

Evaluation und Anwendung von verschiedenen Virtuellen Screening-Methoden am Beispiel von Proteinkinasen

Dissertation

zur Erlangung des
Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät I – Biowissenschaften –

der Martin-Luther-Universität
Halle-Wittenberg,

vorgelegt

von Herrn Erenkamp, Daniel German

geb. am 09.01.1979 in Mönchengladbach

Angefertigt im	Institut für Pharmazie Martin-Luther-Universität Halle-Wittenberg
Tag der Verteidigung	19. Dezember 2014
Erster Gutachter	Univ.-Prof. Dr. W. Sippl
Zweiter Gutachter	PD Dr. A. Hilgeroth
Dritter Gutachter	Univ.-Prof. Dr. G. Ecker

INHALTSVERZEICHNIS

Abbildungsverzeichnis	v
Tabellenverzeichnis	vii
1. Einleitung	1
1.1. Virtuelles Screening	1
1.2. Kinasen	1
1.2.1. Systematik	2
1.2.2. Struktur und Funktion	4
1.2.3. Kinaseinhibitoren	6
1.3. Kinase-Familien mit Bezug zur Arbeit	11
1.3.1. VEGFR, EGFR, IGFR	11
1.3.2. MYT1	13
1.3.3. PRK1	14
1.4. Zielsetzung	14
2. Material und Methoden	15
2.1. Virtuelles Screening	15
2.2. Strukturbasierte VS Methoden	15
2.2.1. Suchalgorithmen	15
2.2.2. Scoring-Funktion	16
2.2.3. Consensus-Score	20
2.3. Datenbank Filterung	21
2.4. Ähnlichkeitsbasierte Suche	22
2.5. Moleküldynamik-Simulation	23
2.5.1. Molecular Mechanics Poisson-Boltzmann/surface area (MM-PB/SA)	24
2.5.2. Molecular Mechanics Generalized-Born/surface area (MM-GB/SA)	25
2.5.3. Linear Response Molecular Mechanics Poisson-Boltzmann/Surface Area	26

2.6.	Gütemaße	26
2.6.1.	Receiver Operating Characteristic Kurve	26
2.6.2.	Area Under Curve	27
2.6.3.	Robust Initial Enhancement	28
2.6.4.	Enrichment Factor	28
2.7.	Erstellen der Substanzdatenbanken	28
2.8.	Vorbereiten der Proteinstrukturen für das Docking	29
2.8.1.	Virtuelles Screening	29
2.8.2.	Moleküldynamik-Simulation	29
2.9.	Verwendete Programme	29
2.10.	Kinase Assays	30
2.10.1.	PRK1 Assay	30
2.10.2.	MYT1-Assay: LanthaScreen Assay mittels FRET	30
2.10.3.	IGFR, EGFR und VEGFR Assay	31
3.	Ergebnisse und Diskussion	33
3.1.	Dockingstudien an Proteinkinasen	33
3.1.1.	Betrachtung der einzelnen Programme und Scoring-Funktion	33
3.1.2.	Generierung der Datenbanken	33
3.1.3.	Prüfen der Dockingeinstellungen	35
3.1.4.	Fingerprintanalyse von aktiven und inaktiven Liganden	35
3.1.5.	Auswertung der AUC-Werte	36
3.1.6.	Auswertung der EF	49
3.1.7.	Auswertung RIE	59
3.1.8.	Diskussion	61
3.2.	PRK1	65
3.2.1.	Dockingstudien	65
3.2.2.	Diskussion	70
3.3.	MYT1-Kinase	72
3.3.1.	Docking	72
3.3.2.	Moleküldynamik-Simulation an der MYT1	74
3.3.3.	Diskussion	77

3.4. VEGFR, EGFR, IGFR	80
3.4.1. Docking	81
3.4.2. MM-PB/SA	84
3.4.3. LR-MMPBSA	91
3.4.4. Diskussion	91
4. Zusammenfassung	93
5. Literatur	95
Anhang	i
5.1. Screening	ii
5.1.1. AUC	ii
5.1.2. Enrichmentfaktor 5	vii
5.1.3. RIE	xiii
5.1.4. RIE für die jeweilige Methode	xviii
5.2. Moleküldynamik-Simulation	xxiv
5.2.1. Vorbereitung der Simulation	xxiv
5.3. IGFR, EGFR, VEGFR	xxvi
5.3.1. Durchführung des Affinitäts Assay	xxvi
5.4. PRK1	xxvii

ABBILDUNGSVERZEICHNIS

1.1	Dendrogramm humanes Kinom	2
1.2	Struktur der Kinasedomäne	4
1.3	Hydrophobe Taschen I und II	5
1.4	DFG-Motiv	5
1.5	Phosphorylierung	6
1.6	Hinge-Region mit ATP	7
1.7	Kinaseinhibitoren	8
1.8	Typ-I- und Typ-II-Inhibitoren	9
1.9	Bisindolylmaleimide	10
1.10	Bindemodus von Staurosporin	10
1.11	Allosterische Hemmung der MEK1	11
1.12	Zellzyklus	13
2.1	Atomare Wechselwirkungen	17
2.2	Freie Bindungsenergie	24
2.3	ROC-Kurve und AUC	27
2.4	TR-FRET	31
3.1	Performance des Dockings	36
3.2	Aktive und inaktive Inhibitoren PIM	40
3.3	Aktive und inaktive Inhibitoren CDK4	41
3.4	Aktive und inaktive Inhibitoren CDK7	43
3.5	Aktive und inaktive Inhibitoren CDK2	44
3.6	Aktive und inaktive Inhibitoren CDK7	46
3.7	Aktive und inaktive Inhibitoren C-JUN	47
3.8	Aktive und inaktive Inhibitoren MET	49
3.9	Aktive und inaktive Inhibitoren CDC7	50
3.10	Aktive Liganden der PRK1	65

3.11 Dockinglösungen in PRK1	68
3.12 Dockinglösungen in PRK1	69
3.13 Design von Inhibitoren der MYT1	72
3.14 Dockinglösungen in MYT1	74
3.15 RMSD-Plots der ausgewählten Komplexe	75
3.16 Lage von Dasatinib	78
3.17 Lage von Tyrphostin AG1478	79
3.18 Getestete Liganden	80
3.19 EGFR Vergleich Dockinglösungen	82
3.20 Ausgewählte Dockingposen	83
3.21 Vergleich der Lage nach MD von Substanz 27 in EGFR	85
3.22 Vergleich der Lage nach MD von Substanz 19 in VEGFR	85
3.23 Vergleich der Lage nach MD von Substanz 24 in IGFR	86
3.24 RMSD-Plots der Komplexe	87
3.25 Korrelationen	89
3.26 Korrelationen	90

TABELLENVERZEICHNIS

3.1	Retrospektives VS	34
3.2	Fingerprintanalyse	37
3.3	Mittelwerte der AUC	38
3.4	AUC für den GOLD und GOLDScore	39
3.5	AUC für GLIDE und GLIDEScore (SP)	42
3.6	AUC für PARADOCKS und PMF	45
3.7	AUC für PARADOCKS und PSCORE	48
3.8	AUC für ROCs	51
3.9	AUC für CONSENSUS-SCORE	52
3.10	Anzahl der Verbesserungen und Verschlechterungen der AUC	52
3.11	Mittelwerte der AUC	53
3.12	Mittelwerte des EF5	53
3.13	EF bei 5 % für GOLD und GOLDScore	54
3.14	EF bei 5 % für GLIDE und GLIDEScore (SP)	55
3.15	EF bei 5 % für PARADOCKS und PMF	56
3.16	EF bei 5 % für PARADOCKS und PSCORE	57
3.17	EF für ROCs	58
3.18	Anzahl der Verbesserungen und Verschlechterungen des EF	58
3.19	Mittelwerte RIE	59
3.20	Anzahl der Verbesserungen und Verschlechterungen von RIE	60
3.21	Dockinglösungen an der PRK1	66
3.22	Ergebnisse der Dockings mit GOLD und GLIDE	73
3.23	ΔG -Werte von BS09, BS10 und BS13	76
3.24	Ergebnisse der biologischen Testung	76
3.25	K_i -Werte aus der biologischen Testung	81
3.26	Dockingergebnisse	81
3.27	Berechnete freien Energien	88

3.28 LR-MMPBSA	91
5.1 Brasca AUC	ii
5.2 Peifer AUC	ii
5.3 Matthews AUC	iii
5.4 Huang AUC	iii
5.5 Cheeseright AUC	iii
5.6 Ermoli AUC	iii
5.7 Aliagas-Martin AUC	iii
5.8 Tsou AUC	iv
5.9 Qian AUC	iv
5.10 Gaisina AUC	iv
5.11 Surya AUC	iv
5.12 Karcher AUC	iv
5.13 Schroeder AUC	v
5.14 Coumar AUC	v
5.15 Popowycz AUC	v
5.16 Menichincheri AUC	v
5.17 Renhowe AUC	vi
5.18 Howard AUC	vi
5.19 Xia AUC	vii
5.20 Brasca EF5	vii
5.21 Peifer EF5	viii
5.22 Matthews EF5	viii
5.23 Huang EF5	viii
5.24 Cheeseright EF5	viii
5.25 Ermoli EF5	ix
5.26 Aliagas-Martin EF5	ix
5.27 Tsou EF5	ix
5.28 Qian EF5	ix
5.29 Gaisina EF5	ix
5.30 Surya EF5	x
5.31 Karcher EF5	x

5.32 Schroeder EF5	x
5.33 Coumar EF5	x
5.34 Popowycz EF5	xi
5.35 Menichincheri EF5	xi
5.36 Renhowe EF5	xi
5.37 Howard EF5	xii
5.38 Xia EF5	xiii
5.39 Brasca RIE	xiii
5.40 Peifer RIE	xiv
5.41 Matthews RIE	xiv
5.42 Huang RIE	xiv
5.43 Cheeseright RIE	xiv
5.44 Ermoli RIE	xv
5.45 Aliagas-Martin RIE	xv
5.46 Tsou RIE	xv
5.47 Qian RIE	xv
5.48 Gaisina RIE	xv
5.49 Surya RIE	xvi
5.50 Karcher RIE	xvi
5.51 Schroeder RIE	xvi
5.52 Coumar RIE	xvi
5.53 Popowycz RIE	xvii
5.54 Menichincheri RIE	xvii
5.55 Renhowe RIE	xvii
5.56 Howard RIE	xvii
5.57 Xia RIE	xviii
5.58 RIE für GOLDScore	xviii
5.59 RIE für GLIDEScore	xix
5.60 RIE für PMF	xx
5.61 RIE für PScore	xxi
5.62 RIE für ROCs	xxii
5.63 RIE für CONSENSUS-SCORE	xxiii

1. EINLEITUNG

VS Methoden nehmen in der modernen Wirkstoffentwicklung eine immer größere Rolle ein. Durch die zunehmende Leistungsfähigkeit der Computer sind inzwischen auch einfache Arbeitsplatzrechner in der Lage diese Berechnungen durchzuführen. Deswegen tragen virtuelle Screening Methoden auch dazu bei die Kosten in den frühen Phasen der Wirkstoffentwicklung zu reduzieren.

Neben den etablierten Methoden kommen stetig neue Algorithmen und Implementationen derselben hinzu. Deshalb wird es zunehmend schwerer eine Aussage über die Qualität der Ergebnisse zu machen. Daher ist es von Nöten die vorhandenen Methoden an einer großen Menge an bekannten aktiven Substanzen für bestimmte Proteine zu testen, um zu prüfen, ob die Methoden in der Lage sind zwischen aktiven und inaktiven Substanzen zu unterscheiden. Es gibt eine Reihe von Publikationen zu diesem Thema, doch werden bei diesen meistens diverse Proteine verschiedener Familien betrachtet. Ausser Acht gelassen wird dabei jedoch, wie sich die Methoden innerhalb einer Proteinfamilie verhalten. Dies ist dann von besonderem Interesse, wenn die einzelnen Vertreter einer Familie sehr homolog zueinander sind, denn gerade hierbei ist es wichtig, dass neue Wirkstoffe selektiv nur das Zielprotein adressieren, und keine Aktivität an den übrigen Vertretern der Familie zeigen. Dies trifft besonders auf die Proteinfamilie der Kinasen zu. Kinasen sind an vielen Prozessen innerhalb der Zelle beteiligt (z. B. am Zellzyklus, der Signalweiterleitung um nur ein paar Beispiele zu nennen.). Gerade der Umstand, dass sie für einen fehlerfreien Zellzyklus mitverantwortlich sind, macht Kinasen als Wirkstofftarget in der Tumorforschung interessant.

1.1. Virtuelles Screening

VS ist eine computerbasierte Methode im Bereich der Wirkstoffentwicklung, um Moleküldatenbanken nach einem oder mehreren möglichen Inhibitoren eines Enzyms oder Rezeptors zu durchsuchen. Der Grundgedanke ist, aus einer großen Menge an Strukturen die Kandidaten zu filtern, die sich für eine *in vitro*-Testung am besten eignen [1, 2]. Obwohl es auf diese Weise theoretisch möglich ist den gesamten chemischen Raum von 10^{60} denkbaren Verbindungen zu durchsuchen, verwendet man sie doch hauptsächlich zum Durchsuchen von Datenbanken kommerziell verfügbarer Verbindungen und zur Verbesserung von bekannten Substanzen [3].

1.2. Kinasen

Das menschliche Genom kodiert 518 Kinasen, die sich aufgrund ihrer Homologie in sieben Gruppen unterteilen lassen. Damit stellt diese Proteinklasse eine der größten

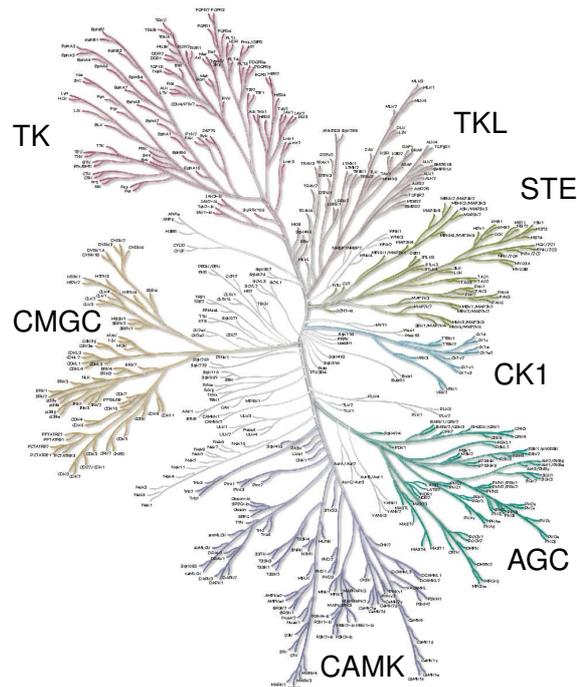


Abbildung 1.1: Dendrogramm humanes Kinom

Gruppennamen: **AGC** umfasst PKA, PKG, PKC, **CAMK** z. B. Calcium/ Camodulin-dependent protein kinase, **CK1** Casein Kinase 1, **CMGC** CDK, MAPK, GSK3, CLK-Familie, **STE** Homologe von Yeast Sterile 7, Sterile 11, Sterile 20 Kinase, **TK** Tyrosin-Kinase, **TKL** Tyrosin-Kinase-ähnlich (nach [4]).

Proteingruppen im Menschen dar. Ebenso konnte gezeigt werden, dass Kinasen innerhalb der Zelle und der Regulation ihrer Funktionen wichtige Rollen spielen [4]. Über die Signaltransduktion durch Regulierung der Phosphorylierung bestimmter Proteine [5] werden wichtige zelluläre Prozesse wie z. B. Zellwachstum, Differenzierung und Apoptose gesteuert [6]. Bei der Phosphorylierung wird die γ -Phosphatgruppe des Adenosin triphosphat (ATP) auf die Hydroxylgruppen der Aminosäuren (AS) Serin, Threonin oder Tyrosin übertragen. Abhängig von den phosphorylierten AS teilt man die Kinasen in Serin/Threonin- oder tyrosinspezifische Kinasen ein. Fehlregulation der Kinaseaktivität kann – direkt oder indirekt – zu verschiedenen Erkrankungen, einschließlich Krebs, führen [7]. Deswegen gibt es derzeit ein steigendes Interesse daran, diese Signalwege durch spezifische Kinase-Inhibitoren zu modulieren. Aus diesem Grund sind Kinasen die zweitwichtigste Gruppe als Wirkstofftarget neben den G-Protein-gekoppelten Rezeptoren [8].

1.2.1 Systematik

Obwohl die Struktur der katalytischen Domäne bei Kinasen sehr ähnlich ist (siehe Kapitel 1.2.2), gibt es Variationen in der Sequenz der Aminosäure (AS). Aufgrund dieser

Unterschiede lassen sich die Kinasen in Gruppen unterteilen, die sich wiederum in Familien aufteilen lassen (Abbildung 1.1).

