was derived by combining the calculated binding free energy and the molecular descriptors in order to predict the activities of the novel compounds. Regression-based QSAR modeling was not able to discriminate between the active and inactive. However, it was able to predict the improvement of the potency for the lead optimization. Fragment growing approach is an efficient method to achieve additional interactions within the binding pocket. Other factors should be considered during the growing process e.g. ligand flexibility, conformational changing, and steric hindrance. Our approach guided the design and synthesis of the first set of diaminopyrimidine derivatives, some of which showed potency in the sub-micromolar range (e.g. compound **5k**). As a consequence, we were able to generate several lead compounds and to assess the structure-activity relationship of the diaminopyrimidine derivatives. The derived results provide the basis for further chemical modifications of the diaminopyrimidine derivatives to design more potent PKMYT1 inhibitors. This allows further analysis of PKMYT1 as target for cancer therapy (**Publication 3**).

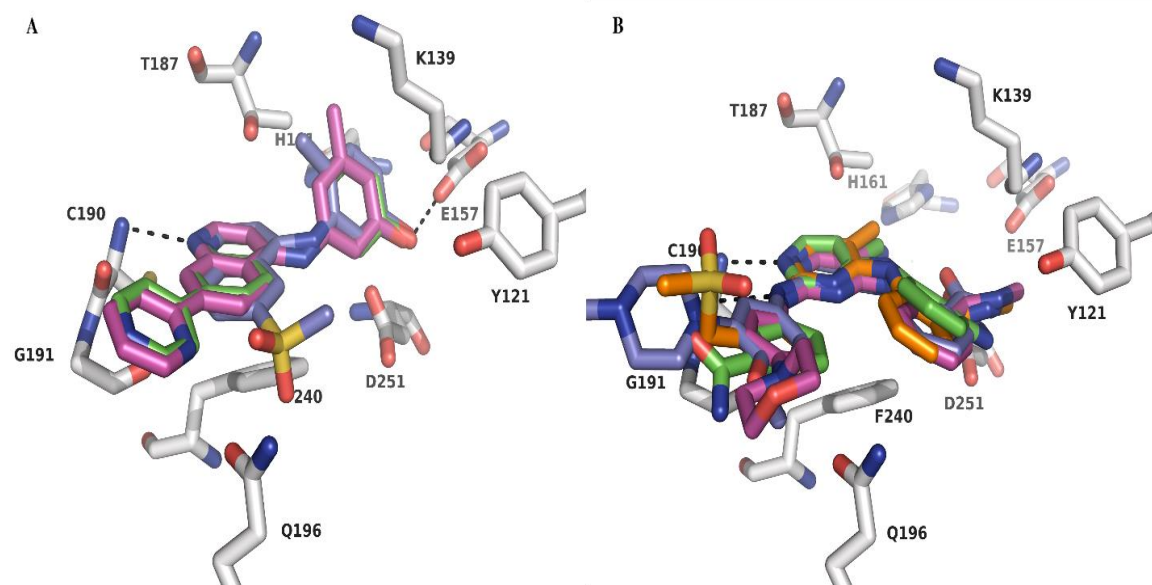


Figure 11. Determined binding modes of the identified PKMYT1 inhibitors. A) Putative binding modes of aminoquinoline compounds; B): Putative binding modes of diaminopyrimidine series.

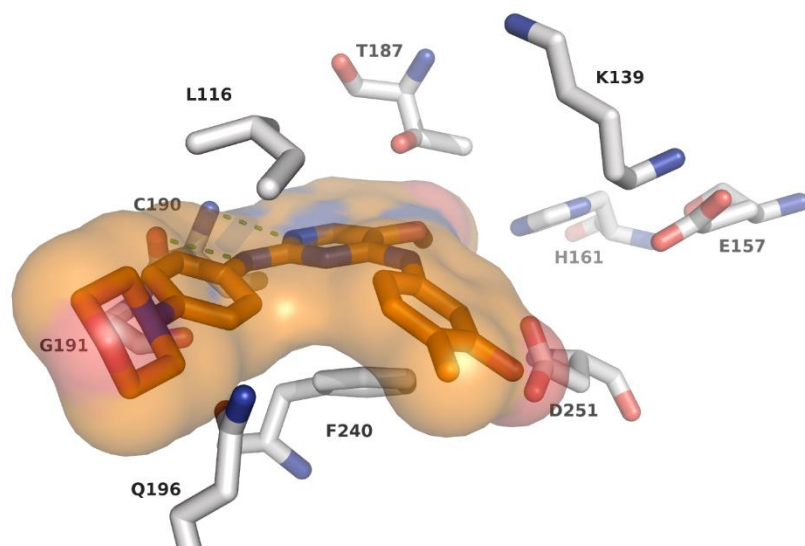


Figure 12. Predicted binding mode of compound **5k** (green) within the PKMYT1 binding pocket (white). Molecular surface (green) display the filled volume of the PKMYT1 pocket by compound **5k**.

4.3. Virtual screening of commercial data sets on PRK1

Protein kinase C-related kinase 1 (PRK1) is a serine/threonine enzyme, which is a regulator of the androgen receptors. In this work, a successful virtual screening of GSK and Selleckchem libraries was performed, leading to the identification of highly potent and novel PRK1 inhibitors. The identified hits are ATP-competitive and bind to the active form of PRK1. While all of them show the conserved interactions with the hinge region residues, the most potent PRK1 inhibitors either form additional hydrogen bonds to polar residues in the ATP-sugar binding region or interact with the hydrophobic back-pocket behind the gatekeeper residue (Met701). The PRK1 inhibitors identified in the current work as well as their predicted binding modes can be further used to guide structure-based optimization. In addition, the obtained results reveal that the combination of molecular docking with post-processing using BFE calculation might help to compensate the disadvantages of each individual method and to improve the screening success. Application of this combined method was able to identify novel nanomolar PRK1 inhibitors. This combined method was also effectively used to prioritize PRK1 hits for further testing. However, the presence of false positives among the selected hits could be due to the ability of the protein to accommodate bigger ligands during the minimization step, or the introduction of larger errors when the energy of the receptor is taken into account. Nevertheless, the rescoring of docking poses using the BFE-based QSAR model performed better than docking scores. For example, only one weakly active compound was found among the ten top-ranked by GlideSP score. Interestingly, the generated QSAR model was able to rank two highly potent inhibitors (with IC_{50} less than 100 nM) among the top-six GSK compounds according to predicted pIC_{50} values. Since the focused databases are

composed of kinase inhibitors, this may explain the good hit rates of the virtual screenings (100% for Selleckchem and 32% for GSK kinase inhibitor set). It was also remarkable that, despite the structural diversity of the GSK kinase inhibitor set, most of the compounds identified as highly-potent PRK1 inhibitors contain an aminofurazan scaffold (**Figure 13 and 14**) which were initially developed as potent inhibitors of AKT and GSK-3 kinases. A number of the identified compounds are also known to target other members of the AGC kinase family. For example, the pan-Akt inhibitor GSK-690693 (ChEMBL494089) is also inhibits PKA, PRK or PKC isozymes.

Analysis of the later released crystal structures of PRK1 in complex some identified inhibitors revealed that the generated homology models could be successfully used to re-predict the binding modes of most known ligand, nevertheless, deviations between the predicted and experimental binding modes of some ligands could be observed. Several factors can have impact on the molecular docking performance, e.g. model quality, absence/presence of water, receptor and ligand flexibility. The comparison of PRK1 homology models with the crystal structures and the analysis of inhibitor binding modes highlighted the important role of kinase domain flexibility, which should be considered for the design and optimization of novel PRK1 inhibitors (**Publication 4 and 5**).

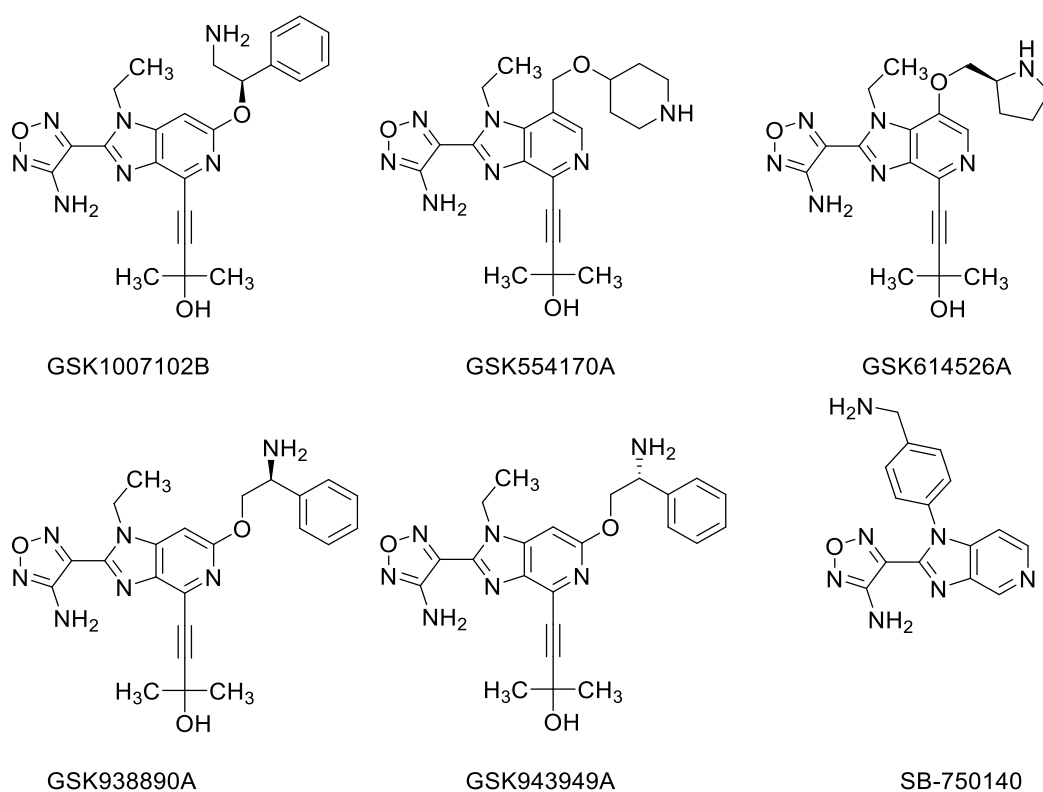


Figure 13. Identified PRK1-inhibitors from GSK PKIS set.