AGC: Diese Gruppe ist nach den Kinasen **PKA**, **PKG** und **PKC** benannt, die alle drei Serin/Threoninspezifische Kinasen sind, die von *second messengern* wie zyklischem Adenosinmonophosphat (cAMP) oder Lipiden reguliert werden. Insgesamt besteht diese Gruppe aus 16 Familien, von denen acht wahrscheinlich schon in frühen Eukaryoten auftreten, zwei (RSK, PKC) sowohl bei Pilzen, als auch bei Metazoen, und die übrigen sechs (PKG, PKN, DMPK, YANK, RSKR, RSKL) nur in Metazoen [9, 10].

CAMK: Benannt nach den **Calcium-** und **Calmodulin-regulierten Kinasen** besteht diese Gruppe aus 22 Familien. Calmodulin steht für **Calcium-modulated protein** und ist ein Calcium bindendes regulatorisches Protein in eukaryotischen Zellen [11, 12].

CK1: Der Namensgeber dieser Gruppe ist die Familie der **Casein Kinase 1**, Serin/Threoninspezifische Kinasen, die am circadiären Rhythmus, Transport von Transkriptionsfaktoren vom Zellkern in das Cytoplasma, DNA Reparatur und DNA Transkription beteiligt sind [13]. Die Gruppe besteht aus vier weiteren Familien VRK, TTBK und TTBKL. Die letzte Familie ist bisher nur in Fadenwurm *C. elegans* gefunden worden, und besteht aus mehreren Subfamilien Worm7-Worm11 und Dual, wobei Letztere dadurch hervorsticht, dass sie über zwei Kinasedomänen pro Protein verfügt. Diese Subfamilien sind derzeit kaum erforscht, weshalb man keine Aussagen über deren Funktion machen kann [14].

CMGC: Wieder steht der Name für eine Reihe von Familien, namentlich **CDK**, **MAPK**, **GSK3** und **CLK**. Alle neun Familien, die zu dieser Gruppe gehören sind in allen Organismen hoch konserviert, was bedeutet, dass wichtige strukturelle Merkmale über alle Familien konserviert und für die gleichen Mechanismen wichtig sind [15].

STE: Diese Gruppe besteht aus drei Familien (STE20, STE11 und STE7), die sich der Reihe nach gegenseitig aktivieren, um schließlich die MAPK Familie zu aktivieren. Vertreter der STE7 Familie übernehmen dabei den letzten Schritt, die Phosphorylierung der MAPK Familie [16, 17].

TK: Diese Gruppe phosphoryliert fast nur Tyrosinreste, ganz im Gegensatz zu den meisten anderen Kinasen, die entweder Serin oder Threonin phosphorylieren. Dabei scheint es sich um die jüngste Gruppe von Kinasen zu handeln, da sie in Pflanzen (z. B. *Dictyostelium*) nicht auftritt [18]. Diese Gruppe ist sehr gut untersucht, und ihre Hauptaufgabe ist die Weiterleitung von extrazellulären Signalen in das Innere der Zelle [19].

TKL: Die sieben Unter-Gruppen (MLK, RAF, STKR, LRRK, LISK, IRAK und RIPK) dieser Gruppe (**Tyrosine kinase like**) unterscheiden sich strukturell am stärksten von den anderen Kinasen, sind aber nah mit dem Mitgliedern der TK-Gruppe verwandt. Sie sind in Metazoen weit verbreitet, fehlen aber in Pilzen völlig und in Pflanzen stellen sie bis zu 80% des Kinoms. Sie phosphorylieren Serin- und Threoninreste.

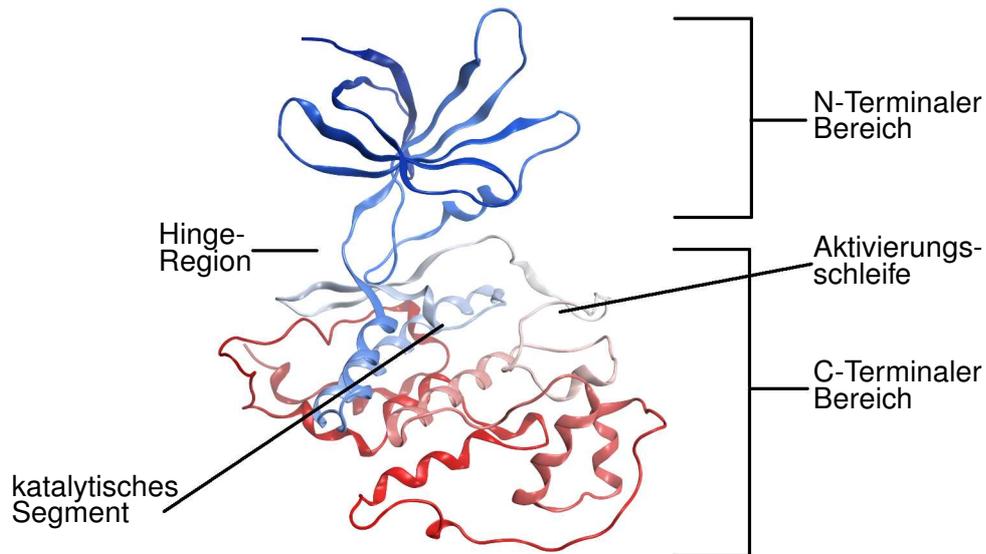


Abbildung 1.2: Struktur der Kinasedomäne

Dargestellt sind die einzelnen Sekundärstrukturelemente (PDB 1Q3D), sowie die einzelnen Bereiche der Kinasedomäne. Gezeigt ist das Proteinrückgrat (ribbon).

Auch innerhalb dieser Gruppe sind die Familien nur schwach miteinander verwandt, aber alle sind den Vertretern der TK-Gruppe ähnlich, obwohl ihnen die für die TK-Gruppe spezifischen Motive fehlen [20–22].

1.2.2 Struktur und Funktion

Allen Kinasen ist der ähnliche Aufbau der katalytischen Domäne gemein. Diese Domäne besteht aus 250 – 300 AS. Zwölf konservierte Subdomänen sind für den Aufbau der räumlichen Struktur der Kinasedomäne verantwortlich [23]. Das aktive Zentrum wird am N-Terminus durch fünf verdrehte β -Faltblätter und die glycinreiche Schleife begrenzt. Die Hinge-Region beschreibt die seitliche Begrenzung der ATP-Bindetasche. Der C-terminale Bereich besteht vorwiegend aus α -Helices (Abbildung 1.2).

Die Kinasedomäne lässt sich weiter in die Bereiche *Hydrophobe Tasche I* und *II* aufteilen, welche sich vor und hinter der Hinge-Region anschließen. Hierbei handelt es sich um Bereiche, die von hydrophoben Aminosäuren gebildet werden. Die Größe der Hydrophoben Tasche II wird durch den *Gatekeeper* gesteuert (Abbildung 1.3). Die Aminosäure an dieser Position ist nicht so wie das DFG-Motiv über die Kinasefamilien konserviert. Zudem gibt es auch innerhalb einer Familie gibt es Unterschiede. So ist z. B. Threonin in den Kinasen der TK- und TKL-Familien am häufigsten, gefolgt von Methionin. Phenylalanin in der CMCG-Familie am häufigsten anzutreffen, hier ebenfalls gefolgt von Methionin [24, 25]. Obwohl etwa 20 % aller Kinasen ein Threonin als Gatekeeper aufweisen, weisen die meisten Kinasen, für welche aktuell Inhibitoren in

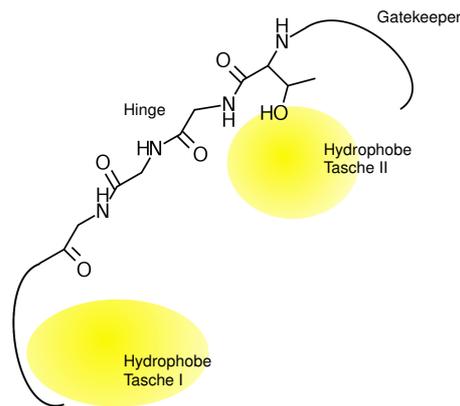


Abbildung 1.3: Hydrophobe Taschen I und II (nach [27]).

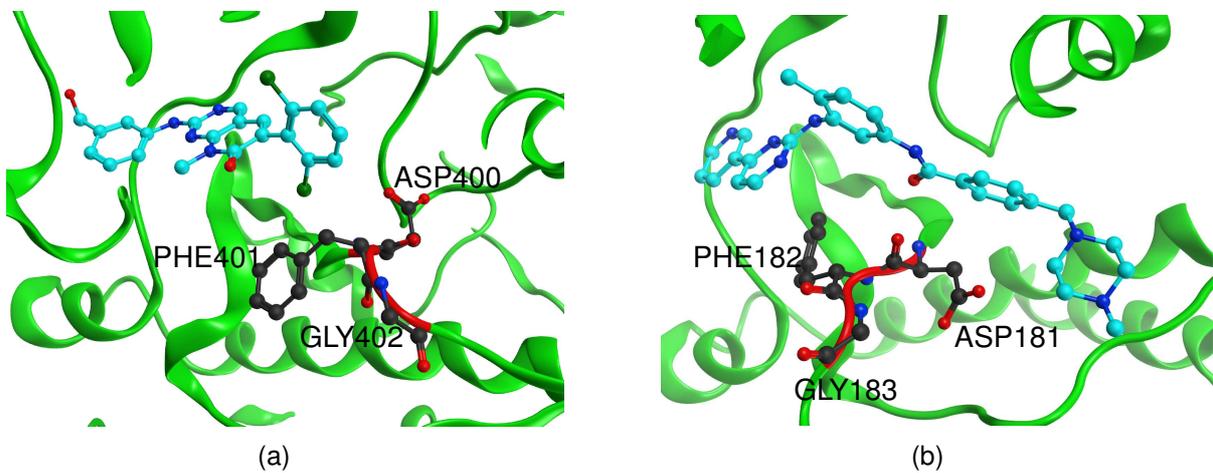


Abbildung 1.4: DFG-Motiv

in beiden Bildern ist der Inhibitor cyan gefärbt; die Kohlenstoffatome des DFG-Motivs in schwarz; das anders geformte Proteinerückgrat in rot; a) *DFG-In* Abl1 Kinase im Komplex mit PD166326 (PDB 1OPK), b) *DFG-Out* Abl1 Kinase im Komplex mit Imatinib (PDB 1IEP). Hier – sowie in der folgenden Abbildungen – erfolgt die Nummerierung anhand der Proteinsequenz.

der klinischen Verwendung sind, ein Threonin an entsprechender Stelle auf. Der Gatekeeper ist somit ein wichtiges Merkmal für die Selektivität der Inhibitoren [26].

Ein weiteres wichtiges Merkmal ist das DFG-Motiv. DFG ist die Ein-Buchstaben-Notation für die AS Asparaginsäure, Phenylalanin und Glycin. Es können zwei Konformationen vorliegen: *DFG-in* und *DFG-out* (Abbildung 1.4). Die *DFG-in*-Konformation bezeichnet man auch als aktive Konformation der Kinase, *DFG-out* entsprechend als inaktive Konformation. Bei der aktiven Konformation ist das Aktivierungssegment phosphoryliert, und hat eine Konformation, die kaum Unterschiede bei den bisherigen Kristallstrukturen aufweist. Wohingegen die inaktive Form nicht phosphoryliert ist, und aufgrund von nun fehlenden Wechselwirkungen vielfältige Konformationen einnimmt [28].

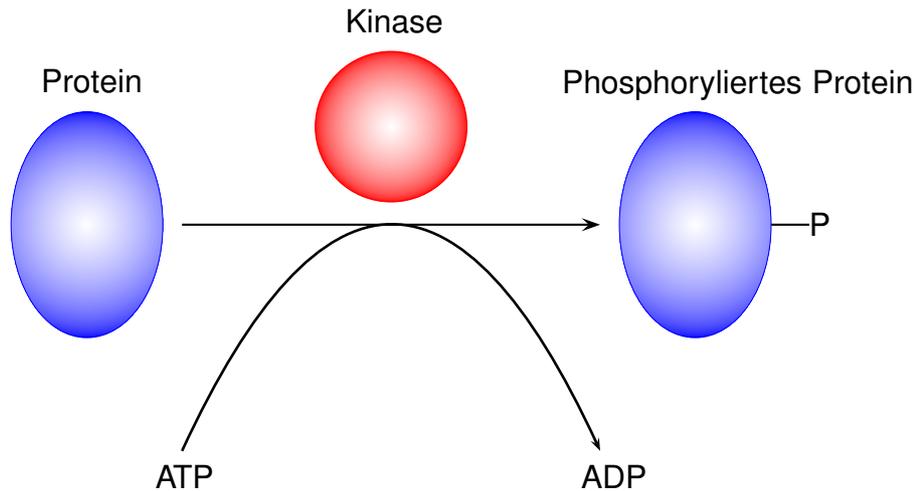


Abbildung 1.5: Phosphorylierung

Die Kinase sorgt dafür, dass die endständige Phosphatgruppe des ATP auf ein Protein übertragen wird

DFG-In: Die Aktivierungsschleife nimmt eine offene Konformation ein. Die Seitenkette des Phenylalanins ragt aus der ATP-Bindetasche heraus, während die Seitenkette der Asparaginsäure nach innen zeigt (Abbildung 1.4(a)).

DFG-Out: Die Aktivierungsschleife nimmt eine geschlossene Konformation ein, wobei die Seitenkette des Phenylalanins um ca. 10° umklappt und somit in die ATP-Bindetasche hereinragt (Abbildung 1.4(b)).

1954 wurde die erste wirkliche Kinaseaktivität beobachtet. Diese beschrieb ein Leberenzym welches Casein phosphoryliert [29]. Erst danach wurde der Mechanismus aufgeklärt der die Phosphorylase a in Phosphorylase b umwandelt [30, 31]. Abbildung 1.5 zeigt schematisch den Ablauf dieser Reaktion.

Damit die Phosphatgruppe übertragen werden kann, muss diese an die Kinase binden. Der Bindungsort dafür ist durch die Hinge-Region (Abbildung 1.6) sowie die glycinreiche-Schleife begrenzt. Diese Schleife verfügt über das konservierte Motiv **GXGX Φ G** (wobei **X** für jede AS stehen kann und Φ in der Regel für Tyrosin oder Phenylalanin steht) [32]. Dieses Motiv ist ebenfalls mit dafür verantwortlich dass die Phosphatgruppen die richtige Orientierung haben und somit der γ -Phosphatrest in Richtung der Aktivierungsschleife zeigt, um dort abgespalten und auf das Substrat übertragen zu werden [33].

1.2.3 Kinaseinhibitoren

Da Kinasen, wenn sie in ihrer Funktion gestört sind, als Onkogene fungieren können [34], sind sie ein begehrtes Ziel in der Wirkstoffentwicklung, insbesondere in der

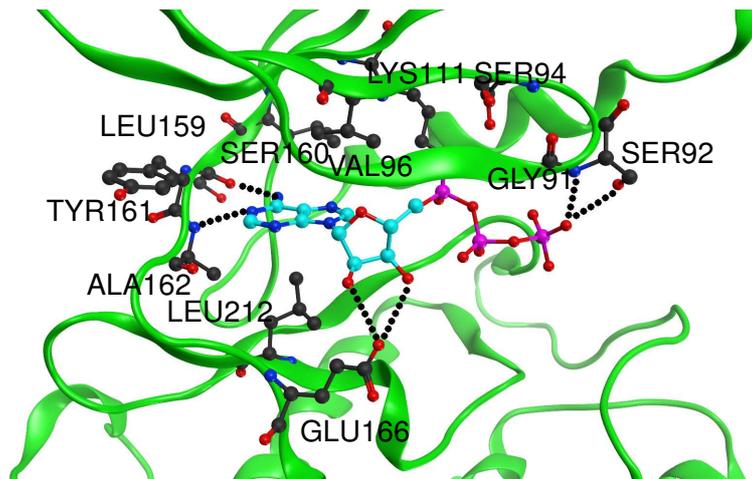


Abbildung 1.6: Hinge-Region mit ATP

die wichtigen Wasserstoffbrücken von ATP zum Proteinrückgrat sind hervorgehoben die Kohlenstoffatome des Rezeptors sind schwarz, die Phosphatreste des ATP magenta und der Adenosinteil cyan (PDB 3HRC, PDK1 Kinase).

Tumorforschung. Onkogene sind Teile des Genoms einer Zelle, die dazu führen können, dass sich das kontrollierte Zellwachstum zu einem ungebremsten Tumorstadium entwickelt. Dies passiert durch Mutationen in den Onkogenen. Kinaseinhibitoren führen dazu dass die Kinase nicht mehr aktiv ist, und die Krebszelle deswegen in ihrem Wachstum gehemmt wird und schließlich die Apoptose eintritt. Dies ist nicht der einzige Bereich in dem Kinaseinhibitoren eine wichtige Rolle spielen. So sind auch Wirkstoffe gegen rheumatische Arthritis, kardiovaskuläre Erkrankungen und Diabetes mellitus in der klinischen Entwicklung. Derzeit sind mindestens neun sog. *small molecule* Kinaseinhibitoren durch die Food and Drug Administration (FDA) der USA gegen Krebserkrankungen zugelassen [35, 36] (s. Abbildung 1.7). Neben diesen gibt es auch noch monoklonale Antikörper wie z. B. Cetuximab und Bevacizumab [37] sowie Pegaptanib, ein RNA-Aptamer [38], welche ebenfalls therapeutisch eingesetzt werden und Kinasen hemmen.