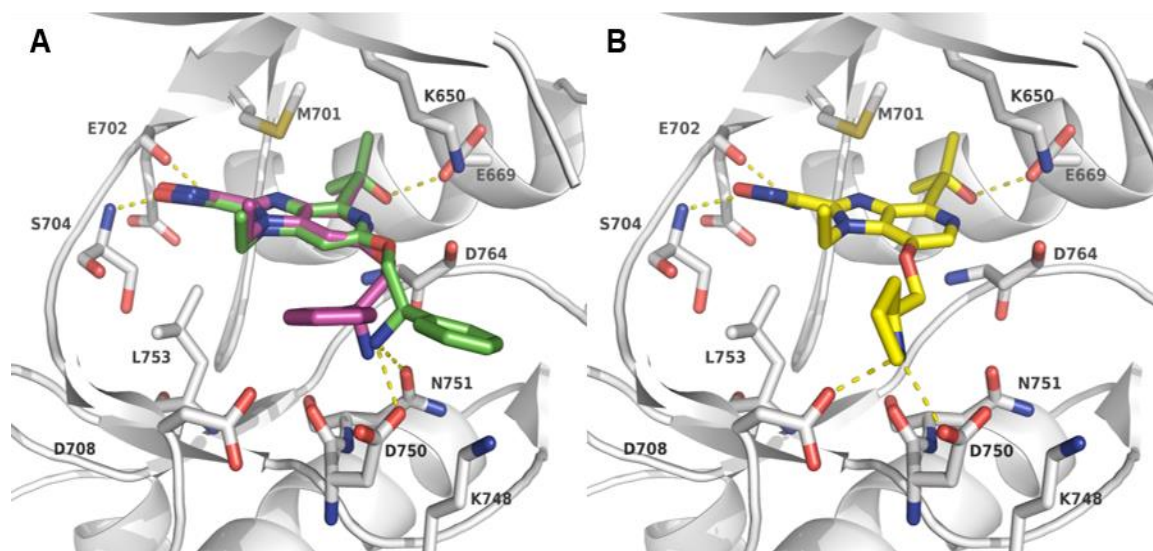


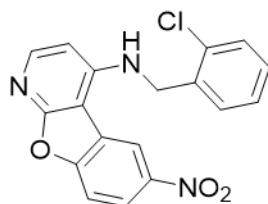
Figure 14. Predicted binding mode for three of the most active PRK1 inhibitors from the GSK kinase set. A) *R*-isomer GSK-943949A (green carbons) and the corresponding *S*-isomer, GSK-938890A (magenta carbons); B) GSK-614526A (yellow carbons).

4.4. Discovery of novel growth factors receptors inhibitors

Expansion and progression of cancer cells depend on the activation of transmembrane receptors harboring kinase activity such as EGFR, IGF-1R, VEGFR2, and PDGFR- β by a variety of growth factors. Thus, inhibition of the receptors of the growth factors offers opportunities for cancer therapy.

In this project, the development of a series of pyrrolopyridine derivatives was rationalized by carrying out molecular docking of the synthesized derivatives. Thus, first dual inhibitors of EGFR and IGF-1R with sub-micromolar and nanomolar activities were identified, which showed a growth inhibition in the EGFR-related breast and non-small lung cancer cell lines making them attractive compounds for further drug development. Novel furo- and pyrrolopyridines inhibit EGFR and IGF-1R activities as ATP-competitive inhibitors. Analysis of the ligand binding of these derivatives within EGFR and IGF-1R were performed in order to provide a potential rationalization for the detected *in vitro* activity. The benzo-anellated furo[2,3-*b*]pyridine and benzo-anellated pyrrolo[2,3-*b*]pyridine scaffolds mimic the adenine ring of ATP and represent hinge-binding motifs. The variation in the gatekeeper size of both kinases (Thr790 in EGFR and Met1079 in IGF-1R) results in different orientation of the active derivatives (**Figure 15**). It is worth mentioning that the diversity of the gatekeeper residues and the accessibility to the hydrophobic back-pocket are mainly responsible for the observed selectivity of some inhibitors. The resulting selectivity of the active compounds were later evaluated against a variety of EGFR-related kinases including HER2 and HER4 as well as the related receptor tyrosine kinases JAK2 and 3, and TIE-2 plus other kinase

families like VEGFR2 and 3, PDGFR- β and GSK-3 β . Interestingly, no the compounds showed no inhibition of these related kinases even at concentrations >100 μ M (**Publication 6**).



13b

EGFR: $K_i = 80$ nM

IGF-1R: $K_i = 280$ nM

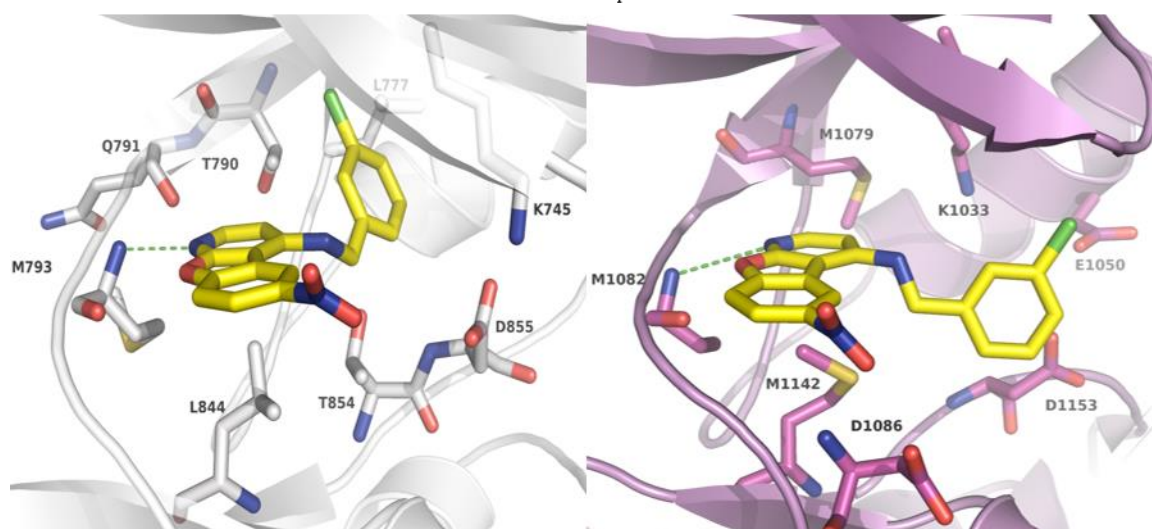


Figure 15. Putative binding mode of the most potent inhibitor (**13b**) in EGFR (light grey) and IGF-1R (violet).

The results obtained from the previous analysis of ligand binding served as guidance to develop the first dual-inhibitors of EGFR and VEGFR2 which showed an inhibitory activity in the nanomolar range. This series contains benzo-anellated pyrrolopyrimidines, which showed favorable *para*- and *ortho*-substitutions of the 4-benzylamino residue (**Figure 16**: compounds **6m** and **6f**). The most potent VEGFR2 inhibitor showed the best anticancer activity. The result reveals the importance of VEGFR2 inhibition for inhibiting cancer cell growth in VEGFR2-relevant tumors. Once again, the selectivity of the novel inhibitors was evaluated and no activity found against MARK1, DARK1, MEK1, MEK2, c-ABL, MLK4 and others. (**Publication 7**)

The third series was found to be dual-inhibitors of the tyrosine kinase receptors EGFR and IGF-1R. The novel active compounds are 6-cyano derivatives of the benzylaminopyrrolopyrimidines and contain favorable 3-amino and the 4-methyl benzamine moieties (**Publication 8**). These novel compounds showed activities in the low nanomolar range against EGFR and in the sub-micromolar range against IGF-1R (**Figure 16**: compound **6a**).

The last series of the 4-anilinopyrimido[4,5-*b*]indole derivatives were investigated against EGFR and PDGFR- β (**Figure 16**: compound **5k**). The molecular docking produced a binding mode that mimics the adenine ring of ATP and displays a conserved hydrogen bond with the hinge region residue. It was found that attaching large substituents at the *para*-position of the aniline residue enhance the activity against of EGFR and PDGFR- β . These dual-inhibitors with favorable EGFR-affinities showed an effective anticancer growth inhibition (**Publication 9**).

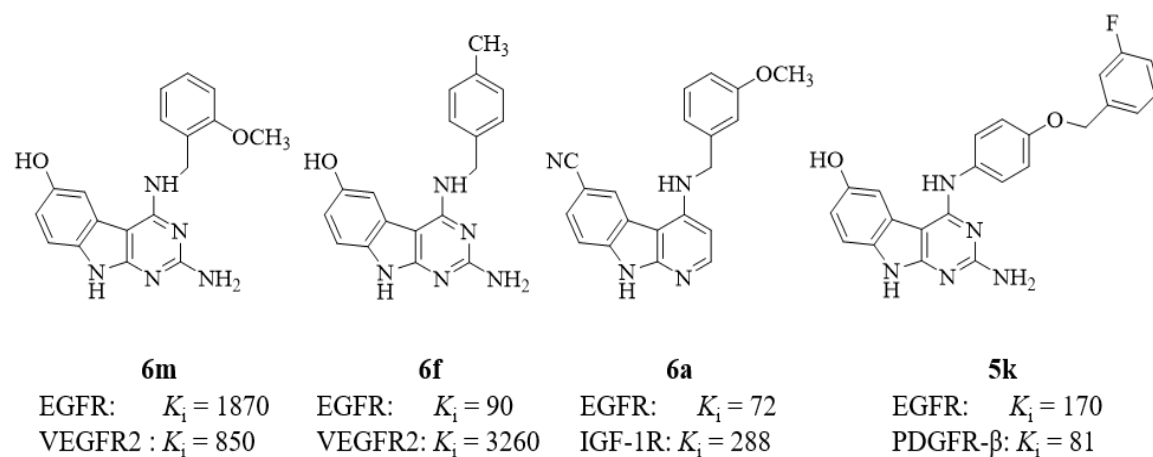


Figure 16. Chemical structures of identified dual-inhibitors of the growth factors receptors (K_i , nM).

CONCLUSIONS

The current study shows the importance of rational drug design to identify novel inhibitors targeting the human kinome. The digitalization of the drug discovery became incredibly widely implemented to accelerate the development of drug candidates. Structure-based drug design is considered as an efficient method to design selective and highly active kinase inhibitors.

Structural analysis of the 3D structures of the kinase domains proved to be a very helpful procedure for studying kinases and developing kinase inhibitors. For instance, analyzing the binding pocket of the studied kinases offered reasonable explanations of the observed activity of the identified hits. Moreover, the analysis of the site map aided to find the relevant residues that interact with the ligands, which helped guide hit optimization studies.

A combination between *in silico* and *in vitro* screening was found as an optimal approach to accelerate the discovery of novel kinase inhibitors. For instance, VS screening of GSK and Selleckchem databases in homology models of PRK1 followed by *in vitro* testing of the selected hits resulted in the identification of nanomolar PRK1 inhibitors. Meanwhile, various computational drug discovery tools assisted us to rationalize the lead optimization process so as to improve the potency of the obtained PKMYT1 leads. Here, the derived computational results plus the structural information of the binding pocket were implemented to chemically modify the lead compounds, to enhance the inhibitory potency towards PKMYT1.

In this study, we could show the ability of the molecular docking to correctly predict the bioactive conformation of small molecules, albeit some limitations of the rigid docking approach were also encountered. Through this approach, we could successfully predict the binding modes of the obtained kinase inhibitors, like PRK1- and PKMYT-inhibitors. Further, we were able to explain the obtained activities and selectivity of developed inhibitors towards different kinase receptors of the growth factors. Indeed, molecular docking offered rational insights to develop highly active and selective lead compounds, and the derived results provided the basis for further chemical modifications of the identified hits to design highly potent kinase inhibitors allowing further analysis of kinases as targets for the cancer therapy

Despite their good performance of the docking software in predicting the binding modes of the obtained inhibitors, docking scores failed, as is often the case, to predict the affinity of the ligands, resulting in a weak correlation between the computed and experimental affinities. Hence, the search for accurate predicting methods was necessary to guide the lead optimization studies. All things considered, we were able to show, as exemplified by our studies on PKMT1- and PRK1-inhibitors, the higher ability of the binding free energy calculations to predict the affinity of the ligands as

compared to the docking scores. Moreover, including 2D molecular descriptors based on the partial charges of the ligands improved the quality of the prediction slightly.

Generation of several homology models also proved to be a useful tool to generate 3D structures of kinases, with no resolved crystal structure, and to consider the flexibility of the kinase domains. Here, molecular dynamics (MD) simulations were used to address and evaluate the conformational changes upon ligand binding. The highly computational costs of running MD simulations, however, restrict its application only for a small set of compounds.

Finally, computational drug discovery and design for the human kinase showed several advantages in guiding the development of novel hits and some limitations. The conserved structure of the kinase domain makes it easy to design highly potent kinase inhibitors, for both the resolved and non-resolved kinase structures. So, structure-based drug design can be performed based on the knowledge of the primary sequences for a target kinase and a homologous kinase structure. However, a huge caveat is encountered with the issue of selectivity. Generally, to design selective ATP-competitive inhibitors, the adjacent pockets of the ATP-binding pocket should be addressed. Since the size of the gatekeeper has a big impact on the accessibility to the hydrophobic back-pocket and can restrict the size of this pocket, it is therefore logical to be considered during the lead optimizations process. The interactions with the P-loop residues can enhance the activity and the selectivity, but the P-loop adopts several conformations depending on the size of the bound ligands, which may have a big impact on the ATP-binding pocket conformation. Some of the identified compounds showed only low or moderate activity. Another issue is of course, the kinase plasticity, which reveals several conformations based on the activation state. Protein plasticity still remains a significant limitation in the computational methods of drug discovery.

REFERENCES

1. Hopkins, A.L. and C.R. Groom, *The druggable genome*. Nat Rev Drug Discov, 2002. **1**(9): p. 727-30.
2. Drews, J., *Drug discovery: a historical perspective*. Science, 2000. **287**(5460): p. 1960-4.
3. Gashaw, I., et al., *What makes a good drug target?* Drug Discov Today, 2011. **16**(23-24): p. 1037-43.
4. Imming, P., C. Sinning, and A. Meyer, *Drugs, their targets and the nature and number of drug targets*. Nat Rev Drug Discov, 2006. **5**(10): p. 821-34.
5. Lipinski, C.A., et al., *Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings*. Advanced Drug Delivery Reviews, 2012. **64**: p. 4-17.
6. Kinch, M.S., et al., *Target selection for FDA-approved medicines*. Drug Discovery Today, 2015. **20**(7): p. 784-789.
7. Santos, R., et al., *A comprehensive map of molecular drug targets*. Nat Rev Drug Discov, 2017. **16**(1): p. 19-34.
8. Overington, J.P., B. Al-Lazikani, and A.L. Hopkins, *How many drug targets are there?* Nature Reviews Drug Discovery, 2006. **5**: p. 993.
9. Rask-Andersen, M., S. Masuram, and H.B. Schioth, *The druggable genome: Evaluation of drug targets in clinical trials suggests major shifts in molecular class and indication*. Annu Rev Pharmacol Toxicol, 2014. **54**: p. 9-26.
10. Manning, G., et al., *The Protein Kinase Complement of the Human Genome*. Science, 2002. **298**(5600): p. 1912.
11. Rask-Andersen, M., et al., *Advances in kinase targeting: current clinical use and clinical trials*. Trends in Pharmacological Sciences. **35**(11): p. 604-620.
12. Zhang, J., P.L. Yang, and N.S. Gray, *Targeting cancer with small molecule kinase inhibitors*. Nature Reviews Cancer, 2009. **9**: p. 28.
13. Sun, C. and R. Bernards, *Feedback and redundancy in receptor tyrosine kinase signaling: relevance to cancer therapies*. Trends in Biochemical Sciences, 2014. **39**(10): p. 465-474.
14. Huang, M., et al., *Molecularly targeted cancer therapy: some lessons from the past decade*. Trends in Pharmacological Sciences. **35**(1): p. 41-50.
15. Clark, J.D., M.E. Flanagan, and J.-B. Telliez, *Discovery and Development of Janus Kinase (JAK) Inhibitors for Inflammatory Diseases*. Journal of Medicinal Chemistry, 2014. **57**(12): p. 5023-5038.
16. Barnes, P.J., *New anti-inflammatory targets for chronic obstructive pulmonary disease*. Nature Reviews Drug Discovery, 2013. **12**: p. 543.
17. Muth, F., et al., *Tetra-Substituted Pyridinylimidazoles As Dual Inhibitors of p38 α Mitogen-Activated Protein Kinase and c-Jun N-Terminal Kinase 3 for Potential Treatment of Neurodegenerative Diseases*. Journal of Medicinal Chemistry, 2015. **58**(1): p. 443-456.
18. Bové, J., M. Martínez-Vicente, and M. Vila, *Fighting neurodegeneration with rapamycin: mechanistic insights*. Nature Reviews Neuroscience, 2011. **12**: p. 437.
19. Kim, E.K. and E.-J. Choi, *Pathological roles of MAPK signaling pathways in human diseases*. Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease, 2010. **1802**(4): p. 396-405.
20. Kikuchi, R., et al., *An antiangiogenic isoform of VEGF-A contributes to impaired vascularization in peripheral artery disease*. Nature Medicine, 2014. **20**: p. 1464.