Da die Kinaseaktivität essentiell für die Funktion der Zelle ist, ist es wichtig, dass zukünftige Inhibitoren nur an der gewünschten Kinase wirken, um unerwünschte Wirkungen auszuschließen. Die Kinaseinhibitoren unterteilen sich in drei Klassen:

- ATP-kompetitive Inhibitoren
- Nicht-ATP-kompetitive Inhibitoren
- Mg^{2+}/Mn^{2+} kompetitive Inhibitoren

ATP-kompetitive Inhibitoren: 1993 wurde die Kristallstruktur der Cyclin-abhängigen Kinase 2 (CDK2) in Verbindung mit ATP aufgeklärt [39]. Inhibitoren, die eben diesen Bindungsmodus imitieren, werden als ATP-kompetitive Inhibitoren bezeichnet. Diese konkurrieren mit ATP an der Bindungsstelle der Kinase. Die Klasse der

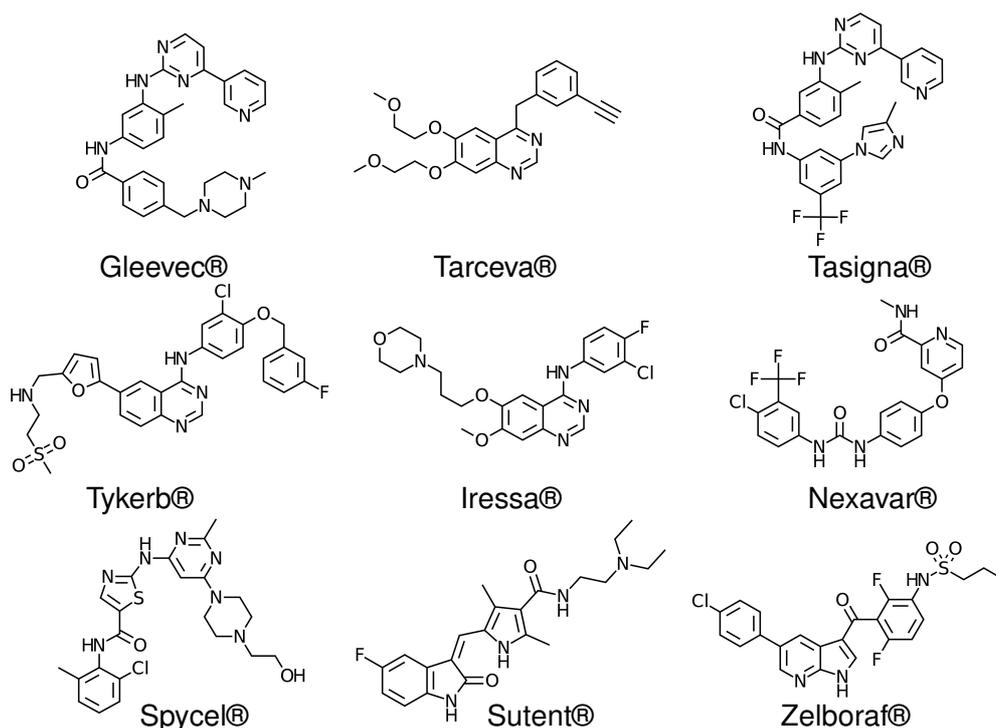


Abbildung 1.7: Kinaseinhibitoren

Von der FDA zugelassene Protein Kinaseinhibitoren (nach Grant [35] und [36]).

ATP-Kompetitiven Inhibitoren lässt sich weiter in *Typ-I-* und *Typ-II-Inhibitoren* einteilen. Diese unterscheiden sich nicht im Bindungsmodus, sondern in der Kinasekonformation an die sie binden und somit hemmen. Als dritte Gruppe gibt es noch die *kovalenten Kinaseinhibitoren*.

Typ-I-Inhibitoren können sowohl an die aktive Konformation als auch an die inaktive Konformation binden, die Typ-II-Inhibitoren binden dagegen nur an die inaktive Konformation. Die Inhibitoren bilden ein bis drei Wasserstoffbrücken zur Hinge-Region aus (Abbildung 1.8b). Da die ATP-Bindetasche einen hohen Grad an struktureller und sequenzieller Ähnlichkeit aufweist, führt dies dazu, dass die meisten Inhibitoren mehrere Kinasen hemmen. Damit sind unerwünschte Nebeneffekte und Toxizität verbunden [40]. Ein Angriffspunkt um die Selektivität zu erhöhen ist der Gatekeeper. Als Beispiel sei hier SB203580 genannt, ein selektiver Typ-I Inhibitor der Kinasen p38 α und p38 β , sowie der eng verwandten p38 γ und p38 δ . Die Selektivität beruht hier auf den unterschiedlichen AS in der Hydrophoben Tasche II (Threonin bei p38 α und p38 β bzw. Methionin bei p38 γ und p38 δ [41, 42]). Eine kleinere oder bewegliche Seitenkette führt zu einer größeren Hydrophoben Tasche II, die nahe Reste freilegt, die nun mit dem Liganden in Wechselwirkung treten können und somit die Selektivität und/ oder die Stärke der Hemmung steigern. Im Gegensatz dazu führen große Reste dazu, dass der Zugang zur Hydrophoben Tasche II eingeschränkt wird.

Eine wichtige Stoffklasse im Bereich der Typ-I-Inhibitoren sind die Bisindolylmaleimide (Abbildung 1.9). Diese leiten sich alle von Arcyriarubin A ab, welches zu einer Rei-

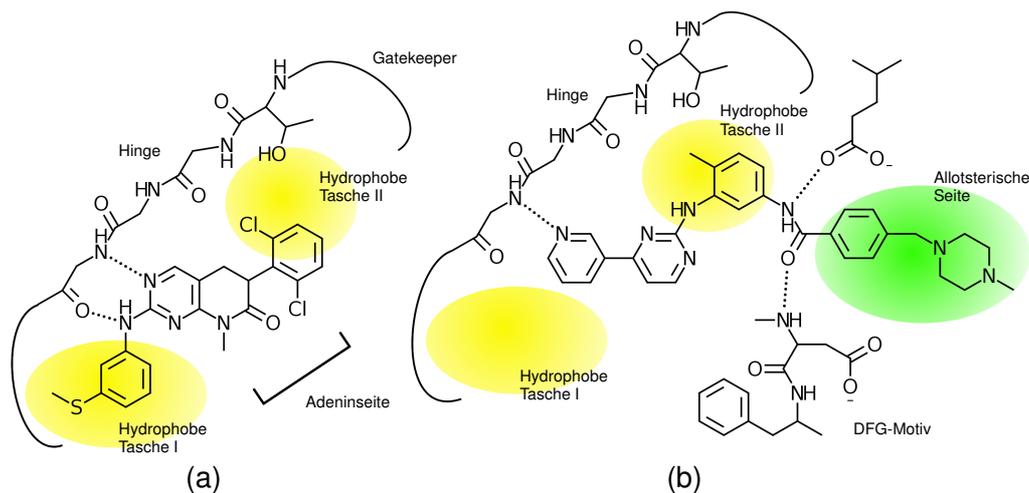


Abbildung 1.8: Typ-I- und Typ-II-Inhibitoren bei ABL1. a) mit PD166326 als Typ-I-Inhibitor b) mit Imatinib als Typ-II-Inhibitor (nach [27]).

he von Bisindolylmaleimiden gehört, die in Schleimpilzen synthetisiert werden [43]. Staurosporin wurde 1977 aus *Streptomyces staurosporeus* isoliert, und war die erste bekannte Substanz aus der Gruppe der Bisindolylmaleimide. Die Struktur von Staurosporin konnte jedoch erst 1994 aufgeklärt werden [44, 45]. Staurosporin hemmt eine Vielzahl verschiedener Kinasen, indem es an die ATP-Bindetasche bindet, und dabei eine höhere Affinität aufweist, als ATP selbst [46]. Somit zeigt Staurosporin, und damit auch die Klasse der Bisindolylmaleimide, einen ähnlichen Bindemodus wie ATP. Abbildung 1.10 zeigt den Bindemodus von Staurosporin zur Hinge-Region von GSK-3 β . Dieser Bindemodus zur Hinge-Region wird von allen Bisindolylmaleimiden eingegangen. Somit weisen alle Derivate dieser Stoffklasse eine inhibitorische Wirkung bei Kinasen auf, die jedoch unterschiedlich stark sein kann. So werden z. B. von Ro 31-8220 nur alle Isoformen der PKC inhibiert [47]. K252a hingegen wirkt sich auf die CaM- und Phosphorylase-Kinase am stärksten aus. Setzt man jedoch eine deutlich höhere Konzentration ein, dann hemmt K252a auch verschiedene Serin-/Threonin-Kinasen [48].

Typ-II-Inhibitoren binden an die inaktive Konformation der Kinase. Imatinib (Gleevec®, Novartis) war der erste Kinaseinhibitor auf dem Markt, der die inaktive Konformation der Kinase adressierte [49]. Er wird für die Behandlung von chronisch myeloischer Leukämie, gastrointestinalen Stromatumoren verwendet, wobei die Tyrosinkinase ABL gehemmt wird. Die DFG-Out Konformation öffnet eine weitere hydrophobe Tasche (Abbildung 1.8 b). Diese Tasche ist weniger konserviert als die ATP-Bindetasche und man geht davon aus, dass man hierüber die Selektivität erhöhen kann.

Kovalente Kinaseinhibitoren binden irreversibel an einen nukleophilen Cysteinrest in der ATP-Bindetasche [50, 51]. Die am meisten erforschten Kinaseinhibitoren des epidermal growth factor receptor (EGFR) sind HKI-272 [52] und CL-387785 [53]. Sie wurden entwickelt um einen Cysteinrest am Rand der ATP-Bindetasche anzugreifen der nur bei wenigen Kinasen vorkommt [54]. Durch eine *Michael-Addition* [55] werden die

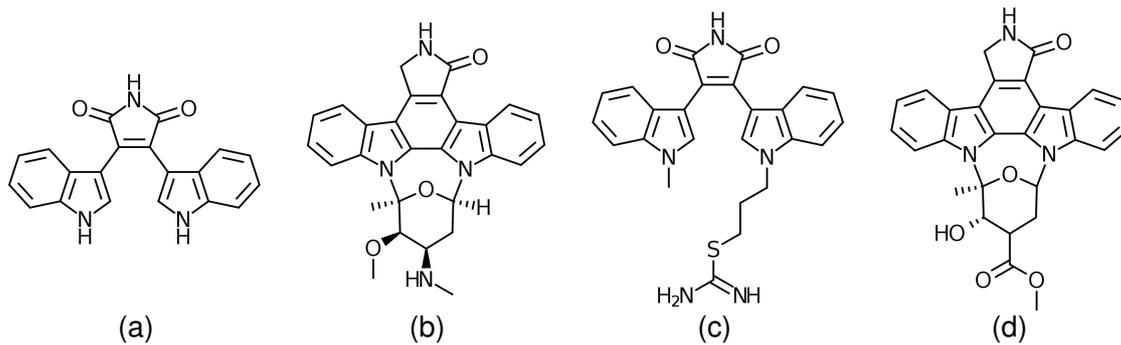


Abbildung 1.9: Bisindolylmaleimide

a) Arcyriarubin A, die Grundstruktur - ein Bis[3,4]-(1H-indol-3-yl)-1H-pyrrol-2,5-dion - von der sich die Klasse der Bisindolylmaleimide ableitet. b) Staurosporin c) Ro 31-8220 d) K252a

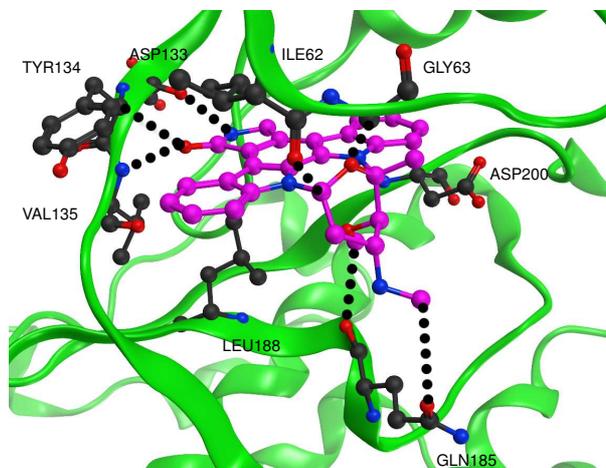


Abbildung 1.10: Bindemodus von Staurosporin

am Beispiel von GSK-3 β (PDB 1Q3D). Die Wasserstoffbrückenbindungen zur Hinge-Region sind die gleichen, wie bei ATP. Das Proteinrückgrat ist in grün dargestellt, Staurosporin in magenta.

Inhibitoren mit dem in das Solvent ragende Cystein verknüpft. Derzeit sind fünf EGFR-Inhibitoren in klinischen Testungen zur Behandlung von Lungenkrebs [51, 56, 57].

Nicht-ATP-Kompetitive Inhibitoren: Die zweite Klasse von Substanzen bindet außerhalb der ATP-Bindetasche und inhibiert die Kinase allosterisch. Substanzen die zu dieser Klasse gehören zeigen die höchste Selektivität, weil sie Bereiche der Kinase angreifen, die für diese einmalig sind. Am Besten ist die Substanz CI-1040 charakterisiert, welche MEK1 und MEK2 inhibiert, indem sie eine Tasche angreift, die mit der ATP-Bindetasche verbunden ist (Abbildung 1.11) [58].

Mg²⁺/Mn²⁺ Kompetitive Inhibitoren: Mg²⁺ oder Mn²⁺ Ionen werden innerhalb der Kinasedomäne benötigt, damit sich das ATP präzise ausrichten kann, um so die Übertragung der Phosphatgruppe auf das Protein zu ermöglichen. Andere Ionen können

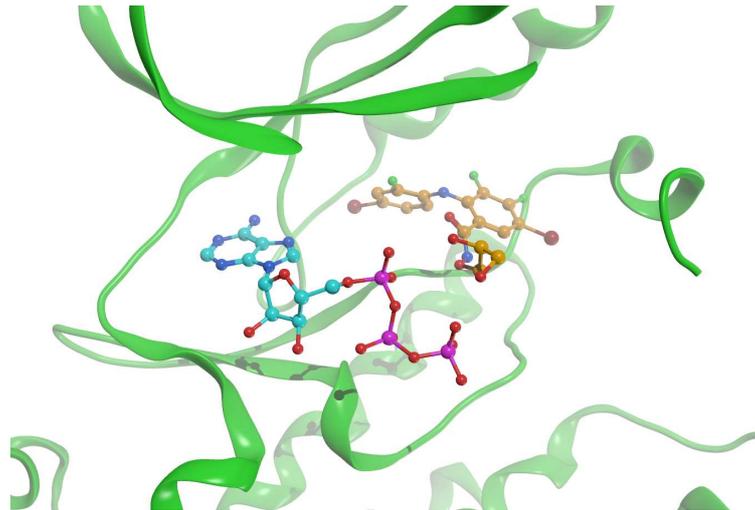


Abbildung 1.11: Allosterische Hemmung der MEK1

Ein Beispiel für die allosterische Hemmung der MEK1. Cyan ATP; orange CL-1040 Analogon (PDB 1S9J).

mit Mg^{2+} oder Mn^{2+} konkurrieren, und somit die Kinase hemmen, weil eine richtige Orientierung des ATP verhindert wird. Ein Vertreter dieser Klasse ist Lithium (Li^{2+}), welches Mg^{2+} aus der Domäne verdrängt, und somit die GSK-3 inhibiert [59].

1.3. Kinase-Familien mit Bezug zur Arbeit

1.3.1 VEGFR, EGFR, IGFR

Der *vascular endothelial growth receptor* (VEGFR), *epidermal growth factor receptor* (EGFR) und *insulin-like growth factor 1 receptor* (IGFR) gehören zur Klasse der Rezeptor-Tyrosin-Kinasen (RTK). Bei dieser Familie handelt es sich um an der Zelloberfläche angelagerte Rezeptoren mit einer hohen Affinität für Wachstumsfaktoren, Polysaccharide oder Hormone. Bisher wurden im menschlichen Genom 90 Gene für Tyrosin-Kinasen identifiziert, von denen 58 RTK sind [19]. RTK spielen nicht nur bei der Regulation von zellulären Prozessen eine wichtige Rolle, sondern ebenso bei der Entstehung von vielen Krebs-Arten [60], bei denen sie in der Regel überexprimiert werden. Die Überexpression führt zu einer permanenten Dimerisierung der einzelnen Tyrosin-Kinase-Untereinheiten und dies dann zu einer andauernden Stimulation und zum Zellwachstum [61, 7].