21. Banks, A.S., et al., *An ERK/Cdk5 axis controls the diabetogenic actions of PPAR γ* . Nature, 2014. **517**: p. 391.
22. Pennisi, E., *ENCODE Project Writes Eulogy for Junk DNA*. Science, 2012. **337**(6099): p. 1159-1161.
23. Brognard, J. and T. Hunter, *Protein kinase signaling networks in cancer*. Curr Opin Genet Dev, 2011. **21**(1): p. 4-11.
24. Roskoski, R., Jr., *Classification of small molecule protein kinase inhibitors based upon the structures of their drug-enzyme complexes*. Pharmacol Res, 2016. **103**: p. 26-48.
25. Hanks, S.K. and T. Hunter, *Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification*. Faseb j, 1995. **9**(8): p. 576-96.
26. Manning, G., et al., *Evolution of protein kinase signaling from yeast to man*. Trends Biochem Sci, 2002. **27**(10): p. 514-20.
27. Knighton, D.R., et al., *Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase*. Science, 1991. **253**(5018): p. 407.
28. Hanks, S.K., A.M. Quinn, and T. Hunter, *The protein kinase family: conserved features and deduced phylogeny of the catalytic domains*. Science, 1988. **241**(4861): p. 42-52.
29. Kannan, N., et al., *Structural and functional diversity of the microbial kinome*. PLoS Biol, 2007. **5**(3): p. e17.
30. Scott, J.D. and T. Pawson, *Cell Signaling in Space and Time: Where Proteins Come Together and When They're Apart*. Science, 2009. **326**(5957): p. 1220-1224.
31. Jin, J. and T. Pawson, *Modular evolution of phosphorylation-based signalling systems*. Philosophical Transactions of the Royal Society B-Biological Sciences, 2012. **367**(1602): p. 2540-2555.
32. Cowan-Jacob, S.W., W. Jahnke, and S. Knapp, *Novel approaches for targeting kinases: allosteric inhibition, allosteric activation and pseudokinases*. Future Med Chem, 2014. **6**(5): p. 541-61.
33. Taylor, S.S. and A.P. Kornev, *Protein kinases: evolution of dynamic regulatory proteins*. Trends Biochem Sci, 2011. **36**(2): p. 65-77.
34. Nolen, B., S. Taylor, and G. Ghosh, *Regulation of protein kinases; controlling activity through activation segment conformation*. Mol Cell, 2004. **15**(5): p. 661-75.
35. Cowan-Jacob, S.W., *Structural biology of protein tyrosine kinases*. Cell Mol Life Sci, 2006. **63**(22): p. 2608-25.
36. Schmidt, M., et al., *Regulation of G2/M Transition by Inhibition of WEE1 and PKMYT1 Kinases*. Molecules, 2017. **22**(12).
37. Fabbro, D., S.W. Cowan-Jacob, and H. Möbitz, *Ten things you should know about protein kinases: IUPHAR Review 14*. British Journal of Pharmacology, 2015. **172**(11): p. 2675-2700.
38. Ubersax, J.A. and J.E. Ferrell, Jr., *Mechanisms of specificity in protein phosphorylation*. Nat Rev Mol Cell Biol, 2007. **8**(7): p. 530-41.
39. Möbitz, H. and D. Fabbro, *Conformational Bias: A key concept for protein kinase inhibition*. Vol. 17. 2012.
40. Kornev, A.P., et al., *Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism*. Proc Natl Acad Sci U S A, 2006. **103**(47): p. 17783-8.
41. Cowan-Jacob, S.W., H. Möbitz, and D. Fabbro, *Structural biology contributions to tyrosine kinase drug discovery*. Current opinion in cell biology, 2009. **21**(2): p. 280-287.
42. Zhu, J.-Y., et al., *Structural basis of Wee kinases functionality and inactivation by diverse small molecule inhibitors*. Journal of Medicinal Chemistry, 2017: p. acs.jmedchem.7b00996-acs.jmedchem.7b00996.

43. Tong, M. and M.A. Seeliger, *Targeting Conformational Plasticity of Protein Kinases*. ACS Chemical Biology, 2015. **10**(1): p. 190-200.
44. Kornev, A.P., S.S. Taylor, and L.F. Ten Eyck, *A helix scaffold for the assembly of active protein kinases*. Proceedings of the National Academy of Sciences, 2008. **105**(38): p. 14377.
45. Oruganty, K., et al., *Identification of a hidden strain switch provides clues to an ancient structural mechanism in protein kinases*. Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(3): p. 924-929.
46. Chen, H., et al., *A molecular brake in the kinase hinge region regulates the activity of receptor tyrosine kinases*. Mol Cell, 2007. **27**(5): p. 717-30.
47. Stamos, J., M.X. Sliwkowski, and C. Eigenbrot, *Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor*. J Biol Chem, 2002. **277**(48): p. 46265-72.
48. Zhou, T., et al., *Structural mechanism of the Pan-BCR-ABL inhibitor ponatinib (AP24534): lessons for overcoming kinase inhibitor resistance*. Chem Biol Drug Des, 2011. **77**(1): p. 1-11.
49. Park, J.H., et al., *Erlotinib binds both inactive and active conformations of the EGFR tyrosine kinase domain*. Biochem J, 2012. **448**(3): p. 417-23.
50. Wu, P., T.E. Nielsen, and M.H. Clausen, *FDA-approved small-molecule kinase inhibitors*. Trends Pharmacol Sci, 2015. **36**(7): p. 422-39.
51. Shibuya, M. and Y. Suzuki, *[Treatment of cerebral vasospasm by a protein kinase inhibitor AT 877]*. No To Shinkei, 1993. **45**(9): p. 819-24.
52. Shibuya, M., T. Asano, and Y. Sasaki. *Effect of Fasudil HCl, a Protein Kinase Inhibitor, on Cerebral Vasospasm*. in *Cerebral Vasospasm*. 2001. Vienna: Springer Vienna.
53. Vasquez, E.M., *Sirolimus: a new agent for prevention of renal allograft rejection*. Am J Health Syst Pharm, 2000. **57**(5): p. 437-48; quiz 449-51.
54. Kelly Patrick, A., et al., *Sirolimus, a New, Potent Immunosuppressive Agent*. Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy, 2012. **17**(6): p. 1148-1156.
55. Traxler, P., et al., *Tyrosine kinase inhibitors: from rational design to clinical trials*. Med Res Rev, 2001. **21**(6): p. 499-512.
56. Buchdunger, E., A. Matter, and B.J. Druker, *Bcr-Abl inhibition as a modality of CML therapeutics*. Biochim Biophys Acta, 2001. **1551**(1): p. M11-8.
57. Dar, A.C. and K.M. Shokat, *The Evolution of Protein Kinase Inhibitors from Antagonists to Agonists of Cellular Signaling*. Annual Review of Biochemistry, 2011. **80**(1): p. 769-795.
58. Monod, J., J.-P. Changeux, and F. Jacob, *Allosteric proteins and cellular control systems*. Journal of Molecular Biology, 1963. **6**(4): p. 306-329.
59. Zuccotto, F., et al., *Through the "Gatekeeper Door": Exploiting the Active Kinase Conformation*. Journal of Medicinal Chemistry, 2010. **53**(7): p. 2681-2694.
60. Gavrin, L.K. and E. Saiah, *Approaches to discover non-ATP site kinase inhibitors*. Med. Chem. Commun., 2013. **4**(1): p. 41-51.
61. Vandana, L. and G. Indraneel, *New Directions in Targeting Protein Kinases: Focusing Upon True Allosteric and Bivalent Inhibitors*. Current Pharmaceutical Design, 2012. **18**(20): p. 2936-2945.
62. Furman, R.R., et al., *Idelalisib and Rituximab in Relapsed Chronic Lymphocytic Leukemia*. New England Journal of Medicine, 2014. **370**(11): p. 997-1007.
63. Ali, K., et al., *Inactivation of PI(3)K p110 δ breaks regulatory T-cell-mediated immune tolerance to cancer*. Nature, 2014. **510**: p. 407.
64. Somoza, J.R., et al., *Structural, biochemical, and biophysical characterization of idelalisib binding to phosphoinositide 3-kinase delta*. J Biol Chem, 2015. **290**(13): p. 8439-46.

65. Wu, P. and Y. Hu, *Small molecules targeting phosphoinositide 3-kinases*. *MedChemComm*, 2012. **3**(11).
66. Sinko, W., S. Lindert, and J.A. McCammon, *Accounting for receptor flexibility and enhanced sampling methods in computer-aided drug design*. *Chem Biol Drug Des*, 2013. **81**(1): p. 41-9.
67. Eriksson, M., et al., *SARConnect: A Tool to Interrogate the Connectivity Between Proteins, Chemical Structures and Activity Data*. *Mol Inform*, 2012. **31**(8): p. 555-568.
68. Muresan, S., et al., *Making every SAR point count: the development of Chemistry Connect for the large-scale integration of structure and bioactivity data*. *Drug Discov Today*, 2011. **16**(23-24): p. 1019-30.
69. McInnes, C., *Virtual screening strategies in drug discovery*. *Curr Opin Chem Biol*, 2007. **11**(5): p. 494-502.
70. Brooijmans, N. and I.D. Kuntz, *Molecular recognition and docking algorithms*. *Annu Rev Biophys Biomol Struct*, 2003. **32**: p. 335-73.
71. Kellenberger, E., et al., *Comparative evaluation of eight docking tools for docking and virtual screening accuracy*. *Proteins*, 2004. **57**(2): p. 225-42.
72. Cross, J.B., et al., *Comparison of several molecular docking programs: pose prediction and virtual screening accuracy*. *J Chem Inf Model*, 2009. **49**(6): p. 1455-74.
73. Ferrara, P., et al., *Assessing scoring functions for protein-ligand interactions*. *J Med Chem*, 2004. **47**(12): p. 3032-47.
74. Kitchen, D.B., et al., *Docking and scoring in virtual screening for drug discovery: methods and applications*. *Nat Rev Drug Discov*, 2004. **3**(11): p. 935-49.
75. Warren, G.L., et al., *A critical assessment of docking programs and scoring functions*. *J Med Chem*, 2006. **49**(20): p. 5912-31.
76. Scior, T., et al., *Recognizing Pitfalls in Virtual Screening: A Critical Review*. *Journal of Chemical Information and Modeling*, 2012. **52**(4): p. 867-881.
77. Kirchmair, J., et al., *Evaluation of the performance of 3D virtual screening protocols: RMSD comparisons, enrichment assessments, and decoy selection—What can we learn from earlier mistakes?* *Journal of Computer-Aided Molecular Design*, 2008. **22**(3): p. 213-228.
78. Pettersen, E.F., et al., *UCSF Chimera—a visualization system for exploratory research and analysis*. *J Comput Chem*, 2004. **25**(13): p. 1605-12.
79. Baker, M., *Fragment-based lead discovery grows up*. *Nat Rev Drug Discov*, 2013. **12**(1): p. 5-7.
80. Hubbard, R.E. and J.B. Murray, *Experiences in fragment-based lead discovery*. *Methods Enzymol*, 2011. **493**: p. 509-31.
81. Congreve, M., et al., *Recent Developments in Fragment-Based Drug Discovery*. *Journal of Medicinal Chemistry*, 2008. **51**(13): p. 3661-3680.
82. Hajduk, P.J. and J. Greer, *A decade of fragment-based drug design: strategic advances and lessons learned*. *Nature Reviews Drug Discovery*, 2007. **6**: p. 211.
83. Congreve, M., et al., *A 'Rule of Three' for fragment-based lead discovery?* *Drug Discovery Today*, 2003. **8**(19): p. 876-877.
84. Chessari, G. and A.J. Woodhead, *From fragment to clinical candidate—a historical perspective*. *Drug Discovery Today*, 2009. **14**(13): p. 668-675.
85. de Kloe, G.E., et al., *Transforming fragments into candidates: small becomes big in medicinal chemistry*. *Drug Discovery Today*, 2009. **14**(13): p. 630-646.
86. Hopkins, A.L., C.R. Groom, and A. Alex, *Ligand efficiency: A useful metric for lead selection*. *Drug Discovery Today*, 2004. **9**(10): p. 430-431.
87. Bembenek, S.D., B.A. Tounge, and C.H. Reynolds, *Ligand efficiency and fragment-based drug discovery*. *Drug Discovery Today*, 2009. **14**(5-6): p. 278-283.

88. Schultes, S., et al., *Ligand efficiency as a guide in fragment hit selection and optimization*. Drug Discovery Today: Technologies, 2010. **7**(3): p. 157-162.
89. Leeson, P.D. and B. Springthorpe, *The influence of drug-like concepts on decision-making in medicinal chemistry*. Nature Reviews Drug Discovery, 2007. **6**(11): p. 881-90.
90. Murray, C.W., M.L. Verdonk, and D.C. Rees, *Experiences in fragment-based drug discovery*. Trends in Pharmacological Sciences, 2012. **33**(5): p. 224-232.
91. Friesner, R.a., et al., *Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy*. Journal of Medicinal Chemistry, 2004. **47**(7): p. 1739-1749.
92. Friesner, R.a., et al., *Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes*. Journal of Medicinal Chemistry, 2006. **49**(21): p. 6177-6196.
93. Friesner Ra, B.J.L.M.R.B. and et al., *Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy*. J. Med. Chem., 2004. **47**(7): p. 1739-1749.
94. Bajorath, J., *Integration of virtual and high-throughput screening*. Nat Rev Drug Discov, 2002. **1**(11): p. 882-94.
95. Martin, Y.C., J.L. Kofron, and L.M. Traphagen, *Do Structurally Similar Molecules Have Similar Biological Activity?* Journal of Medicinal Chemistry, 2002. **45**(19): p. 4350-4358.
96. Bender, A. and R.C. Glen, *Molecular similarity: a key technique in molecular informatics*. Org Biomol Chem, 2004. **2**(22): p. 3204-18.
97. Bender, A., et al., *Molecular Similarity Searching Using Atom Environments, Information-Based Feature Selection, and a Naïve Bayesian Classifier*. Journal of Chemical Information and Computer Sciences, 2004. **44**(1): p. 170-178.
98. Bender, A., et al., *Similarity Searching of Chemical Databases Using Atom Environment Descriptors (MOLPRINT 2D): Evaluation of Performance*. Journal of Chemical Information and Computer Sciences, 2004. **44**(5): p. 1708-1718.
99. Willett, P., *Similarity-based virtual screening using 2D fingerprints*. Drug Discov Today, 2006. **11**(23-24): p. 1046-53.
100. Cheng, T., et al., *Structure-based virtual screening for drug discovery: a problem-centric review*. Aaps j, 2012. **14**(1): p. 133-41.
101. Cole, J., et al., *The Basis for Target-Based Virtual Screening: Protein Structures*. 2011. 87-114.
102. Verdonk, M.L., et al., *Virtual screening using protein-ligand docking: avoiding artificial enrichment*. J Chem Inf Comput Sci, 2004. **44**(3): p. 793-806.
103. Chen, Y.C., *Beware of docking!* Trends Pharmacol Sci, 2015. **36**(2): p. 78-95.
104. Burley, S.K., et al., *RCSB Protein Data Bank: Sustaining a living digital data resource that enables breakthroughs in scientific research and biomedical education*. Protein Sci, 2018. **27**(1): p. 316-330.
105. Jones, G., et al., *Development and validation of a genetic algorithm for flexible docking I* Edited by F. E. Cohen. Journal of Molecular Biology, 1997. **267**(3): p. 727-748.
106. Krieger, E., S.B. Nabuurs, and G. Vriend, *Homology modeling*. Methods Biochem Anal, 2003. **44**: p. 509-23.
107. Ramachandran, S. and N.V. Dokholyan, *Homology Modeling: Generating Structural Models to Understand Protein Function and Mechanism*, in *Computational Modeling of Biological Systems: From Molecules to Pathways*, N.V. Dokholyan, Editor. 2012, Springer US: Boston, MA. p. 97-116.

108. *Activities at the Universal Protein Resource (UniProt)*. Nucleic Acids Res, 2014. **42**(Database issue): p. D191-8.
109. Altschul, S.F., et al., *Basic local alignment search tool*. J Mol Biol, 1990. **215**(3): p. 403-10.
110. Biasini, M., et al., *SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information*. Nucleic Acids Research, 2014. **42**(W1): p. W252-W258.
111. Sali, A. and T.L. Blundell, *Comparative protein modelling by satisfaction of spatial restraints*. J Mol Biol, 1993. **234**(3): p. 779-815.
112. Eswar, N., et al., *Comparative protein structure modeling using Modeller*. Curr Protoc Bioinformatics, 2006. **Chapter 5**: p. Unit-5.6.
113. Shen, M.Y. and A. Sali, *Statistical potential for assessment and prediction of protein structures*. Protein Sci, 2006. **15**(11): p. 2507-24.
114. Laskowski, R., et al., *PROCHECK: A program to check the stereochemical quality of protein structures*. Vol. 26. 1993. 283-291.
115. Greenwood, J.R., et al., *Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution*. Journal of Computer-Aided Molecular Design, 2010. **24**(6): p. 591-604.
116. Shelley, J.C., et al., *Epik: a software program for pK a prediction and protonation state generation for drug-like molecules*. Journal of Computer-Aided Molecular Design, 2007. **21**(12): p. 681-691.
117. Banks, J.L., et al., *Integrated Modeling Program, Applied Chemical Theory (IMPACT)*. J Comput Chem, 2005. **26**(16): p. 1752-80.
118. Hawkins, P.C., et al., *Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database*. J Chem Inf Model, 2010. **50**(4): p. 572-84.
119. Case, D.A., et al., *The Amber biomolecular simulation programs*. J Comput Chem, 2005. **26**(16): p. 1668-88.
120. Gerber, P.R. and K. Muller, *MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry*. J Comput Aided Mol Des, 1995. **9**(3): p. 251-68.
121. D.A. Case, T.A.D., T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R., et al., *AMBER 12*. University of California, San Francisco, 2012.
122. Duan, Y., et al., *A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations*. Journal of Computational Chemistry, 2003. **24**(16): p. 1999-2012.
123. Lee, M.C. and Y. Duan, *Distinguish Protein Decoys by Using a Scoring Function Based on a New AMBER Force Field, Short Molecular Dynamics Simulations, and the Generalized Born Solvent Model*. Proteins: Structure, Function and Genetics, 2004. **55**(3): p. 620-634.
124. Case, D.A., et al., *Amber 12 Reference Manual*. University of California, San Francisco, 2012: p. 350-350.
125. Jakalian, A., D.B. Jack, and C.I. Bayly, *Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation*. Journal of Computational Chemistry, 2002. **23**(16): p. 1623-1641.
126. Wang, J.M., et al., *Development and testing of a general amber force field*. J. Comput. Chem., 2004. **25**(9): p. 1157-1174.
127. Jorgensen, W.L., et al., *Comparison of simple potential functions for simulating liquid water*. The Journal of Chemical Physics, 1983. **79**(2): p. 926-926.
128. Pastor, R.W., B.R. Brooks, and A. Szabo, *An analysis of the accuracy of Langevin and molecular dynamics algorithms*. Molecular Physics, 1988. **65**(6): p. 1409-1419.