VEGFR: VEGFR kommt bei Menschen, Mäusen und weiteren Säugetieren in drei Subtypen (1, 2 und 3) vor. Beim Zebrafisch hingegen gibt es vier Varianten von VEGFR: *Fms-like tyrosine kinase* (Flt) 1 (diese entspricht VEGFR 1), Flt4 (entspricht VEGFR 3) und zwei weitere Gene mit der höchsten Ähnlichkeit zu VEGFR 2, die als *kinase insert domain receptor* (Kdr) *like* und Kdr bezeichnet werden [62]. Durch alternatives Splicing können sie entweder membrangebunden (mbVEGFR) oder gelöst

(sVEGFR) in der Zelle vorliegen [63–65]. Das Substrat der *vascular endothelial growth factor* (VEGF) ist als Signalmolekül wichtig für die Angiogenese, und die Vaskulogenese. Die Bindung des Substrats an VEGFR findet an der N-Terminalen-Region an der extrazellulären Domäne statt. Danach kommt es zu einer Dimerisierung zu einem Hetero- oder Homodimer [66]. Die Dimerisierung wird durch Kontaktpunkte zwischen den extrazellulären Domänen an der Immunglobulin-ähnlichen (Ig) Schleife 7 stabilisiert [67, 68]. Die genaue Ausrichtung der Monomere wird dabei von den Transmembrandomänen beeinflusst [69]. Die daraus resultierende starre Ausrichtung der Monomere zueinander ist wichtig, damit sich die intrazellulären Kinasedomänen ausrichten [67] und somit durch Auto- oder Transphosphorylierung aktiviert werden können.

VEGFR ist aktuell ein vielversprechendes Ziel in der Behandlung des Glioblastoms, eines bösartigen Hirntumors. Derzeit sind verschiedene Kinaseinhibitoren in weiteren Untersuchungen, wovon Cediranib aktuell in einer klinischen-Phase-III-Studie [70, 71] ist.

EGFR: Der EGFR ist ein membrangebundenes Protein und der Rezeptor für den *epidermal growth factor* (EGF) [72]. Der Rezeptor gehört zur Familie der ERBB Rezeptoren, einer Untergruppe von vier eng verwandten RTKs: EGFR (ERBB-1), HER2/c-neu (ERBB-2), HER3 (ERBB-3) und HER4 (ERBB-4) [73]. EGF ist für das Zellwachstum von Hautzellen und das Verhindern der Apoptose verantwortlich [74]. Ein weiterer Ligand für EGFR ist der *transforming growth factor* α ($TGF\alpha$), der besonders das Zellwachstum von Nervenzellen im erwachsenen Gehirn anregt [75]. Die Aktivität des EGFR kommt durch den gleichen Mechanismus wie bei VEGFR zustande. Hier führt die Dimerisation zu einer Stimulierung der intrinsischen intrazellulären Kinase Aktivität. Deswegen kommt es zu einer Autophosphorylierung von mehreren Tyrosinresten in der C-Terminalen Domäne [76], die dann zu einer downstream Aktivierung von verschiedenen Proteinen führt, die mit den phosphorylierten Tyrosinen in Wechselwirkung steht [77].

EGFR ist ein vielversprechendes Ziel zur Behandlung eines Lungenkarzinoms. Hierbei ist der Wirkstoff Icotinib derzeit in klinischen-Phase-III-Studien [78, 79].

IGFR: Der IGFR wird von den Wachstumsfaktoren *insuline-like growth factor 1* (IGF-1) und dem verwandten *insuline-like growth factor 2* (IGF-2) aktiviert. Er gehört ebenfalls zur Klasse der Tyrosin-Kinasen und ist in der Zellmembran lokalisiert [80]. Dabei hat der IGFR eine etwa 70 %ige Homologie zu dem Insulin-Rezeptor und obwohl IGFR IGF-2 und Insulin zusammen mit IGF-1 bindet, zeigt er eine geringere Affinität. Die Affinitätsraten für IGF-1, IGF-2 und Insulin betragen 1000:10:1 [81, 82]. IGFR ist aus zwei α - und zwei β -Untereinheiten aufgebaut. Die α - und die β -Untereinheit werden aus der gleichen Precursor-mRNA synthetisiert. Danach wird der Precursor glycosyliert, proteolytisch gespalten, und mittels *cross-linking* über Cysteinreste gebunden um eine funktionfähige transmembrane $\alpha\beta$ -Kette zu bilden [83]. Ebenso sind die beiden α -Untereinheiten über Disulfidbrücken miteinander verknüpft, um so ein funktionfähiges $\alpha_2\beta_2$ -Heterotetramer zu bilden, bei dem die α -Kette im extrazellulären Bereich liegt, während die β -Untereinheit im Zellinneren ist und die Membran durchspannt [84].

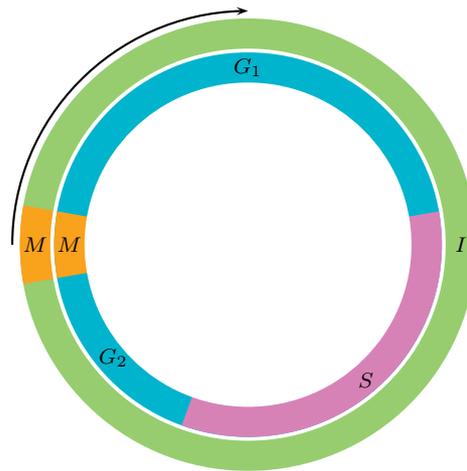


Abbildung 1.12: Zellzyklus

in einer schematischen Darstellung. (Die einzelnen Phasen der Mitose sind nicht gezeigt.) Der äußere Ring zeigt die einfache Unterteilung in Mitose und Interphase. Der innere Ring stellt die einzelnen Schritte der Interphase G₁-, S- und G₂-Phase dar. Eine typische eukaryotische Zelle in einer Kultur teilt sich etwa alle 24 Stunden einmal. Dieser Zyklus lässt sich in zwei Bereiche unterteilen: die *Interphase*, in der sich die Zelle zu etwa 95% innerhalb eines Zyklusses befindet und die *Mitose*. Beide Bereiche können in weitere Phasen unterteilt werden. Die Interphase teilt sich dabei noch in die G₁-Phase, S-Phase – hier findet die Replikation der DNA für die spätere Mitose statt – und die G₂-Phase auf. Innerhalb der G₂-Phase kommt es zur Synthese von verschiedenen Proteinen, die für die anschließende Mitose benötigt werden. Im Anschluss an die G₂-Phase folgt dann die eigentliche Mitose.

IGFR wird mit verschiedenen Krebs-Arten wie Brust-, Prostata-, und Lungenkrebs in Verbindung gebracht [85–87].

1.3.2 MYT1

Die Myelin transcriptions factor 1-Kinase (MYT1) zählt zu den WEE-Kinasen, zeigt aber deutliche Unterschiede zu den anderen Vertretern dieser Familie. Im Gegensatz zu WEE-1 und WEE-1b, die im Zellkern vorliegen, ist die MYT1 am Endoplasmatischen Retikulum im Zytoplasma lokalisiert, und phosphoryliert Cdc2 an Thr14 und Tyr15 in Abhängigkeit von Cyclin, während WEE-1 nur Tyr15 phosphoryliert. Durch die Phosphorylierung wird Cdc2 inaktiviert, dadurch ist die Zelle nicht der Lage in die mitotische Phase überzugehen und sammelt so beim G₂/M-Übergang den vierfachen DNA Gehalt an. MYT1 ist eine dualspezifische Threonin/Tyrosin-Kinase und spielt als Regulator im G₂/M-Übergang eine wichtige Rolle [88–91]. Abbildung 1.12 gibt eine Übersicht über den Zellzyklus.

Abbildung 1.12 zeigt eine vereinfachte Darstellung. Auf die weiteren Phasen der Mitose wird hier nicht weiter eingegangen, da sie für das Verständnis nicht relevant sind.

1.3.3 PRK1

PRK1, oder auch PKN1 genannt [92], ist eine Serin/Threonin-spezifische Proteinkinase und gehört zur Familie der Proteinkinase C. Die Kinasen der PRK-Familie kontrollieren verschiedene Funktionen innerhalb der Zelle, wie z. B. Zellzyklus [93], Transport von Vesikeln [94] und Apoptose [95]. Jedoch ist die Aktivierung der Genexpression des Androgenrezeptors eine der Hauptaufgaben der PRK1 und ist somit ein interessantes Ziel für die Behandlung von Prostatakrebs [96]. Ebenfalls ist PRK1 an der Milchsekretion beteiligt, in dem sie die *tight junction*, oder auch Schlussleiste, in Milchdrüsen reguliert [97]. Aktiviert wird PRK1 von der GTPase Rho oder von Fettsäuren wie z. B. Arachidonsäure [98]. Die Mitglieder dieser Familie definieren sich durch ihre Homologie in der aminoterminalen Kinase-Domäne und durch eine Consensussequenz der zu phosphorylierenden Region **L(I)XXQXTG** [99].

1.4. Zielsetzung

Das Ziel dieser Arbeit ist die Evaluierung von Docking und Ligandenbasierten virtuellen Screening Methoden für Proteinkinasen. Als Grundlage für diese Evaluierung werden aus der Literatur 19 Datensätze mit bekannten aktiven Liganden entnommen. Mit diesen Datensätzen werden nicht nur unterschiedliche Virtuelles Screening Methoden betrachtet, sondern auch der Einfluß der jeweiligen Scoring-Funktion auf das Ergebnis analysiert. In dieser Arbeit wird zum ersten mal das virtuelle Screening auf eine große, homogene Proteinfamilie – die der Proteinkinasen – angewandt. Diese Proteinfamilie eignet sich besonders gut für einen genauen Vergleich von unterschiedlichen virtuellen Screening Methoden, da die einzelnen Vertreter alle über einen konservierten Bereich – die Hinge-Region – verfügen, der für die Aktivität wichtig ist. Zusätzlich soll der Einfluß kristallisierter Wassermoleküle in der Bindetasche untersucht werden.

Diese daraus gewonnenen Ergebnisse werden im Anschluß auf die Untersuchung von weiteren Kinasen angewandt. Zum einen wird dabei untersucht, ob die aufgrund von Dockinlösungen gefundenen potentiellen Inhibitoren auch bei einer biologischen Testung eine inhibitorische Wirkung haben, und zum anderen ob sich anhand der Dockinglösung die schon bekannte inhibitorische Wirkung erklären lässt. Hierzu werden Moleküldynamiksimulationen verwendet, um zu testen, ob eine Korrelation zwischen berechneten freien Energien und pK_i -Werten existiert. Diese Fragestellungen sind für alle Proteinkinasen interessant. Jedoch gibt es nicht für jede Kinase eine Kristallstruktur, und somit müssen Homologiemodelle erstellt werden, an denen die Dockingstudien und die folgenden Moleküldynamiksimulationen durchgeführt werden. Deshalb liegt ein weiterer Aspekt dieser Arbeit darauf, ob es für die Aussagekraft der Ergebnisse einen Unterschied macht, ob die Daten über den räumlichen Aufbau des Protein aus einer Kristallstruktur kommen, oder aus einem Homologiemodell.

Die validierten Methoden und Modelle sollten im Falle von PRK1, EGFR und MYT1 für die Entwicklung neuer Inhibitoren eingesetzt werden.

2. MATERIAL UND METHODEN

2.1. Virtuelles Screening

Es gibt verschiedene Methoden des VS, aber alle lassen sich in zwei Gruppen unterteilen. Zum einen gibt es die ligandenbasierten Ansätze (LBVS), zum anderen strukturbasierte Methoden (SBVS). Ligandbasierte Methoden funktionieren nach dem Prinzip, dass eine ähnliche Struktur eines Liganden eine ähnliche biologische Wirkung hervorruft. Schon bekannte Liganden dienen bei dieser Ähnlichkeitssuche als Vorlage. Bei Strukturbasierten Methoden ist die Struktur des Zielproteins von größerer Bedeutung. [100, 101].

2.2. Strukturbasierte VS Methoden

Unter *Docking* versteht man vereinfacht ausgedrückt ein *Einpassen* eines Liganden in ein Protein, um eine favorisierte Orientierung zu finden, die einen stabilen Komplex bildet. Ein Dockingprogramm unterteilt sich in mindestens drei Abschnitte:

- Konformationsgenerierung
- Platzieren der Konformationen in der Bindetasche
- Bewerten der gefundenen Dockinglösung

Jeder dieser Schritte ist auf seine Art und Weise wichtig für die Funktionalität des Programms. Nicht immer sind diese Schritte diskret voneinander getrennt, sondern gehen fließend ineinander über oder überschneiden sich.

2.2.1 Suchalgorithmen

Die ersten Suchalgorithmen für dieses Problem entstanden in den 80er Jahren des 20. Jahrhunderts [102], deren erste Anwendung in dem Dockingprogramm DOCK veröffentlicht wurde [103]. Dieser erste Algorithmus geht dabei sowohl von einer rigiden Protein- als auch von einer rigiden Ligandenstruktur aus. Mittlerweile ist eine Vielzahl von weiteren Algorithmen in der wissenschaftlichen Anwendung etabliert, um das Dockingproblem zu lösen. Aktuelle Algorithmen sind zudem in der Lage nicht nur die Liganden als flexible Moleküle zu betrachten, sondern auch einige Seitenketten des Zielproteins. Um die Position des Liganden in der Bindetasche zu verbessern gibt es verschiedene Optimierungsverfahren. Die wichtigsten sind der *Genetische*

Algorithmus (GA) [104], die *Partikelschwarm Optimierung* (PSO) [105], *Monte Carlo* (MC) [106] und *Ameisenkolonie Optimierung* (ACO) [107].

Der GA nähert sich durch Evolution an eine optimale Lösung an. Nur die Individuen einer Population, die über eine bestimmte Fitness verfügen dürfen sich vermehren. Zusätzlich kommt es zu Rekombination und Mutation. Dies wiederholt sich solange, bis ein Abbruchkriterium erfüllt ist. Meist ist dies eine vorher bestimmte Anzahl an Wiederholungen. Dieser Algorithmus findet z. B. im Programm GOLD Anwendung [108].

Bei der PSO gibt es zu Beginn eine Menge an Lösungen, die als Partikel dargestellt werden, die sich alle im Suchraum bewegen. Diese Bewegung wird durch die jeweils bekannte beste Lösung beeinflusst, die der jeweilige Partikel kennt, und durch die beste Lösung im gesamten System. Mit jeder Iteration werden die Werte aktualisiert, und somit nähern sich alle Partikel der optimalen Lösung an. Die PSO wird bei SODOCK, einer Erweiterung zu AutoDock, und PARADOCKS eingesetzt [109].

MC Methoden gehen von einer Lösung aus, die zufällig über eine festgelegte Menge an Schritten verändert wird. Nach jedem Schritt wird anhand einer Bewertungsfunktion geprüft, ob sich die neue Lösung gegenüber der vorherigen verbessert hat. Wenn ja, dann wird mit der neuen Lösung weiter gemacht. Dieser Algorithmus ist z. B. in GLIDE implementiert [110].

2.2.2 Scoring-Funktion

Neben den Suchalgorithmen stellt die *Scoring-Funktion* (SF), oder auch Bewertungsfunktion, einen weiteren wichtigen Bestandteil für die Programme dar. Neben dem einfachen Messen von sterischen und elektrostatischen Ergänzungen, welches meist ein initialer Schritt beim Docking ist, berechnen viele SF die freie Bindungsenergie eines Rezeptor-Ligand-Komplexes in wässriger Lösung. Die wenigsten Methoden durchlaufen dabei den vollständigen thermodynamischen Zyklus, der ein Teil des Bindungsprozesses ist [111]. Allgemein werden die Beiträge der Rezeptor-Liganden-Bindung additiv betrachtet [112].

Beim Docking gibt es drei Anwendungen für SF. Als erstes muss eine SF in der Lage sein, den Bindungsmodus von einem Liganden in einem Protein vorherzusagen zu können. Während eines Dockings werden viele Posen des Liganden in der Binde-tasche generiert. Dabei ist es die Aufgabe der SF, die Posen zu bewerten und mit einem Rang zu versehen. Dabei wird bestimmt, wie stark die Bindung der einzelnen Elemente dieses Komplexes ist. Im optimalen Fall würde ein experimentell bestimmter Bindungsmodus von einer SF am besten bewertet [113]. Eine weitere Aufgabe ist es die Bindungsaffinität zwischen Ligand und Protein vorherzusagen. Die dritte Aufgabe einer SF ist es potentielle neue Leitstrukturen durch VS für ein Zielprotein zu finden [114]. D. h. die Strukturen einer Datenbank entsprechend der zu erwartenden Aktivität zu bewerten.

Derzeit gibt es verschiedene Scoringmethoden:

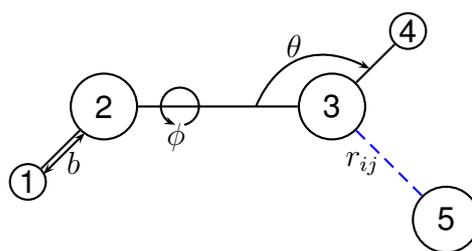


Abbildung 2.1: Atomare Wechselwirkungen

die in der allgemeinen Form eines Kraftfeldes berücksichtigt werden. Nummerierte Kreise stehen für Atome, die schwarzen Linien stehen für kovalente Bindungen, die blaue Linie für eine nicht-kovalente Wechselwirkung. b , θ , ϕ , r_{ij} sind die Einzelbeiträge zur allgemeinen Gleichung, modifiziert nach [117]

Kraftfeldbasierte Methoden: Kraftfeldbasierte Methoden versuchen die atomaren Interaktionen in einem System annähernd zu beschreiben. Das Kraftfeld dient dazu die Wechselwirkungen zwischen den Atomen zu beschreiben. Dabei versucht man die freie Energie der Bindung in Lösung zu bestimmen, indem man die Enthalpie des Systems in einer Gasphase einschätzt [115]. Auch wenn es sich hierbei nur um eine Abschätzung handelt, wurde schon gezeigt, dass eine gute Korrelation vorliegen kann. Jedoch berücksichtigen einige Kraftfelder dabei nicht die Polarisierbarkeit der Umgebung, weshalb die elektrostatischen Interaktionen von geladenen Atomen nur unzureichend betrachtet werden, was dazu geführt hat, dass polarisierbare Kraftfelder entwickelt wurden [116].

Hierbei unterscheidet man zwischen gebundener (E_{geb}) und nicht-gebundener (E_{ngeb}) Energien, die jedoch in der Summe die gesamte Energie (E_{ges}) des Systems beschreiben (siehe Formel 2.1). Abbildung 2.1 zeigt die jeweils einzelnen Teilbeiträge, aus denen sich die gebundene Energie (siehe Formel 2.2) und ungebundene Energie (siehe Formel 2.3) zusammensetzt.

$$E_{ges} = E_{geb} + E_{ngeb} \quad (2.1)$$

$$E_{geb} = E_{Bindung} + E_{Tor} + E_{Winkel} \quad (2.2)$$

$$E_{ngeb} = E_{Ele} + E_{vdW} \quad (2.3)$$

Die einzelnen Teilterme für die gebundene Energie sind die *Bindungslängenergie* ($E_{Bindung}$), *Bindungsrotationsenergie* (E_{Tor} , Torsion), *Winkelenergie* (E_{Winkel}) und die nichtgebundene Energie wird aus den Teiltermen für die *van-der-Waals-Energie* (E_{vdW} , Lennard-Jones-Term) und der Coulombenergie (E_{Ele}) gebildet.

Diese einzelnen Terme lassen sich wie folgt weiter beschreiben. In den Formeln 2.4–2.6 sind die jeweiligen Einzelbeiträge für die gebundenen Energien aufgeführt. Die Bindungslängen- (siehe Formel 2.4) und Winkelenergie (siehe Formel 2.5) werden durch harmonische Potentiale beschrieben während die Bindungsrotationsenergie (siehe Formel 2.6) durch eine Cosinusfunktion beschrieben wird. Die Veränderung der

Bindungslänge b , des Winkels ϕ , θ und der Gleichgewichtsgröße (b_0, θ_0, ϕ_0) fließt in die Berechnung mit ein [118–120]:

$$E_{Bindung} = \sum_{ij} k_{ij}^b (b_i - b_j)^2 \quad (2.4)$$

$$E_{Winkel} = \sum_{ijk} k_{ijk}^\theta (\theta_{ijk} - \theta_{ijk0})^2 \quad (2.5)$$

$$E_{Tor} = \sum_{ijkl} k_{ijkl}^\phi (1 + \cos(n\phi_{ijkl} - \phi_{ijkl0})) \quad (2.6)$$

Die Formeln 2.7–2.8 beschreiben die Beiträge zu den nicht-gebundenen Wechselwirkungen. Die van-der-Waals-Energie wird durch ein Lennard-Jones-Potential beschrieben. Ein Lennard-Jones-Potential beschreibt die Wechselwirkung zwischen ungeladenen, nicht kovalent aneinander gebundenen Atomen, wobei in (2.7) in Bezug auf den Abstand (r) die anziehenden Beiträge (A_{ij}) und die abstoßenden Beiträge (B_{ij}) mit einfließen.

Die elektrostatischen Interaktionen zwischen den einzelnen Ladungen q_i und q_j werden durch den Coulombterm (2.8) beschrieben. Zusätzlich dazu sind die Dielektrizitätskonstante ε_0 und der Abstand r im Divisor.

$$E_{vdW} = \sum_{\text{nichtgebundenePaare}} \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \quad (2.7)$$

$$E_{Ele} = \sum_{\text{nichtgebundenePaare}} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \quad (2.8)$$

Die Energierterme für nicht-gebundene Wechselwirkungen der Programme AMBER [121, 122] und CHARMM [123] werden in abgewandelter Form in unterschiedlichen Docking-Programmen als SF verwendet [124].

Empirische Scoring-Funktionen: Empirische Scoring-Funktionen basieren darauf, dass sie die Anzahl der unterschiedlichen Interaktionen zwischen zwei an der Bindung beteiligten Partnern zählen [125] und so die freie Bindungsenergie (ΔG_{Bind}) abschätzen, wobei die Summe um einen Gewichtungskoeffizient ΔG_i (siehe Formel 2.9) erweitert wird. f_i ist eine geometrische Funktion. Dabei dient entweder die Anzahl der an der Bindung beteiligten Atome vom Ligand und Rezeptor als Grundlage, oder aber die Änderung der Solvent-Accessible-Surface-Area ($\Delta SASA$) des Komplexes, verglichen mit den jeweiligen unkomplexierten Protein und Liganden.

$$\Delta G_{Bind} = \sum \Delta G_i f_i(r_l, r_p) \quad (2.9)$$

Die empirischen SF stellen jeweils einen eigenen Term für Wasserstoffbrücken, ionische Interaktionen, hydrophobe Interaktionen und die Bindungsentropie und werden jeweils auf andere Weise berechnet:

Wasserstoffbrücken werden im allgemeinen nur anhand der Anzahl der Donatoren und Akzeptoren bewertet, die einen bestimmten Abstand zueinander haben [126, 125]

Ionische Interaktionen werden auf die gleiche Art wie die Wasserstoffbrücken bewertet, wobei Salzbrücken und Wasserstoffbrücken zwischen geladenen Atomen besser bewertet werden als reine elektrostatische Wechselwirkungen

Hydrophobe Wechselwirkungen werden anhand der Kontaktfläche zwischen Protein und Ligand im Komplex berechnet und stellen die treibende Kraft dar, die zur Bildung eines Komplexes führt [126]

Entropie wird meist nur aus der Anzahl der frei drehbaren Bindungen des Liganden abgeleitet. Diese Vorgehensweise ist zwar stark vereinfacht, reicht in den meisten Fällen aber aus. Zudem ist die Entropie so leichter zu berechnen, da sie nur eine Funktion des Liganden darstellt [126, 125]

Wissensbasierte Methoden: In 3D-Datenbanken, in denen Ligand-Protein-Komplexe gespeichert sind, konnte man die intermolekularen Wechselwirkungen der am Komplex beteiligten Teilchen zueinander auswerten, und daraus *potentials of mean-force* (PMF) (siehe Formel 2.10) ableiten. Diese Idee beruht auf der Annahme, dass nahe, intermolekulare Wechselwirkungen zwischen bestimmten Atomtypen oder funktionellen Gruppen häufiger auftauchen als bei einer zufälligen Verteilung und somit energetisch günstiger sind. Die Idee der mean-force stammt aus der Physik, und kann mit einem einfachen Fluidsystem mit N Partikeln beschrieben werden. Die Positionen der Partikel sind r_1, \dots, r_N [127]. Beschreibt es einen Wert für $w^{(n)}(r_1, \dots, r_N)$ durch

$$g^{(n)}(r_1, \dots, r_N) \equiv e^{-\beta w^{(n)}(r_1, \dots, r_n)} \quad (2.10)$$

wobei $g^{(n)}$ als Korrelationsfunktion bezeichnet wird und $\beta = 1/k_B T$ ist, wobei k_B die Boltzmann-Konstante ist, und T die absolute Temperatur des Systems. Dann kann das System wie folgt beschrieben werden [127]

$$-\nabla_j w^{(n)} = \frac{\int \dots \int e^{-\beta U} (\nabla_j U) dx_{n+1} \dots dx_N}{\int \dots \int e^{-\beta U} dx_{n+1} \dots dx_N}, \quad j = 1, 2, \dots, n \quad (2.11)$$

Hierbei ist U die gesamte potentielle Energie des Systems und kann nach (2.12) bestimmt werden

$$U = \sum_{i < j}^N u(r_{ij}) \quad (2.12)$$

wobei $u(r_{ij})$ das Interaktionspotential zwischen den Atomen ist, r_{ij} der Abstand zwischen den zwei Partikeln i und j , und die Summe über alle möglichen Atompaare im System gebildet wird. Da $\nabla_j U$ die Kraft ist, die auf den Partikel j in einer bestimmten Konfiguration r_1, \dots, r_N einwirkt, ist $-\nabla_j w^{(n)}$ die mean-force über alle Konfigurationen der Partikel $n+1, \dots, N$ die auf den Partikel j wirken, wenn die Partikel $1, \dots, n$ an einem festen Punkt sind. $w^{(n)}$ stellt also das Potential dar, dass mit der mean-force auf

den Partikel j wirkt, und ist somit das PMF [128].

Ausgehend von der Physik wurde das PMF als erstes verwendet um Protein-Struktur-Modelle zu evaluieren [129], und schrittweise erweitert zu werden [130–134].

Verschiedene Wissensbasierte Methoden haben sich in den letzten Jahren etabliert. Beispiele dafür sind:

PMF04 von Muegge ist eine verbesserte Version von PMF99. Für die neue Version wurden über 7000 Protein-Liganden-Komplexe untersucht, und daraus entsprechende Wechselwirkungen der verschiedenen Atomtypen zueinander abgeleitet [135]

DrugScore verwendet Paarpotentiale von Protein-Ligand-Komplexen. Desweiteren wird hier noch ein Term für die Solventzugängliche Oberfläche (*SAS*, solvent-accessible surface) verwendet [136]

DrugScore^{CSD} verwendet zusätzlich zu den aus Protein-Ligand-Komplexen ermittelten Wechselwirkungen des DrugScore noch zusätzlich Paarpotentiale, die aus ungebundenen Wechselwirkungen kleiner Moleküle zueinander abgeleitet wurden [137]

2.2.3 Consensus-Score

Der **CONSENSUS-SCORE** ist eine Kombination verschiedener normalisierter Scoring-Funktionen. Dabei werden im einfachsten Fall die Scores der jeweiligen Docking-Lösungen addiert und überprüft, ob sich die Rangfolge verändert. Um die einzelnen Werte der unterschiedlichen Scoring-Funktionen vergleichbarer zu machen und um zu verhindern, dass eine Scoring-Funktion einen zu großen Einfluß auf die Summe nimmt, werden die Werte normalisiert [138–140]. Neben dieser Methode gibt es noch weitere Varianten des **CONSENSUS-SCORE** [141].

rank-by-number verwendet den Mittelwert aller Scores für eine Substanz

rank-by-rank beruht darauf, dass die Substanzen zuerst nach ihren Scores für die jeweilige Scoring-Funktion (SF) sortiert werden, und danach der Mittelwert des Ranges von allen SF der jeweiligen Substanz gebildet wird

rank-by-vote verfolgt einen anderen Ansatz. Hierbei wird nur geprüft, ob eine Substanz innerhalb der z. B. ersten 5% für eine SF zu finden ist. Trifft dies zu, dann bekommt dieser Kandidat eine 1 für diese SF, wenn nicht, dann eine 0. Somit können die getesteten Kandidaten Werte zwischen 0 und der Summe aller SF annehmen.

Dies lässt sich noch weiter differenzieren, ob man nur die jeweils am besten bewertete Pose eines Liganden betrachtet, oder aber mehrere Posen des gleichen Liganden in die Bildung des **CONSENSUS-SCORE** einfließen lässt.

2.3. Datenbank Filterung

An die Strukturen in den Moleküldatenbanken werden besondere Anforderungen im Bezug auf ihre Bioverfügbarkeit gestellt. Hierzu gibt es verschiedene Methoden diese Eigenschaften abzuschätzen. Eine der bekanntesten Regeln hierzu ist die *Lipinski's Rule of five*. Diese Regel sagt für chemische Substanzen eine gute Bioverfügbarkeit voraus, wenn sie mindestens drei der folgenden Bedingungen erfüllt:

- nicht mehr als fünf Wasserstoffbrückendonatoren
- eine molare Masse von nicht mehr als 500 Da
- nicht mehr als zehn Wasserstoffbrückenakzeptoren
- einen Verteilungskoeffizient (Log P) zwischen Oktanol und Wasser von nicht mehr als fünf

Diese Regel beruht auf der Verteilung der berechneten Eigenschaften von mehreren tausend Wirkstoffen. Es gibt allerdings auch Wirkstoffe, die diese Regeln nicht erfüllen. Einige Klassen, die ausserhalb dieser Beschreibung liegen sind:

- Antibiotika
- Antimykotika
- Vitamine
- Herzglykoside

Man geht davon aus, dass die orale Bioverfügbarkeit dieser Wirkstoffklassen daher rührt, dass sie strukturelle Merkmale aufweisen, die es ihnen erlaubt als Substrate für Transporterproteine zu dienen [142].

Es gibt auch Erweiterungen dieser Regel. Ghose *et al.* haben so die „*druglikeness*“ näher beschrieben [143]:

- Log P liegt in einem Bereich zwischen -0.4 und +5.6
- molare Masse liegt zwischen 160 und 480 Da
- Anzahl der Schweratome liegt zwischen 20 und 70
- molekulare Brechung liegt zwischen 40 und 130

Nach der Rule of five und den Kriterien für die druglikeness kann man vorhandene Strukturdatenbanken schon vor dem eigentlichen Screening filtern. Dies hat den Vorteil, dass man gleich zu Beginn solche Substanzen nicht weiter betrachtet, die diese Bedingungen nicht erfüllen und so für eine biologische Testung nicht in Frage kommen. Insgesamt liegt der Vorteil des VS darin, dass schnell und verhältnismäßig günstig potentielle Hits getestet werden können, ohne dass diese aufwendig synthetisiert werden müssen [144, 145].

2.4. Ähnlichkeitsbasierte Suche

Verschiedene Methoden der ähnlichkeitsbasierten Suche haben in den letzten Jahren an Popularität gewonnen. Da ebenfalls die Leistung von Computern zugenommen hat, können diese Methoden auch auf sehr große Datenbanken angewendet werden, um eine schnelle und zuverlässige Aussage über potentielle neue Kandidat-Strukturen zu treffen [146]. Es können zwei grundsätzlich verschiedene Arten von ähnlichkeitsbasierten Methoden beschrieben werden:

Atombasiert: Hierbei gibt es wieder eine Unterteilung und zwar in 2D- und 3D-Methoden. Bei den 2D-Methoden werden die Moleküle als Graph dargestellt und es wird versucht, die Atome aus Molekül 1 in Molekül 2 wiederzufinden [147]. Bei den 3D-Methoden versucht man eine Superpositionierung des gesamten Moleküls. Dabei kann man sich entweder nur auf die Atome beschränken [148], oder sich auf die Oberfläche bzw. das Volumen beziehen [149–151]. Bei allen Methoden wird die Ähnlichkeit aus den sich überlappenden Bereichen berechnet. Nach dem Schlüssel-Schloss-Prinzip ist die Form des Liganden wichtig, damit er vom Protein erkannt wird. Somit müssen potentielle Inhibitoren auch einen gewissen Grad an Ähnlichkeit mit dem regulären Substrat haben. Es gibt eine Reihe von mathematischen Methoden um die 3D Ähnlichkeit von zwei Molekülen zu bestimmen [152]. Diese sind zum Teil sehr komplex und rechenintensiv, deswegen wurde eine Gauß'sche-Beschreibung der molekularen Form entwickelt, die auf einem Koaleszenz-Theorem beruht, um molekulare Volumina und Bereiche zu berechnen [153]. Nachdem die Sphären bestimmt wurden, werden die sich überlappenden Bereiche maximiert. Dazu werden nur die Schweratome berücksichtigt. Es ist möglich, nicht nur die Form zu betrachten, sondern auch noch chemische Eigenschaften bei der Superpositionierung mit zu berücksichtigen.