129. Darden, T., D. York, and L. Pedersen, *Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems*. The Journal of Chemical Physics, 1993. **98**(12): p. 10089-10089.
130. Ryckaert, J.P., G. Ciccotti, and H.J.C. Berendsen, *Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes*. Journal of Computational Physics, 1977. **23**(3): p. 327-341.
131. Still, W.C., et al., *Semianalytical treatment of solvation for molecular mechanics and dynamics*. Journal of the American Chemical Society, 1990. **112**(16): p. 6127-6129.
132. Srinivasan, J., et al., *Application of a pairwise generalized Born model to proteins and nucleic acids: inclusion of salt effects*. Theoretical Chemistry Accounts, 1999. **101**(6): p. 426-434.
133. Luo, R., L. David, and M.K. Gilson, *Accelerated Poisson-Boltzmann calculations for static and dynamic systems*. Journal of Computational Chemistry, 2002. **23**(13): p. 1244-1253.
134. Gleeson, M.P. and D. Gleeson, *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**(6): p. 1437-1448.
135. Hayik, S.A., R. Dunbrack, and K.M. Merz, *A Mixed QM/MM Scoring Function to Predict Protein-Ligand Binding Affinity*. Journal of chemical theory and computation, 2010. **6**(10): p. 3079-3091.
136. Miller, B.R., et al., *MMPBSA.py: An efficient program for end-state free energy calculations*. Journal of Chemical Theory and Computation, 2012. **8**(9): p. 3314-3321.
137. Ferrari, A.M., et al., *Validation of an automated procedure for the prediction of relative free energies of binding on a set of aldose reductase inhibitors*. Bioorganic and Medicinal Chemistry, 2007. **15**(24): p. 7865-7877.
138. Hou, T., et al., *Assessing the Performance of the MM / PBSA and MM / GBSA Methods . I . The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations*. J. Chem. Inf. Model, 2010. **51**(1): p. 69-82.
139. Wichapong, K., et al., *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**(9): p. 1574-88.
140. Uciechowska, U., et al., *Binding free energy calculations and biological testing of novel thiobarbiturates as inhibitors of the human NAD⁺ dependent histone deacetylase Sirt2*. MedChemComm, 2012. **3**(2): p. 167-167.
141. Genheden, S. and U. Ryde, *The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities*. Expert Opinion on Drug Discovery, 2015. **0441**(October): p. 1-13.
142. Slynko, I., et al., *Identification of Highly Potent Protein KinaseC-Related Kinase1 Inhibitors by Virtual Screening, Binding Free Energy Rescoring, and invitro Testing*. Chemmedchem, 2016. **11**(18): p. 2084-2094.
143. Karaman, B. and W. Sippl, *Docking and binding free energy calculations of sirtuin inhibitors*. European Journal of Medicinal Chemistry, 2015. **93**: p. 584-598.
144. Kollman, P.a., et al., *Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models*. Accounts of Chemical Research, 2000. **33**(12): p. 889-897.
145. Gilson, M.K. and H.X. Zhou, *Calculation of protein-ligand binding affinities*. Annual Review of Biophysics and Biomolecular Structure, 2007. **36**: p. 21-42.
146. Connolly, M.L., *Analytical molecular surface calculation*. Journal of Applied Crystallography, 1983. **16**(5): p. 548-558.
147. Hawkins, G.D., C.J. Cramer, and D.G. Truhlar, *Pairwise solute descreening of solute charges from a dielectric medium*. Chemical Physics Letters, 1995. **246**(1-2): p. 122-129.

148. Hawkins, G.D., C.J. Cramer, and D.G. Truhlar, *Parametrized Models of Aqueous Free Energies of Solvation Based on Pairwise Descreening of Solute Atomic Charges from a Dielectric Medium*. The Journal of Physical Chemistry, 1996. **100**(51): p. 19824-19839.
149. Tsui, V. and D.A. Case, *Theory and applications of the Generalized Born solvation model in macromolecular simulations*. Biopolymers, 2000. **56**(4): p. 275-291.
150. Onufriev, A., D. Bashford, and D.A. Case, *Exploring Protein Native States and Large-Scale Conformational Changes with a Modified Generalized Born Model*. Proteins: Structure, Function and Genetics, 2004. **55**(2): p. 383-394.
151. Mongan, J., et al., *Generalized Born model with simple, robust molecular volume correction*. NIH Public Access, 2010. **3**(1): p. 156-169.
152. Homeyer, N. and H. Gohlke, *Free Energy Calculations by the Molecular Mechanics Poisson–Boltzmann Surface Area Method*. Molecular Informatics, 2012. **31**(2): p. 114-122.
153. Pearlman, D.A., et al., *AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules*. Computer Physics Communications, 1995. **91**(1): p. 1-41.
154. Giordanetto, F., et al., *Novel scoring functions comprising QXP, SASA, and protein side-chain entropy terms*. Journal of Chemical Information and Computer Sciences, 2004. **44**(3): p. 882-893.
155. Raha, K. and K.M. Merz, *A Quantum Mechanics-Based Scoring Function: Study of Zinc Ion-Mediated Ligand Binding*. Journal of the American Chemical Society, 2004. **126**(4): p. 1020-1021.
156. Thompson, D.C., C. Humblet, and D. Joseph-McCarthy, *Investigation of MM-PBSA rescoring of docking poses*. Journal of Chemical Information and Modeling, 2008. **48**(5): p. 1081-1091.
157. Wichapong, K., et al., *Application of docking and QM/MM-GBSA rescoring to screen for novel Myt1 kinase inhibitors*. Journal of Chemical Information and Modeling, 2014. **54**(3): p. 881-893.
158. MR., B., *Status of the NCI preclinical antitumor drug discovery screen*. Princ Pract Oncol, 1989. **3:1**.
159. Monks A, S.D., Skehan P, Boyed M., *Implementation of a pilot-scale, high flux anticancer drug screen utilizing disease-oriented panels of human tumor cell lines in culture*. Proc Am Assoc Cancer Res, 1989. **30:607**.
160. Cohen, P., *Protein kinases--the major drug targets of the twenty-first century?* Nat Rev Drug Discov, 2002. **1**(4): p. 309-15.

APPENDIX

5.1A. FDA-approved protein kinase inhibitors

FDA-approved protein kinase inhibitors, its indications and classifications compiled by Robert Roskoski Jr. Updated in 20th of April 2018.

<i>Drug¹</i>	<i>Company</i>	<i>Inhibitor type²</i>	<i>Year approved</i>	<i>Known targets</i>	<i>Class³</i>	<i>Disease</i>	<i>PDB</i>
<i>Abemaciclib, LY2835219, Verzenio</i>	Lilly	I1/2B	2017	CDK4/6	S/T	Breast Ca	5L2S
<i>Acalabrutinib ACP-196</i>	Acerta Pharma	?	2017	Bruton tyrosine kinase	NRY	Mantle cell lymphoma	
<i>Afatinib, BIBW 2992, Tovok, OWN</i>	Boehringer Ingelheim	VI	2013	EGFR, ErbB2, ErbB4	RY	NSCLC	4G5J
<i>Alectinib, CH5424802, Alecensa, EMH</i>	Hoffman-LaRoche	I1/2B, IIB	2015	ALK and RET	RY	NSCLC, ALK-positive	3AOX
<i>Axitinib, AG-013736, Inlyta, AXI</i>	Pfizer	IIA	2012	VEGFR1/2/3, PDGFR β	RY	RCC, advanced	4AG8
<i>Bosutinib, SKI-606, BOSULIF, DB8</i>	Pfizer	I, IIB	2012	BCR-Abl, Src, Lyn, and Hck	NRY	CML	3UE4
<i>Brigatinib, AP 26113, Alunbrig</i>	Ariad	I1/2B, IIB	2017	ALK, ROS1, IGF-1R, Flt3, EGFR	RY	NSCLC, ALK-positive	5J7H
<i>Cabozantinib, XL-184, BMS-907351, Cometriq</i>	Excelixis	IIA	2012	RET, MET, VEGFR1/2/3, Kit, TrkB, Flt3, Axl, Tie2	RY	Thyroid cancer, metastatic medullary	
<i>Ceritinib, LDK378, Zykadia, 4MK</i>	Novartis	I	2014	ALK, IGF-1R, InsR, ROS1	RY	NSCLC, ALK-positive after crizotinib resistance	4MKC
<i>Cobimetinib, GDC-0973, EUI, Cotellic</i>	Genentech	III	2015	MEK1/2	T/Y	Melanoma with BRAF mutations together with vemurafenib	4AN2
<i>Crizotinib, PF 2341066, VGH, Xalkori</i>	Pfizer	I, I1/2B	2011	ALK, MET (HGFR), ROS1, MST1R	RY	NSCLC, ALK-positive or ROS1-positive	2XB2 3ZBF