Deskriptorbasiert: Bei dieser Methode werden die Moleküle durch eine Reihe von ein- oder mehrdimensionalen Deskriptoren näher beschrieben, z. B. Molekulargewicht, Anzahl der Ringe, Wasserstoffbrückenakzeptoren und Ladung. Der Vorteil von Deskriptoren ist, dass sie aufgrund der molekularen Struktur berechenbar sind [154–156]. Ebenso können Substrukturen als Deskriptor verwendet werden. Dabei werden diese entweder als *fingerprints*, *Pharmacophormodelle* oder SMARTS gespeichert.

Ein fingerprint stellt die molekulare Struktur und chemische Eigenschaften, die durch die Deskriptoren beschrieben werden, typischerweise als einen binären String dar. Dies bedeutet, dass man in einem eindimensionalen String eine Folge von 1 und 0 speichert, wobei der Wert 1 bedeutet, dass eine bestimmte Eigenschaft in dem Molekül vorhanden ist. Jede Stelle in diesem String wird als Bit bezeichnet. Somit ergibt sich für jedes denkbare Molekül eine charakteristische Bitfolge. Damit ein Vergleich verschiedener fingerprints zuverlässig möglich ist, müssen diese die gleiche Länge haben. Eine der ersten Anwendungen für den Vergleich von einem fingerprint mit einer Datenbank entstand 1996 unter dem Namen Stigmata [157]. Bei diesen Vergleichen wird aus den Gemeinsamkeiten

der Strings ein Koeffizient, z. B. Tanimoto [158] gebildet, der ein Maß für die Ähnlichkeit der Strukturen darstellt.

$$S_{A,B} = \frac{c}{[a + b - c]} \quad (2.13)$$

$S_{A,B}$ ist die Ähnlichkeit vom Molekül A zu B , c ist die Anzahl der gemeinsamen Bits mit dem Wert 1, a und b die jeweilige Anzahl der Bits mit dem Wert 1 [159]. Da die Speicherung als String erfolgt, sind fingerprints für Vergleichsoperationen sehr effizient und eignen sich besonders für das schnelle Screening einer Datenbank [160, 161].

Das Programm ROCS (*Rapid Overlay of Chemical Structures*) verwendet die Methode der Superpositionierung der 3D-Struktur. ROCS maximiert dabei nicht nur die sich überlappenden Bereiche der Moleküle, sondern bezieht auch noch die chemischen Eigenschaften der Moleküle mit in die Bewertung für die Ähnlichkeit mit ein [162].

2.5. Moleküldynamik-Simulation

Bei einer Moleküldynamik-Simulation (MD) handelt es sich um eine Computersimulation der Bewegungen und Wechselwirkungen von Atomen und Molekülen innerhalb eines festgelegten Systems. Die MD wird für eine bestimmte Zeit durchgeführt. Diese Zeitspanne variiert stark in der Abhängigkeit von der Komplexität des zu simulierenden Systems, befindet sich aber meist im Nanosekundenbereich. Dies geschieht unter Zuhilfenahme des *Newtonschen Gesetzes* (siehe Formel 2.14; F_i ist die Kraft, die auf das Atom i wirkt, m_i dessen Masse, a_i die Beschleunigung des Atoms i). So erhält man eine Anzahl von Konformationen, die als Funktion der Zeit durch eine *Trajektorie* beschrieben werden.

$$F_i = m_i a_i \quad (2.14)$$

Diese rein mechanische Betrachtung erlaubt es Simulationen von großen Systemen mit mehr als 100.000 Atomen durchzuführen. Deswegen ist dies eine gute Methode um die Interaktionen von einem Liganden mit einem Protein zu bestimmen. Da es sich hierbei um ein bewegliches System handelt, ist es möglich die freie Bindungsenergie des Komplexes zu bestimmen [163]. Hierzu bedarf es eines Kraftfeldes, das in einer allgemeinen Form in (2.15) beschrieben ist [123].

$$E = \sum_{\text{Bindung}} K_b (b - b_0)^2 + \sum_{\text{Winkel}} K_\theta (\theta - \theta_0)^2 + \sum_{\text{Tor}} K_\phi [1 + \cos(n\phi - \delta)] \\ + \sum_{\text{vdW}} \varepsilon_{ij} \left[\left(\frac{R_{\text{min},ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{\text{min},ij}}{r_{ij}} \right)^6 \right] + \sum_{\text{ele}} \frac{q_i q_j}{4\pi \varepsilon_0 r_{ij}} \quad (2.15)$$

Die gebundenen Interaktionen schließen die Bindungs-, Winkel- und Drehterme (b , ϕ und ε) mit ein. Die nichtgebundenen Interaktionen umfassen den van-der-Waals-Energie-Term, dargestellt durch das Lennard-Jones Potential, und die elektrostatischen Wechselwirkungen, die als Coulomb-Wechselwirkungen mit einer Punktladung

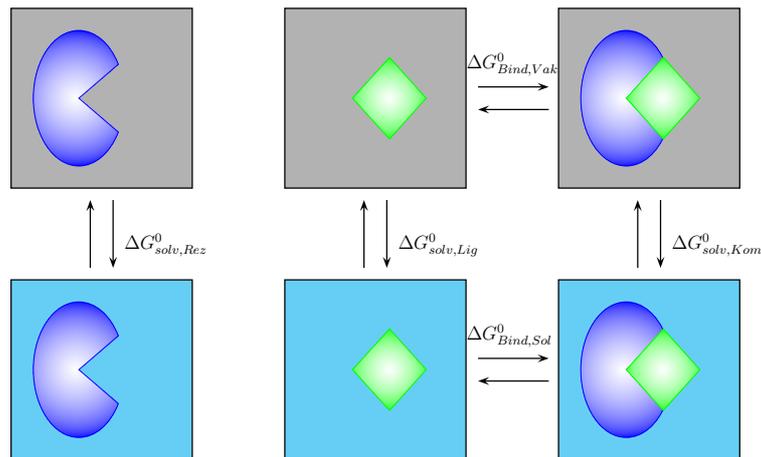


Abbildung 2.2: Freie Bindungsenergie

Die Berechnung der freien Bindungsenergie wird in mehrere Schritte unterteilt. Gezeigt ist ein thermodynamischer Kreislauf, in dem getrennt die ΔG^0 für den Liganden, den Rezeptor und dem Komplex vom Vakuum ins Solvent gerechnet wird. Dabei ist $\Delta G^0_{Bind,Sol} = \Delta G^0_{Bind,Vak} + \Delta G^0_{Solv,Kom} - (\Delta G^0_{Solv,Lig} + \Delta G^0_{Solv,Rez})$ grau = Vakuum, cyan = Solvent

im Zentrum des Atoms, behandelt werden. Wasser kann in der gleichen Art und Weise behandelt werden, wie das Protein. Dabei haben sich eine Reihe von verschiedenen Modellen mit explizitem Solvent entwickelt, um die absolute oder relative freie Bindungsenergie von einem Protein-Ligand-Komplex zu bestimmen [164]. Hierbei muss jedoch beachtet werden, dass das *sampling* von Konformationen sehr viel Rechenzeit in Anspruch nimmt. Deswegen haben sich Modelle mit impliziten Solvent etabliert, die das Wasser als dielektrisches Kontinuum behandeln. Somit kann die benötigte Rechenzeit, die man auf die Berechnung der Wechselwirkungen mit dem Solvent verwenden muss, reduziert werden. Die verbreitetsten Modelle hierfür sind die *Poisson-Boltzmann/Surface Area* (PB/SA)[165] und *Generalized Born/Surface Area* (GB/SA)[166, 167]. GB-Modelle können als halb-analytische Annäherung an die PB-Gleichung angesehen werden. Aus diesen Modellen hat sich die *Molecular Mechanics Poisson-Boltzmann/Surface Area* MM-PB/SA entwickelt. Sie wird verwendet, um die freie Bindungsenergie des Liganden zu bestimmen und zeigt eine hohe Übereinstimmung mit experimentellen Daten [168–171].

2.5.1 Molecular Mechanics Poisson-Boltzmann/surface area (MM-PB/SA)

MM-PB/SA dient zur Bestimmung der freien Bindungsenergie (ΔG_b^0). Sie ist definiert als die Differenz aus der freien Bindungsenthalpie (ΔH_b^0), dem Produkt der Temperatur (T) und der Bindungsentropie (ΔS_b^0) wie in (2.16) beschrieben ist[172].

$$\Delta G_b^0 = \Delta H_b^0 - T \Delta S_b^0 \quad (2.16)$$

Die Berechnung von ΔG_b^0 erfolgt anhand der Trajektorien des Systems, nach einer MD mit explizitem Solvent. Aus den Trajektorien wird ein Ensemble von Schnappschüssen erstellt, die jeweils den Zustand des Systems nach zu einem bestimmten Zeitpunkt darstellen. Diese Schnappschüsse dienen dazu, die gemittelte freie Bindungsenergie ($\langle G \rangle$) nach (2.17) zu berechnen [173, 174].

$$\langle G \rangle = \langle E_{MM} \rangle + \langle G_{PB SA} \rangle - TS_{MM} \quad (2.17)$$

$\langle E_{MM} \rangle$ beschreibt die Wechselwirkungen der Atome des Solvents untereinander, und $\langle G_{PB SA} \rangle$ die freien Solvatationsenergien. Die Werte für $\langle E_{MM} \rangle$ (2.18) berechnen sich analog wie in (2.1):

$$\langle E_{MM} \rangle = \langle E_{Bindung} \rangle + \langle E_{Winkel} \rangle + \langle E_{Tor} \rangle + \langle E_{vdW} \rangle + \langle E_{Ele} \rangle \quad (2.18)$$

Wie schon erwähnt, wird bei MM-PB/SA-Modellen das Solvent als dielektrisches Kontinuum betrachtet und das Solut als eine Ansammlung von Punktladungen. Dadurch lässt sich die freie Enthalpie ($\langle G_{PB SA} \rangle$) abschätzen. Diese setzt sich aus einem apolaren und einem polaren Beitrag zusammen. Der polare Beitrag wird durch die numerische Lösung der Poisson-Boltzmann-Gleichung berechnet (2.19) [156]. Sie beschreibt den Zusammenhang zwischen dem elektrostatischen Potential $\phi(r)$ innerhalb eines homogenen Mediums mit der Dielektrizitätskonstante ϵ und der Ladungsdichte $p(r)$.

$$\nabla^2 \phi(r) = -\frac{4\pi p(r)}{\epsilon} \quad (2.19)$$

Durch Berücksichtigung des Oberflächeninhalts der Bindetasche wird der apolare Beitrag $\langle G_{apolar} \rangle$ ausgedrückt [175]. Dieser wird aus der für das Lösungsmittel zugänglichen Oberfläche (SASA) von Protein, Ligand und Komplex berechnet (2.20) [176].

$$G_{apolar} = \gamma SASA + b \quad (2.20)$$

2.5.2 Molecular Mechanics Generalized-Born/surface area (MM-GB/SA)

Das Generalized Born/Surface Area (GB/SA)-Modell ist eine Näherungsgleichung der genauen Poisson-Boltzmann-Gleichung. Dabei wird das Protein als eine Sammlung von Sphären betrachtet, die sich durch ihre Dielektrizitätskonstanzen von Solvent unterscheiden. Dieses Modell wird durch die Formel (2.21) beschrieben:

$$G_s = \frac{1}{8\pi} \left(\frac{1}{\epsilon_0} - \frac{1}{\epsilon} \right) \sum_{i,j}^N f_{i,j} \frac{q_i q_j}{f_{GB}} \quad (2.21)$$

wobei

$$f_{GB} = \sqrt{r_{ij}^2 + a_{ij}^2} e^{-D} \quad (2.22)$$

und

$$D = \left(\frac{r_{ij}}{2a_{ij}} \right)^2, a_{ij} = \sqrt{a_i a_j} \quad (2.23)$$

ϵ_0 ist die elektrische Feldkonstante, ϵ ist die Dielektrizitätskonstante des Solvents. q_i ist die elektrische Ladung des Teilchens i , r_{ij} ist der Abstand zwischen den Teilchen i und j . a_i ist der Born-Radius [177, 178].

Zur Verbesserung wird dieses Modell noch durch den Term für die SASA erweitert. GB/SA ist einer der am weitesten verbreiteten Methoden für die Berechnungen von MDs mit impliziten Solvent. Für kurze Peptide konnte dieses Modell schon erfolgreich die native Struktur von kurzen Peptiden bestimmen [179], was für Modelle mit expliziten Solvent nicht funktioniert hat [180].

2.5.3 Linear Response Molecular Mechanics Poisson-Boltzmann/Surface Area

LR-MMPBSA ist eine Variante der normalen MM-PB/SA Berechnung. Dabei werden im Solvationsmodell verschiedene Dielektrizitätskonstanten für das Lösungsmittel und das Solut verwendet. Der elektrostatische Effekt, der bei einer gegebenen Ladung auftritt, kann durch die Poisson-Gleichung beschrieben werden

$$\nabla \cdot \epsilon(r) \nabla \phi(r) + \rho = 0 \quad (2.24)$$

Wobei r den Positionsvektor darstellt, ρ die Ladungsdichte ϕ das elektrostatische Potential und ϵ die Dielektrizitätskonstante. Die linearisierte Form von 2.24 lässt sich in der Form

$$\nabla \cdot \epsilon(r) \nabla \phi(r) = -\rho(r) - k_0^2 \phi(r) \quad (2.25)$$

aufstellen, wobei k_0 eine Konstante darstellt, und wie folgt definiert ist

$$k_0^2 = 2e^2 N_A \rho_0 I / kT \quad (2.26)$$

Die Gleichung 2.25 wird durch die Finite-Differenzen-Methode gelöst [181–183].

2.6. Gütemaße

Um die Ergebnisse eines Dockings zu bewerten kann man verschiedene Gütemaße betrachten. Die einfachste Betrachtung ist hierbei die reine Rangfolge der Ergebnisse anhand der Scores. Da diese Scores jedoch, abhängig von der SF, in vielen Fällen auf Schätzungen beruhen, ist es nicht zu empfehlen, sich nur auf die Rangfolge der Ergebnisse zu verlassen [184]. Als Bewertungsmaße für VS werden hier die *Receiver Operating Characteristic Kurve*, die *Area Under the Curve*, *Robust Initial Enhancement* und der *Enrichment Factor* weiter verwendet.

2.6.1 Receiver Operating Characteristic Kurve

Die Receiver Operating Characteristic Kurve (ROC-Kurve) wird benutzt, um Analyse-Strategien zu optimieren. Dabei wird die Sensitivität (die Rate der *true positives* oder *richtig Positiven*) gegen die *false negatives* (*falsch negativen* oder auch $1 - \text{Spezifität}$)

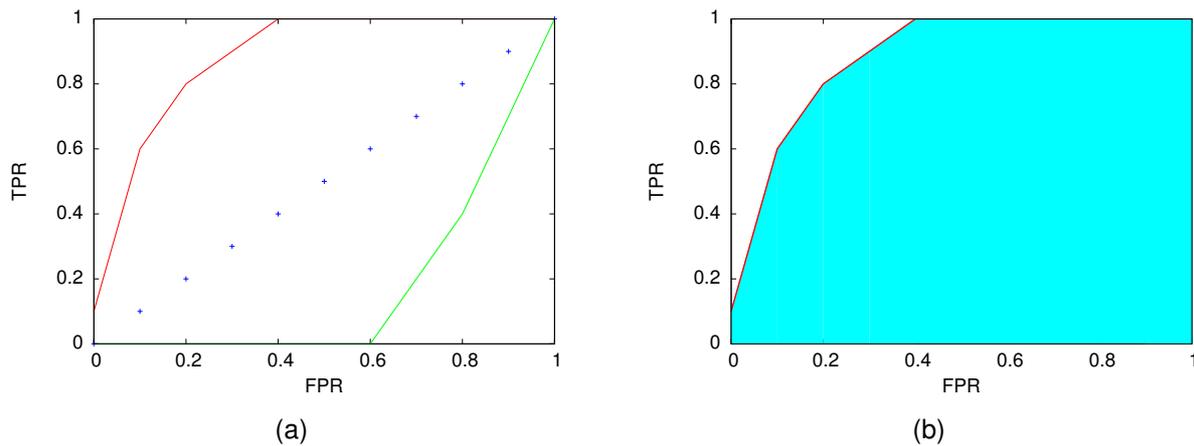


Abbildung 2.3: ROC-Kurve und AUC

a) Zwei beispielhafte Verläufe von ROC-Kurven. Die rote Linie zeigt einen guten Verlauf, da die Trefferquote schnell ansteigt. Bei einem optimalen Verlauf würde die Kurve direkt zu Beginn senkrecht ansteigen. Die grüne Linie zeigt einen schlechten Verlauf, da zuerst viele falsch Positive gefunden werden, bevor es zu echten Treffern kommt. Die blauen Kreuze deuten den Verlauf an, wenn die Ergebnisse zufällig verteilt wären. b) AUC der blau markierte Bereich wird berechnet und beschreibt die Fläche unter der Kurve.

aufgetragen. Die ROC-Kurve kann somit gleichbedeutend dargestellt werden, indem man den Anteil der wahr Positiven aus der Menge aller Positiven, gegen den Anteil der falsch Negativen aus der Menge aller negativen aufträgt. Es gilt somit für die richtig Positiven (2.27), und die richtig Negativen (2.28):

$$TPR = \frac{TP}{P} = \frac{TP}{TP + FN} \quad (2.27)$$

$$FPR = \frac{FP}{N} = \frac{FP}{FP + TN} \quad (2.28)$$

Abbildung 2.3(a) gibt einige Beispiele zum besseren Verständnis.

2.6.2 Area Under Curve

Als weiteres Merkmal einer ROC-Kurve wird ebenfalls die *area under curve* (AUC) bestimmt. Man erhält einen Wert für die Fläche, und dieser Wert kann zwischen 0 und 1 liegen. 1 stellt dabei den besten Wert dar, da hierbei alle richtig positiven direkt am Anfang gefunden wurden. Ein Ergebnis von 0.5 steht für den schlechtesten Wert, da hier eine zufällige Verteilung der Ergebnisse vorliegt und man deswegen keine Aussage über die Messung machen kann. Ist der Wert jedoch kleiner als 0.5 und nähert sich 0 an, so kann man eine negative Aussage über den Test machen. Formel 2.29 [185] beschreibt, wie die Fläche berechnet wird

$$G_1 = 1 - \sum_{k=1}^n (X_k - X_{k-1})(Y_k + Y_{k-1}) \quad (2.29)$$

2.6.3 Robust Initial Enhancement

Um die frühe Anreicherung gut bestimmen zu können gibt es die Methode des *Robust Initial Enhancement* (RIE), eine Weiterentwicklung des *initial enhancement* (IE). Das IE beschreibt das Verhältnis der beobachteten Nummern aktiver Liganden an einer bestimmten Stelle in der Liste zu der Anzahl der erwarteten aktiven Liganden, wenn diese zufällig über die Liste verteilt wären. Diese Methode funktioniert jedoch nur bei einer großen Anzahl von untersuchten Substanzen (> 300), weil sonst eine Verschiebung von ein oder zwei aktiven Liganden innerhalb der Liste einen zu großen Einfluß auf des Ergebnis hat. Deshalb wird für das RIE zuerst eine Glättungsfunktion auf die geordnete Liste angewand:

$$S = \bullet_i^{\text{actives}} \exp(-\text{rank}(i)/a) \quad (2.30)$$

Wobei gilt $a \ll N$. Daraus ergibt sich für das RIE

$$RIE = S / \langle S \rangle \quad (2.31)$$

$\langle S \rangle$ steht dabei für den Durchschnitt von S aus 1000 versuchen, bei denen die Ränge der aktiven Liganden zufällig über die Liste verteilt waren. [186]

2.6.4 Enrichment Factor

Der EF spiegelt die Fähigkeit eines VS wieder aus allen getesteten Substanzen die *richtig positiven* zu finden, und besser zu bewerten als die *wahr negativen*. Dazu wird der Quotient aus der Konzentration der aktiven Liganden aus der betrachteten Teilmenge und der Konzentration der aktiven Liganden in der gesamten Datenbank gebildet:

$$EF_x = \frac{\frac{L_x}{N_x}}{\frac{L_{tot}}{N_{tot}}} \quad (2.32)$$

L_x ist die Anzahl an aktiven Liganden, die innerhalb der ersten $x\%$ der Ergebnisdatenbank gefunden wurden, und N_x die Anzahl aller Liganden innerhalb der ersten $x\%$. L_{tot} und N_{tot} ist die Anzahl aller aktiven Liganden bzw. aller Liganden in der Datenbank [187]. Wichtig hierbei ist, dass dieser Wert immer nur für einen Teil der Datenbank, und sich somit an jeder Stelle ändert.

2.7. Erstellen der Substanzdatenbanken

Um die Leistung der einzelnen Scoring-Funktionen bewerten zu können, wird eine hinreichend große Menge an sogenannten *Decoys* benötigt. Diese sollen den aktiven Substanzen ähnlich sein, damit es für die verwendete SF nicht zu leicht ist, zwischen aktiven von inaktiven Substanzen zu differenzieren. Die Decoys sind aus dem *Directory of Useful Decoys* (DUD) [184] entnommen. Dazu wurden verschiedene Eigenschaften der aktiven Substanzen bestimmt (TPSA, logP, Anzahl der Donatoren,

Anzahl der Akzeptoren, Gewicht, Anzahl der Ringe, Anzahl der Schweratome, Anzahl von Halogenen). Um aus dem DUD jeweils passende Substanzen zu filtern, wurden die gefundenen Werte angepasst. Die Ober- und Untergrenzen der Werte für TPSA, logP, Gewicht wurden jeweils um 5% nach oben und unten erweitert. Die Anzahl der Ringe, Halogene, Akzeptoren und Donatoren musste gleich sein, und die Anzahl der Schweratome insgesamt durfte um 2 größer oder kleiner sein. Anschließend wurden die Strukturen aus dem DUD verworfen, die nicht innerhalb dieser Grenzen lagen und aus den verbliebenen Strukturen wurden zufällig neunmal mehr Decoys ausgewählt, als aktive Substanzen vorhanden waren. Für das Screening mit ROCS wurden von jeder Substanz 250 zusätzliche Konformationen erzeugt.

2.8. Vorbereiten der Proteinstrukturen für das Docking

2.8.1 Virtuelles Screening

Soweit möglich wurden nur Kristallstrukturen ausgewählt, deren Auflösung weniger als 3 Å betrug. In Kapitel 1.2.3.1 wurde gezeigt, dass Kinasen zwei Konformationen einnehmen können. Darum wurden ebenfalls beide Konformationen für das VS verwendet, wenn es denn für beide Konformationen eine Kristallstruktur gab. Sofern nicht anders angegeben, wurden alle Proteinstrukturen nach dem gleichen Protokoll mit dem Programmpaket MOE2010.10 [188] vorbereitet. Dabei wurden die fehlenden Protonen addiert und darauf folgend die Bindetasche mehrfach einer Energieminimierung unterzogen, wobei in jedem Schritt eine schwächere Kraft gelegt wurde, um die Schweratome an ihren Positionen zu halten. Dabei wurde das MMFF94-Kraftfeld und *conjugate gradient* als Minimierungsalgorithmus so lange verwendet, bis ein Gradient von $0,01 \text{ kcal}\cdot\text{mol}^{-1}$ erreicht wurde. Soweit vorhanden wurden Wassermoleküle im Umkreis von 5 Å um die Bindetasche nicht entfernt und mit ins Docking einbezogen. Insgesamt wurden die Dockings einmal mit und einmal ohne Wasser gemacht.

2.8.2 Moleküldynamik-Simulation

Alle in MDs benutzten Proteine wurden nach dem gleichen Protokoll mittels *LEaP* aus dem AMBER-Programmpaket vorbereitet. Dabei wurde als Kraftfeld FF99 [189] verwendet. Die detaillierte Auflistung der Skripte ist im Anhang Kapitel 5.2.1 zu finden.

2.9. Verwendete Programme

Folgende Dockingprogramme und Scoring-Funktionen wurden für das VS verwendet:

- GOLD [108]
 - GOLDScore
 - CHEMScore

- GLIDE [110]
 - GLIDEScore SP
- PARADOCKS [190]
 - PScore
 - PMF
- ROCS [162]
 - TANIMOTOCOMBOscore
- DOCK [191]
 - AMBERscore [192]

Für die Moleküldynamik-Simulationen wurden AMBER [193] und AmberTools verwendet.

2.10. Kinase Assays

2.10.1 PRK1 Assay

Für die biologische Testung an der PRK1 wurde ein LanthaScreen™Eu Kinase Binding Assay Kit von Invitrogen verwendet. Die Assaykonzentrationen waren 5 nM für die PRK1, 2 nM LanthaScreen Eu-anti-GST Antikörper und 10 nM Kinase Tracer 236. Alle getesteten Substanzen wurden bei einer Konzentration von 100 nM getestet.

2.10.2 MYT1-Assay: LanthaScreen Assay mittels FRET

Die Aktivitätsuntersuchungen an der MYT1-Kinase wurden mit einem LanthaScreen™Eu Kinase Binding Assay Kit von Invitrogen durchgeführt. Für die Messung wird die Technik des *time-resolved fluorescence energy transfer* (TR-FRET) genutzt. Allgemein wird dabei die Energie eines angeregten Farbstoffes (Donator) strahlungsfrei auf einen Akzeptor übertragen. Wichtig ist, dass hierbei kein Transfer von Elektronen stattfindet. Dies geschieht über Dipol-Dipol-Wechselwirkungen und beruht auf dem Coulombschen Gesetz. Die Reichweite dieser Übertragung ist jedoch beschränkt. Der Abstand zwischen Donator und Akzeptor muss weniger als 10 nm betragen [194, 195]. Mit der hier verwendeten Methode ist man in der Lage den zeitlichen Verlauf des Energietransfers zu beobachten. Wird der Tracer von einem aktiven Liganden aufgrund der höheren Affinität des Liganden aus der Bindetasche verdrängt, sinkt die Signalintensität im Laufe der Zeit [196].

Als Chromophor am Antikörper wird hier Europium verwendet, und der Akzeptorfarbstoff am Tracer ist AlexaFluor®647.

In Abbildung 2.4 ist die Funktionsweise des MYT1-Bindungsassay dargestellt

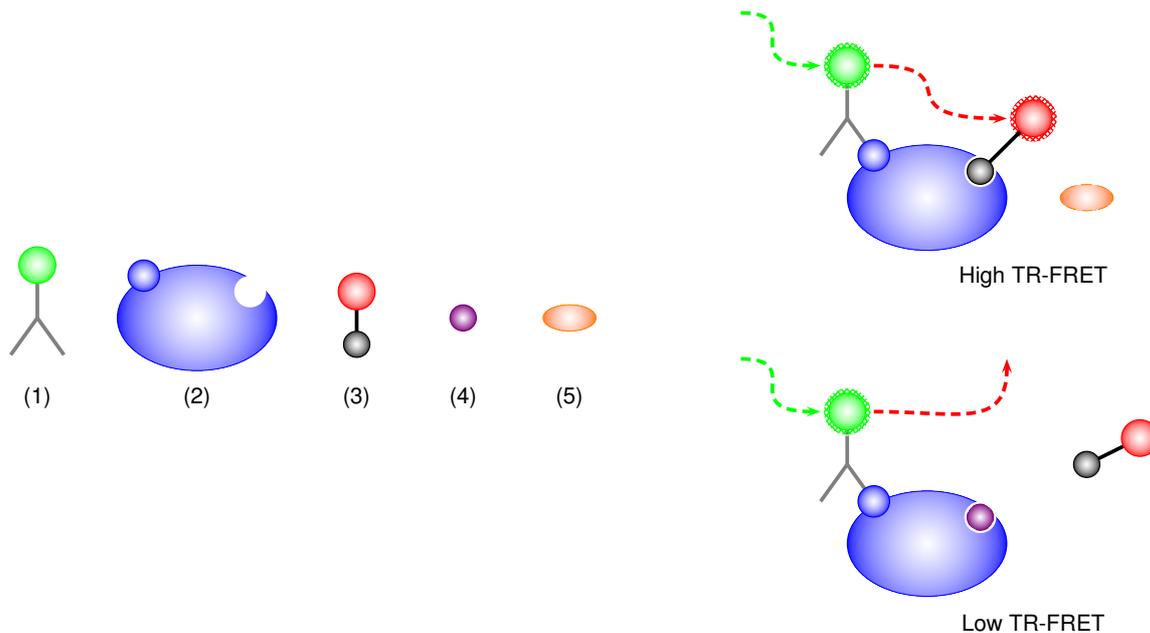


Abbildung 2.4: TR-FRET

Dargestellt sind die Funktionsweise eines TR-FRET, mit High und Low FRET.

1) Eu-markierter Antikörper 2) Protein mit Epitop, das vom Antikörper erkannt wird 3) Tracer 4) Inhibitor 5) Nicht-Binder

2.10.3 IGFR, EGFR und VEGFR Assay

Bei der biologischen Testung von VEGFR, EGFR und IGFR wurde ein Affinitäts Assay durchgeführt. Hierbei wurde überprüft, wie sich die Phosphorylierungsrate der Kinasen in Anwesenheit der zu testenden Substanzen änderte. Dazu wurde die emittierte Strahlung von radioaktivem Phosphat ($^{33}\text{P}_i$) gemessen, das auf das Substrat übertragen wurde. Aus der verbliebenen Aktivität des Enzyms wurden die IC_{50} -Werte ermittelt, aus denen im Anschluß die K_i -Werte wie folgt berechnet wurden [197]:

$$\text{IC}_{50} = [E_t]/2 + K_i x (1 + [S]/K_M) \quad (2.33)$$

3. ERGEBNISSE UND DISKUSSION

3.1. Dockingstudien an Proteinkinasen

3.1.1 Betrachtung der einzelnen Programme und Scoring-Funktion

Anhand von 19 Datensätzen aus der Literatur (siehe Anhang) wurde eine retrospektive Analyse durchgeführt, welche Methode sich am besten eignet, um die bekannten aktiven Liganden von den ebenfalls bekannten inaktiven Liganden zu unterscheiden. Diese Datensätze haben sich besonders für diese Untersuchungen geeignet, da für alle der getesteten Kinasen Kristallstrukturen mit Inhibitoren verfügbar waren. Ebenso lagen gut dokumentierte Struktur-Aktivitäts-Beziehungen (SAR, structure-activity-relationship) vor. Zudem sprach die Veröffentlichung in Zeitschriften wie z. B. *J Med Chem*, mit einem *impact factor* von 5,6 für die Qualität der durchgeführten Untersuchungen. Bei den 19 Datensätzen handelte es sich um Aktivitätsmessungen von 660 Liganden an unterschiedlichen Proteinkinasen. Um eine ausreichendgroße Menge an Decoys zu haben, wurde die jeweils neunfache Menge von aktiven Liganden aus der DUD entnommen.

Sowohl die aktiven Liganden als auch die Decoys wurden mit verschiedenen Dockingprogrammen und SF in die jeweiligen Proteine gedockt. Nicht immer war anhand der ausgewählten Studien zu erkennen, ob die Kinase in der aktiven, oder in der inaktiven Konformation vorlag. Deshalb wurde – wenn möglich – in beide Konformationen gedockt. Zudem wurde analysiert, wie sich eventuell vorhandenes Kristallwasser auf die Dockings auswirkt. Auch hier wurden die Dockings jeweils zweimal durchgeführt; einmal mit Wasser und einmal ohne Wasser im Umkreis von 5 Å um die ATP-Bindetasche.

Insgesamt wurden mit jeder SF 70 Dockingstudien an 40 Kristallstrukturen von 19 Kinasen durchgeführt. Von den daraus entstanden 420 Ergebnisdatenbanken, wurden 280 nochmals einer Neubewertung durch den CHEMSCORE und GLIDESCORE unterzogen, um zu prüfen, ob ein Rescoring der Ergebnisse die Unterscheidung zwischen aktiven und inaktiven Liganden verbessert.

3.1.2 Generierung der Datenbanken

Die Datenbank für die Dockingstudien wurde wie in Kapitel 2.7 beschrieben erzeugt. Dabei wurden für jeden Datensatz die physicochemischen Eigenschaften der aktiven Substanzen als Grenzwerte genommen, um aus der DUD ähnliche inaktive Substanzen zu filtern. Es entstand für jeden Datensatz eine neue Teildatenbank mit Substanzen aus der DUD. Aus den jeweiligen Unterdatenbanken wurde die neunfache Menge an inaktiven Substanzen zufällig ausgewählt, als aktive Substanzen vorhanden waren.

In Tabelle 3.1 sind für alle Datensätze die verwendeten Kristallstrukturen aufgelistet, ebenso die aktivitätsbereiche der Inhibitoren, und die Anzahl der aktiven und inaktiven Verbindungen.

Für die Untersuchungen mit ROCS wurde unter Zuhilfenahme des Programms *OMEGA* [162] von jeder Struktur 250 zusätzliche Konformationen erzeugt.