							2WGJ
<i>Dabrafenib</i> , GSK2118436, <i>Tafinlar</i> , P06	GSK	II	2013	B-Raf	S/T	Melanoma and NSCLC with <i>BRAF</i> mutations	5CSW
<i>Dasatinib</i> , BMS-354825, <i>Sprycel</i> , 1N1	Bristol-Myers Squibb	I, I1/2A	2006	BCR-Abl, Src, Lck, Yes, Fyn, Kit, EphA2, PDGFR β	NRY	CML	2GQG
<i>Erlotinib</i> , CP-358774, <i>OSI-774</i> , <i>Tarceva</i> , AQ4	Genentech	I, I1/2B	2004	EGFR	RY	NSCLC; pancreatic cancer	4HJO 1M17
<i>Everolimus</i> , RAD001, <i>Afinitor</i>	Novartis	IV	2009	FKBP12/mTOR	S/T	Breast cancer, HER2-negative; PNET; RCC; renal angiomyolipoma; subependymal giant cell astrocytoma	
<i>Fostamatinib</i> , R788, <i>Tavalisse</i>	Rigel	?	2018	Syk, Spleen tyrosine kinase	RY	Thrombo-cytopenia	
<i>Gefitinib</i> , ZD1839, <i>Iressa</i> , IRE	Astra Zeneca	I	2003-2005, 2015	EGFR	RY	NSCLC	4WKQ 2ITY
<i>Ibrutinib</i> , PCI-32765, <i>Imbruvica</i>	Pharmaceuticals and J&J	VI	2013	Bruton tyrosine kinase	NRY	Mantle cell lymphoma; CLL; Waldenstrom's macroglobulinemia; marginal zone lymphoma; graft vs. host disease	5P9J
<i>Imatinib</i> , STI571, <i>Gleevec</i> , STI	Novartis	IIA	2001	BCR-Abl, Kit, and PDGFR	NRY	CML and ALL, Ph ⁺ aggressive systemic mastocytosis; CEL; DFSP; HES; GIST; MDS/MDP	2HYY 1IEP
<i>Lapatinib</i> , GW572016, <i>Tykerb</i> , FMM	GSK	I, I1/2A	2007	EGFR, ErbB2	RY	Breast cancer	1XKK

Lenvatinib , AK175809, <i>Lenvima</i> , <i>LEV</i>	Easai	I1/2A	2015	VEGFR1/2/3, PDGFR, FGFR, Kit, RET	RY	Different-iated thyroid cancer	3WZD
Midostaurin , PKC412, <i>CPG</i> 41251, <i>Rydapt</i>	Novartis	?	2017	Flt3, PDGFR, VEGFR2, PKC	RY	Acute myeloid leukemia, mastocytosis, mast cell leukemia	
Neratinib , HKI-272	Puma	?	2017	ErbB2/HER2	Ry	HER ⁺ breast cancer	2JIV
Nilotinib , AMN107, <i>Tasigna</i> , <i>NILK</i>	Novartis	IIA	2007	BCR-Abl, PDGFR, DDR1	NRY	CML, Ph ⁺	3CS9
Nintedanib , BIBF-1120, <i>Vargatef</i> , <i>XIN</i>	Boehringer Ingelheim	IIB	2014	FGFR1/2/3, PDGFR α/β , VEGFR1/2/3, Flt3	RY	Pulmonary fibrosis, idiopathic	3C7Q
Osimertinib , AZD-9292, <i>Tagrisso</i>	AstraZeneca	?	2015	EGFR T970M	RY	NSCLC	
Palbociclib , PD-0332991, <i>Ibrance</i> , <i>LQQ</i>	Park Davis	I	2015	CDK4/6	S/T	Breast cancer, ER ⁺ and HER2 ⁺	5L2I
Pazopanib , GW786034, <i>Votrient</i>	GSK	?	2009	VEGFR1/2/3, PDGFR α/β , FGFR1/3, Kit, Lck, Fms, Itk	RY	RCC; soft tissue sarcomas	
Ponatinib , AP 24534, <i>Iclusig</i> , <i>OLI</i>	Ariad	I	2012	BCR-Abl, BCR-Abl T315I, VEGFR, PDGFR, FGFR, EphR, Src family kinases, Kit, RET, Tie2, Flt3	NRY	CML or ALL, Ph ⁺	4UOI 3OXZ 1UWH
Regorafenib , BAY 73-4506, <i>Stivarga</i>	Bayer	IIA	2012	VEGFR1/2/3, BCR-Abl, B-Raf, B-Raf (V600E), Kit, PDGFR α/β , RET, FGFR1/2, Tie2, and Eph2A	RY	CRC	
Ribociclib , LEE011, <i>Kisqali</i>	Novartis	?	2017	CDK4/6	S/T	Breast cancer	5L2T
Ruxolitinib , INCB-018424, <i>Jakafi</i> , <i>RXT</i>	Incyte	I	2011	JAK1/2	NRY	Myelofibrosis; polycythemia vera	4U5J
Sirolimus , <i>Rapamycin</i>	Wyeth	IV	1999	FKBP/mTOR	S/T	Renal transplant; lymphangio- leiomyomatosis	

Sorafenib , BAY 43-9006, <i>Nexavar</i> , BAX,	Onyx	IIA	2005	VEGFR1/2/3, B-/C-Raf, mutant B-Raf, Kit, Flt3, RET, and PDGFR β	RY	Thyroid cancer, differentiated; hepatocellular carcinoma; RCC	4ASD
Sunitinib , SU-11248, <i>Sutent</i> , B49	Pfizer	I1/2B, IIB	2006	PDGFR α/β , VEGFR1/2/3, Kit, Flt3, CSF-1R, and RET	RY	RCC; GIST; pancreatic neuroendocrine tumors	4AGD
Temsirolimus , CCI-779, <i>Torisel</i>	Wyeth	IV	2007	FKBP12/mTOR	S/T	RCC, advanced	
Tofacitinib , CP-690550, <i>Tasocitinib</i> , MII	Pfizer	I	2012	JAK3	NRY	Rheumatoid arthritis	3EYG 3LXK
Trametinib , <i>Mekinist</i>	GSK	III	2013	MEK1/2	T/Y	Melanoma	
Vandetanib , ZD6474, <i>Zactima</i> , ZD6	Astra-Zeneca	I	2011	EGFRs, VEGFRs, RET, Brk, Tie2, EphRs, and Src family kinases	RY	Thyroid cancer, medullary	2IVU
Vemurafenib , PLX-4032, <i>Zelboraf</i> , O32	Genentech	I1/2A	2011	A/B/C-Raf and B-Raf (V600E)	S/T	Melanoma with the <i>BRAF</i> ^{V600E} mutation	4RZV 3OG7

ALL, acute lymphoblastic leukemia; CEL, chronic eosinophilic leukemia; CLL, chronic lymphocytic leukemia; CML, chronic myelogenous leukemia; CRC, colorectal cancer; DDR1, Discoidin domain receptor family, member 1; DFSP, dermatofibrosarcoma protuberans; GIST, gastrointestinal stromal tumor; HES, hypereosinophilic syndrome; HGFR, hepatocyte growth factor receptor; MDS/MPD, myelodysplastic/myeloproliferative diseases; MST1R, macrophage-stimulating protein receptor aka RON (Recepteur d'Origine Nantais); NSCLC, non-small cell lung cancer; PNET, progressive neuroendocrine tumors of pancreatic origin; Ph⁺, Philadelphia chromosome positive; RCC, renal cell carcinoma.

¹All drugs listed are orally effective except for temsirolimus, which is given intravenously. Generic name, code, trade name, PubChem abbreviation, and X-ray PDB ID

²Types I, II, III, IV, V, and VI from Roskoski, *Pharmacological Research* **103**, 26-48 2016

³NRY, non-receptor protein-tyrosine kinase; RY, receptor protein-tyrosine kinase; S/T, protein-serine/threonine protein kinase; T/Y, dual specificity protein kinase

5.2A. List of transferases: Phosphorous-containing groups (kinase) (EC 2.7.-.-)

Retrieved from Wikipedia, blue highlighted section refers to the targeted kinase families by the current work.

2.7.1-2.7.4: <i>phosphotransferase</i> <i>/ kinase (PO4)</i>	<i>2.7.1: OH acceptor</i>	Hexo-, Gluco-, Fructo-, Hepatic, Galacto-, Phosphofructo-(1, Liver, Muscle, Platelet, 2) Riboflavin, Shikimate, Thymidine, ADP-thymidine, NAD+, Glycerol, Pantothenate, Mevalonate, Pyruvate, Deoxycytidine, PFP, Diacylglycerol, Phosphoinositide 3, (Class I PI 3, Class II PI 3), Sphingosine, Glucose-1,6-bisphosphate synthase	
	<i>2.7.2: COOH acceptor</i>	Phosphoglycerate, Aspartate kinase	
	<i>2.7.3: N acceptor</i>	Creatine	
	<i>2.7.4: PO4 acceptor</i>	Phosphomevalonate, Adenylate, Nucleoside-diphosphate, Uridylate, Guanylate, Thiamine-diphosphate	
2.7.6: <i>diphosphotransferase</i>	Ribose-phosphate diphosphokinase, Thiamine diphosphokinase		
2.7.7: <i>nucleotidyltransferase</i> <i>(PO4-nucleoside)</i>	<i>Polymerase</i>	DNA	DNA-directed DNA polymerase: I, II, III, IV, V
		polymerase	RNA-directed DNA polymerase: Reverse transcriptase (Telomerase) DNA nucleotidyltransferase/Terminal deoxynucleotidyl transferase
	nucleotidyltransferase	RNA	RNA polymerase/DNA-directed RNA polymerase/ RNA polymerase: I, II, III, IV, V, Primase, RNA-dependent RNA polymerase
<i>Phosphorolytic</i>		PNPase	
		RNase PH, PNPase	

	<i>3' to 5' exoribonuclease</i>	
	<i>Nucleotidyltransferase</i>	UTP—glucose-1-phosphate uridylyltransferase, Galactose-1-phosphate uridylyltransferase
	<i>Guanylyltransferase</i>	mRNA capping enzyme
	<i>Other</i>	Recombinase (Integrase), Transposase
2.7.8: miscellaneous	<i>Phosphatidyltransferases</i>	CDP-diacylglycerol—glycerol-3-phosphate 3-phosphatidyltransferase, CDP-diacylglycerol—serine O-phosphatidyltransferase, CDP-diacylglycerol—inositol 3-phosphatidyltransferase, CDP-diacylglycerol—choline O-phosphatidyltransferase
	<i>Glycosyl-1-phosphotransferase</i>	N-acetylglucosamine-1-phosphate transferase
2.7.10-2.7.13: protein kinase	<i>2.7.10: protein-tyrosine</i>	tyrosine kinases
	<i>2.7.11: protein-serine/threonine</i>	serine/threonine-specific protein kinases
	<i>2.7.12: protein-dual-specificity</i>	serine/threonine-specific protein kinases
	<i>2.7.13: protein-histidine</i>	Protein-histidine pros-kinase, Protein-histidine tele-kinase, Histidine kinase

5.3A. Full-text publications

DECLARATION OF AUTHORSHIP

Selbstständigkeitserklärung

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 wortlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Die Arbeit wurde ausschlieslich der Mathematisch-Naturwissenschaftlichen Fakultät der Martin-Luther-Universität Halle-Wittenberg vorgelegt und an keiner anderen Universität oder Hochschule weder im In- und Ausland zur Erlangung des Doktorgrades eingereicht.

Halle (Saale), den 04.09.2018

Abdulkarim Najjar

Curriculum Vitae



PERSONAL INFORMATION

Name Abdulkarim Najjar
Address An der Schwemme 03
06108 Halle(Saale), DE
Phone +49 176 37236129
Email abdulkarim.najjar@pharmazie.uni-halle.de
Birthplace/ date Bennesh, 01.01.1987
Citizenship Syrian
Languages English, German, Arabic

EDUCATION

- 10.2014 – 10.2018 **The Martin-Luther University of Halle-Wittenberg.** Halle (Saale), DE
Medicinal Chemistry Department, Institute of Pharmacy.
- PhD Candidate
- 04.2014 - 10.2014 **The Martin-Luther University of Halle-Wittenberg.** Halle (Saale), DE
Medicinal Chemistry Department, Institute of Pharmacy.
- Diploma Degree.
- 10.2010 - 11.2013 **University of Aleppo** Aleppo, SY
Pharmaceutical Chemistry Department, Faculty of Pharmacy.
- Master in Applied Pharmaceutical Chemistry.
- 09.2005 - 08.2010 **University of Aleppo** Aleppo, SY
Faculty of Pharmacy.
- Doctor of Pharmacy (Pharm.D.)

INTERNSHIPS

- 05.2018 **University of Hamburg** Hamburg, DE
Applied Cheminformatics and Molecular Design, Center for Bioinformatics.
- 11.2016 - 07. 2017 **University of Konstanz** Konstanz, DE
Department of Politics and Public Administration
- E-Learning courses and Face-to-Face workshops
- 07.2009 **University of Assiut** Assiut, EG
Medicinal Chemistry Department, Faculty of Pharmacy.

WORKING EXPERIENCE

RESEARCH

01-26.05.2018	Internship at Center for Bioinformatics, University of Hamburg.	Ham. DE
04.2014 - Present	PhD Candidate at Institute of Pharmacy, MLU.	Halle (Saale), DE
10.2010- 11.2013	Master Student at Faculty of Pharmacy, Aleppo University.	Aleppo, SY
01-31.07.2009	Internship at Faculty of Pharmacy, Assiut University.	Assiut, EG

TEACHING

01.2017- 05.2017	Lab-assistant at Institute of Pharmacy, MLU.	Halle (Saale), DE
02.2013- 07.2013	Lecturer at Faculty of Pharmacy, Private Ebla University.	Edlib, SY
10.2010- 07.2013	Teacher-assistant at Faculty of Pharmacy, Aleppo University.	Aleppo, SY

PUBLICATIONS

- Najjar, A., Platzer, C., Luft, A., Assmann, C., ElGhazawy, N., Erdmannb, F., Sippl, W., & Schmidt, M., Computer-aided design, synthesis and biological characterization of novel inhibitors for PKMYT1. *Manuscript Submitted, European Journal of Medicinal Chemistry*.
- Platzer, C., Najjar, A., Rohe, A., Erdmannb, F., Sippl, W., & Schmidt, M. Identification of protein kinase PKMYT1 inhibitors by screening the GSK published protein kinase inhibitor set I and II. *Bioorganic Med. Chem.* 2018.
- Fischer, T., Najjar, A., Totzke, F., Schachtele, C., Sippl, W., Ritter, C., & Hilgeroth, A. Discovery of novel dual inhibitors of receptor tyrosine kinases EGFR and PDGFR- β related to anticancer drug resistance. *J. Enzyme Inhib. Med. Chem.* 2018.
- Schmidt, M., Rohe, A., Platzer, C., Najjar, A., Erdmann, F., & Sippl, W. Regulation of G2/M transition by inhibition of Wee1 and PKPKMYT1 kinases. *Molecules*, 2017.
- Fischer, T., Kruger, T., Najjar, A., Totzke, F., Schachtele, C., Sippl, W., Ritter, C., & Hilgeroth, A. Discovery of novel substituted benzo-anellated 4-benzylamino pyrrolopyrimidines as dual EGFR and VEGFR2 inhibitors. *Bioorg. Med. Chem. Lett.* 2017.
- Najjar, A., Ntie-Kang, F., & Sippl, W. Application of Computer Modeling to Drug Discovery: Case study of PRK1 Kinase Inhibitors as Potential Drugs in Prostate Cancer Treatment. Unique Aspects of Anti-cancer Drug Development. *InTech Open*, 2017.
- Hempel, C., Totzke, F., Schachtele, C., Najjar, A., Sippl, W., Ritter, C., & Hilgeroth, A. Discovery of novel dual inhibitors of receptor tyrosine kinases EGFR and IGF-1R. *J. Enzyme Inhib. Med. Chem.* 2017.
- Hempel, C., Najjar, A., Totzke, F., Schachtele, C., Sippl, W., Ritter, C., & Hilgeroth, A. Discovery of dually acting small-molecule inhibitors of cancer-resistance relevant receptor tyrosine kinases EGFR and IGF-1R. *MedChemComm*, 2016.

- Slynko, I., Schmidtkunz, K., Rumpf, T., Klaeger, S., Heinzlmeir, S., Najjar, A., Metziger, E., Kuster, B., Schele, R., Jung, M., & Sippl, W. Identification of Highly Potent Protein Kinase C-Related Kinase 1 Inhibitors by Virtual Screening, Binding Free Energy Rescoring, and in Vitro Testing. *ChemMedChem*, 2016.

CONFERENCE PRESENTATIONS

TALKS

- Najjar, A. et al. Computer-aided Drug Design for Cancer Treatments. *14th Proteome Forum*, University Medicine **Rostock, DE**, 2017.
- Najjar, A. et al. Computational Approaches in Drug Discovery for Cancer Treatment. *6th Workshop of Theoretical Chemistry and Computational Modelling (TCCM)*, **Barcelona, ES**, 2017.

POSTERS

- Najjar, A. et al. Applications of Binding Free Energy Calculations and QSAR Modeling to Design Novel Inhibitors for Human Kinase. *11th International Conference on Chemical Structures (ICCS-2018)*, **Noordwijker., NL**, 2018.
- Najjar, A. et al. Chemoinformatics-assisted Identification and Development of Novel PKMYT1 Kinase Inhibitors. *Frontiers in Medicinal Chemistry*, **Jena, DE**, 2018.
- Najjar, A. et al. Structure- & Fragment-Based Design of Novel Diaminopyrimidine Derivatives as Inhibitors for Human Myt1 Kinase. *13th German Conference on Chemoinformatics*, **Mainz, DE**, 2017.
- Najjar, A. et al. Structure- & Fragment-Based Design of Novel Diaminopyrimidine Derivatives as Inhibitors for Human Myt1 Kinase. *9th Vienna Summer School on Drug Design*, **Vienna, AT**, 2017.
- Najjar, A. et al. Structure/Fragment-Based Design of Novel Inhibitors for Human Myt1 Kinase. *11th European Conference on Theoretical and Computational Chemistry (EUCCO-TCC)*, **Barcelona, ES**, 2017.
- Najjar, A. et al. Structure-Based and Fragment-Based Design of Novel Diaminopyrimidine Derivatives as Inhibitors for Human Myt1 Kinase. *37th European School of Medicinal Chemistry*, **Urbino, IT**, 2017.
- Najjar, A. et al. Fragment-based Design of Novel Inhibitors for Human Myt1 Kinase. *6th RSC-BMCS Fragment-based Drug Discovery meeting*, **Vienna, AT**, 2017.
- Najjar, A. et al. Structure-based Development of Novel Inhibitors for the Epigenetic Target EZH2. *Central German Meeting on Bioinformatics*, **Halle, DE**, 2015.