Tabelle 3.1: Retrospektives VS Aufgeführt sind alle Veröffentlichungen, aus denen die im VS gestesteten Verbindungen entnommen sind. Die PDB-Codes sind danach sortiert, ob das DFG-Motiv in einer *in*- oder *out*-Konformation vorliegt

Kinase	Quelle	#Aktive	Aktivität [μM]	#Decoys	PDB-Codes	
					DFG-in	DFG-out
<i>ABL</i>	[198]	50	$3,8 \cdot 10^{-3} - 10$	450	2GQG	1IEP
					3KF4	2HIW
<i>Aurora</i>	[199]	6	$1,0 \cdot 10^{-3} - 0,05$	54	2W1I	1MUO
	[200]	32	$0,033 - >50$	288	2F4J	2J5O
					2WI1	1MUO
[201]	13	$0,0014 - 5,5$	117	2F4J	2J5O	
					2WI1	1MUO
					2F4J	2J5O
<i>CDC7</i>	[202]	47	$0,001 - 0,064$	423	3DU8	
<i>CDK2</i>	[203]	27	$2,0 \cdot 10^{-3} - >10$	243	2FVD	1KE6
	[204]	6	$0,06 - >100$	54	1P2A	1DI8
1E9H						
					1FQ1	
<i>CDK4</i>	[205]	76	$2,0 \cdot 10^{-3} - >50$	684	2F2C	1BI8
	[204]	6	$0,06 - >100$	54	1XO2	1G3N
1XO2					1BI7	
					1JOW	1G3N
<i>CDK5</i>	[204]	6	$0,06 - >100$	54	1UNL	
					1UNG	
<i>CDK7</i>	[206]	40	$3,0 \cdot 10^{-3} - >10$	360	1UA2	
	[204]	6	$0,06 - >100$	54	1UA2	
<i>CDK9</i>	[204]	6	$0,06 - >100$	54	3BLR	
					3MIA	
<i>CHK1</i>	[207]	15	$0,29 - 74$	135	2WMV	
					2XEZ	
<i>CK18</i>	[208]	20	$4,0 \cdot 10^{-3} - 2,52$	180	1EH4	

Fortsetzung nächste Seite ...

Kinase	Quelle	#Aktive	Aktivität [μM]	#Decoys	DFG-in	DFG-out
					2IZR	
<i>C-JUN</i>	[209]	39	0,4–84	351	3G90 2GMX	
<i>GSK</i>	[210]	40	$3,5 \cdot 10^{-4}$ –4,09	360	3F88 3L1S	
<i>MET</i>	[211]	10	$1,0 \cdot 10^{-3}$ – $5,9 \cdot 10^{-3}$	90	2WGJ 2WD1	3F82 3CTH
<i>P38</i>	[212]	58	0,44–4,71	522	3GC9 3FLQ	2GTM 2BAQ
	[213]	19	0,04–2,73	171	3GC9 3FLQ	2GTM 2BAQ
	[208]	20	$4,0 \cdot 10^{-2}$ –25	180	3GC9 3FLQ	2GTM 2BAQ
<i>PIM</i>	[214]	27	0,01–>10	243	2J2I 3BWF	
	[215]	49	0,024–>100	441	2J2I 3BWF	
<i>VGFR</i>	[216]	42	$5,0 \cdot 10^{-3}$ –>25	378	3B8R 2P2H	3EWH 3CPC
Σ		660		5940		

3.1.3 Prüfen der Dockingeinstellungen

Für GOLD (GOLDScore), GLIDE (GLIDEScore), PARADOCKS (PMF und PScore) wurden die jeweiligen Standardeinstellungen übernommen und wurden an jeder für die Dockingstudien verwendeten Kristallstrukturen getestet. Bei GLIDE wurde nur die *standard precision* verwendet. Als Zielregion diente ein Bereich von 15 Å um den Gatekeeper. Damit wurde getestet, ob die Programme mit den jeweiligen SF in der Lage waren die Position des co-kristallisierten Liganden wiederzufinden. Abbildung 3.1 zeigt wie die Genauigkeit der einzelnen Programme war. Alle RMSD-Werte lagen unter 2,5 Å, somit galt die Kristallstruktur des Liganden als gefunden, und mit den erfolgreichen Einstellungen wurden die nachfolgenden Untersuchungen durchgeführt.

3.1.4 Fingerprintanalyse von aktiven und inaktiven Liganden

Es wurde ebenfalls untersucht, wie sehr sich die inaktiven von den aktiven Liganden unterscheiden. Dazu wurde mittels MOE eine Fingerprintanalyse durchgeführt. Hierbei wurden die aktiven Liganden in *Cluster* eingeteilt. Diese Cluster wurden dann mit den inaktiven Liganden verglichen, und betrachtet, wieviele Decoys den aktiven zu 60 und

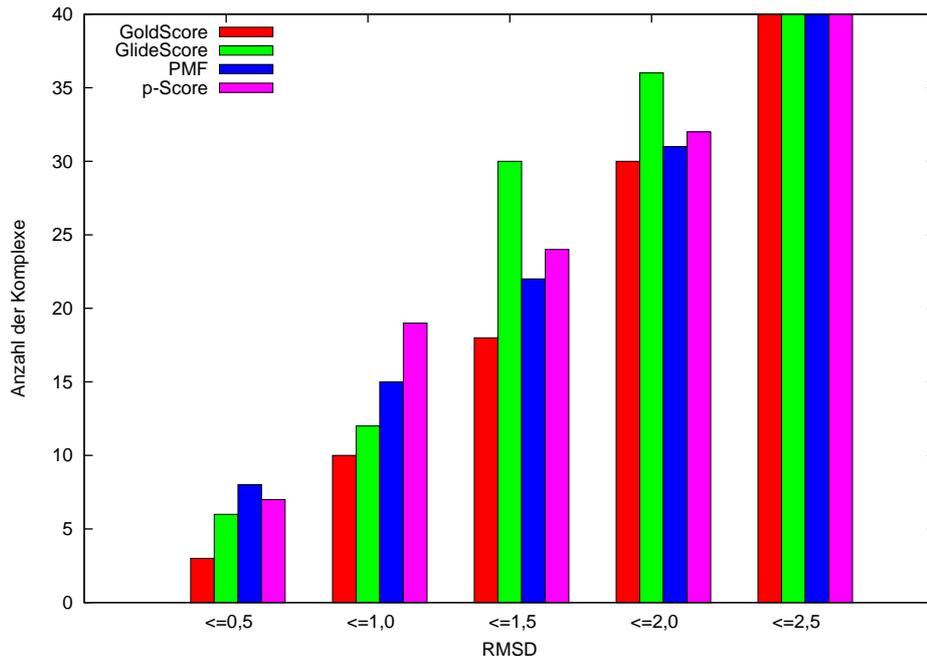


Abbildung 3.1: Performance des Dockings mit den verwendeten Dockingprogrammen und Scoring-Funktionen. Getestet wurde, ob die Programme in der Lage waren die Bindemodus der kristallisierten Liganden wiederzufinden. Getestet wurden alle 40 verwendeten Kristallstrukturen.

70% übereinstimmen. Die Ergebnisse hiervon sind in Tabelle 3.2 dargestellt. Für die inaktiven Verbindungen, die für die Datensätzen von Pfeifer, Qian, und Xia gab es keine Substanzen die mindestens eine 60%ige Ähnlichkeit mit den aktiven Liganden aufwiesen. Und nur für sieben Datensätze wurden inaktive Liganden gefunden, die eine Ähnlichkeit von mindestens 70% mit den aktiven Verbindungen haben. Die ausgewählten Substanzen entsprechen den aktiven Liganden somit in den physico-chemischen Eigenschaften, sind aber Strukturell doch verschieden.

3.1.5 Auswertung der AUC-Werte

Aus den gewonnenen Daten wurden nun die Mittelwerte samt Standardabweichung für die verwendete Methode bei der jeweils erhaltenen AUC ermittelt (s. Tabelle 3.3). Die Ergebnisse für die einzelnen Kinasen sind im Anhang 5.1 aufgeführt. Als Grundlage dafür dienten die nach dem Bewertungsmaß sortierten Listen der getesteten Verbindungen. Es wurde wie in den Kapiteln 2.6.1 bis 2.6.2 beschrieben zuerst die ROC-Kurve-Kurve aufgetragen, und danach die Fläche unter der Kurve ermittelt.

Es zeigt sich, dass bei diesem Screening das Programm Rocs am besten zwischen aktiven und inaktiven Liganden unterscheiden kann. Dies zeigt sich daran, dass man für Rocs die größte AUC erhält.

Tabelle 3.2: Fingerprintanalyse

der aktiven und inaktiven Liganden. Angegeben ist die Anzahl der Cluster für die aktiven Liganden, und die Anzahl der inaktiven Liganden, die eine Ähnlichkeit von 60 bzw. 70 % mit den aktiven.

Datensatz	Cluster	Anzahl Decoys	60 %	70 %
Brasca	1	243	2	0
Pfeifer	9	180	19	5
Matthews	8	135	0	0
Huang	3	451	4	0
Cheeseright	19	522	16	1
Ermoli	8	360	33	4
Aliagas	3	54	0	0
Tsou	5	694	15	2
Qian	8	243	0	0
Gaisina	2	720	8	0
Surya	12	351	4	0
Karcher	13	171	1	0
Schroeder	1	90	4	0
Coumar	5	288	128	58
Popowycz	3	54	2	2
Menichincheri	5	423	2	0
Renhowe	4	378	119	15
Howard	4	117	7	0
Xia	19	441	0	0

So ähnlich die Ergebnisse auch im Mittel für die jeweiligen Programm sind, unterscheiden sie sich doch in Hinblick auf die jeweiligen Kinasen an denen die Programm getestet wurden. Da nicht bekannt ist, an welche Konformation der Kinase die jeweiligen aktiven Liganden binden wurde in den Folgenden Tabellen (3.4 bis 3.9 auf Seiten 39–52) auch keine Unterscheidung dahingehend gemacht. Es wurden die Mittelwerte für die AUC von allen Varianten ermittelt, ebenso für Kinasen mit Kristallwasser und ohne Kristallwasser. Es konnten nicht für alle Kinasen entsprechende Kristallstrukturen mit kristallisiertem Wasser in der PDB gefunden werden, deshalb gibt es hier weniger Ergebnisse. Teilweise gab es nur eine Kristallstruktur, die unseren Anforderungen entsprach. Aus diesem Grund konnte nicht immer eine Standardabweichung ermittelt werden.

GOLD/ GOLDScore: Für die aktiven Liganden aus dem Datensatz von Qian (PIM) erreichte GOLD einen Mittelwert von allen AUC von 0,90, die Standardabweichung liegt bei 0,03. Mehr als die Hälfte aller aktiven Ligand ist in den oberen 10 % der Bewertung. Abbildung 3.2 zeigt die zwei am besten und die schlechtesten bewerteten aktiven Liganden.

In Abbildung 3.2(a) sind Lösungen für drei aktive Liganden dargestellt. Die Verbindungen *Qian14* und *Qian15* wurden auf die beiden ersten Plätze bewertet, aber beide

Tabelle 3.3: Mittelwerte der AUC und Standardabweichung aller getesteten virtueller Screeningmethoden

	Gesamt		ohne Wasser		mit Wasser	
	Ø	Std	Ø	Std	Ø	Std
GOLDScore	0,65	0,21	0,66	0,20	0,63	0,24
GLIDEScore (SP)	0,67	0,20	0,67	0,21	0,66	0,19
PMF	0,60	0,27	0,65	0,27	0,61	0,28
PScore	0,66	0,23	0,67	0,21	0,63	0,26
ROCS	0,73	0,18	0,73	0,18	–	–
CONSENSUS-SCORE	0,63	0,16	0,63	0,17	0,59	0,15

zeigen keine Interaktionen mit der Hinge-Region. Jedoch bilden die Pyrazinringe dieser Verbindungen Kation- π -Interaktionen zu den Aminosäuren *Ile185* und *Asp186* aus. Zusätzlich werden die jeweiligen primären Amine durch den Carboxyrest von *Asp186* und durch die Acylgruppe des *Asn172* in ihrer Position stabilisiert. Bei *Qian6* ist die einzige Wechselwirkung mit dem Protein die Kation- π -Interaktion, die schon die beiden anderen erwähnten Liganden aufweisen. Diese Verbindung wurde insgesamt am schlechtesten bewertet.

Zum Vergleich sind in Abbildung 3.2(b) die beiden am besten bewerteten Decoys dargestellt. Beide bilden über die Thiazolgruppe zwei Wasserstoffbrücken zur Hinge-Region aus. Zudem haben sie auch die gleiche Kation- π -Interaktion mit dem *Asp186* und eine weitere Wasserstoffbrücke zu der Carboxygruppe des *Glu89*. Diese beiden aus der DUD entnommenen Liganden weisen durch ihre Bindung zur Hinge-Region die Wechselwirkung auf, die von aktiven Liganden erwartet wird.

Am schlechtesten konnte GOLD die Liganden von Tsou an der CDK4 unterscheiden. Dort wurde eine AUC von 0,23 mit einer Standardabweichung von 0,06 erreicht. In Abbildung 3.3(a) sind zum Vergleich die beiden besten Lösungen für die aktiven Substanzen dargestellt *Tsou5jZ* und *Tsou5hZ* sowie die am schlechtesten Bewertete aktive Substanz *Tsou17aZ*. Die beiden am besten bewerteten Substanzen haben jeweils die gleichen Interaktionen zum Protein. Es gibt jeweils zwei Wasserstoffbrücken zwischen dem *His100* der Hinge-Region und einer Ketogruppe am Pyrrolring. Desweiteren gibt es zwei Kation- π -Interaktionen mit dem *Val27*. Zusätzlich bildet *Tsou5jZ* noch eine Brücke zur Acylgruppe des *Asn150* aus. Bei *Tsou5hZ* wird diese Brücke mit dem Schwefel der Thiolgruppe gebildet.

Auch hier bilden die beiden inaktiven Substanzen ähnliche Wechselwirkungen zum Protein aus (s. Abbildung 3.3(b)).

GLIDE/ GLIDEScore (SP): GLIDEScore hat die Inhibitoren von Ermoli für die CDK7 am besten bewertet. Es wurde eine AUC von 0,96 erreicht. Ein Wert von über 0,9 wird auch für die CDK7 mit den aktiven Liganden von Popwycz erreicht, ebenso wie für die CHK1 und die dazu untersuchten Verbindungen von Matthews. In Abbildung 3.4 sind jeweils die beiden am bestbewerteten aktiven und inaktiven Verbindungen gezeigt. Bei

Tabelle 3.4: AUC für den GOLD und GOLDScore für die jeweilige Kinase.

	Gesamt		ohne Wasser		mit Wasser	
	Ø	SD	Ø	SD	Ø	SD
CDK2	0,83	0,31	0,97	0,05	0,63	0,42
CK18	0,67	0,24	0,81	0,10	0,53	0,29
P38	0,59	0,15	0,60	0,12	0,56	0,26
CHK1	0,63	0,07	0,63	0,07	–	–
ABL1	0,73	0,12	0,79	0,07	0,66	0,15
P38	0,63	0,05	0,63	0,06	0,63	0,00
CDK7	0,49	–	0,49	–	–	–
Aurora-2	0,78	0,12	0,77	0,13	0,84	–
CDK4	0,23	0,06	0,23	0,06	–	–
PIM	0,90	0,03	0,89	0,04	0,90	0,01
GSK	0,80	0,06	0,75	0,03	0,85	0,01
C-Jun	0,52	0,05	0,52	0,05	–	–
P38	0,67	0,16	0,66	0,16	0,70	0,22
MET	0,82	0,06	0,80	0,06	0,87	–
Aurora	0,53	0,08	0,55	0,07	0,44	–
CDK2	0,56	0,17	0,56	0,17	–	–
CDK4	0,73	0,15	0,73	0,15	–	–
CDK5	0,82	0,06	0,80	0,05	0,87	–
CDK7	0,73	–	0,73	–	–	–
CDK9	0,80	0,18	0,80	0,18	–	–
CDC7	0,44	0,05	0,41	–	0,47	–
VGFR	0,37	0,07	0,38	0,09	0,35	0,05
Aurora	0,79	0,06	0,81	0,07	0,74	–
PIM	0,27	0,06	0,29	0,01	0,24	0,09
Durchschnitt	0,65	0,21	0,66	0,20	0,63	0,24

den aktiven Substanzen liegt ein hydrophober Ring in der ATP-Bindetasche, und es werden die typischen Wasserstoffbrückenbindungen ausgebildet. Es ist zu erkennen, dass sowohl *Ermoli32Z* und *Ermoli14Z* mit ihren Azaindolen zwei Wasserstoffbrücken zu den Aminosäuren *Asp92* und *Met94* der Hinge-Region bilden. Ebenso bilden beide mit dem Carbonylsauerstoff des Imidazols eine Wasserstoffbrücke zu *Lys41*, sowie mit dem Amidstickstoff zu *Asn141*. Die letzte gemeinsame Wasserstoffbrückenbindung wird mit dem sekundären Amin des Linkers zu *Asp97* ausgebildet.

Die beiden in Abbildung 3.4(b) gezeigten inaktiven Substanzen weisen ein ähnliches Bindungsmuster auf, wie die in Abbildung 3.4(a) gezeigten aktiven Substanzen. Es werden drei Wasserstoffbrücken zu den Aminosäuren *Asp92*, *Lys41* und *Glu20*. Es fehlen aber die stabilisierend wirkenden Wechselwirkungen zu den anderen Aminosäuren in der ATP-Bindetasche. Dabei liegt der 2-Benzopyranonest nicht tief genug in der Tasche. Somit kann der Sauerstoff im Ringsystem keine weitere Wechselwirkung mit der Hinge-Region eingehen. Dafür wird in diesem Fall jedoch eine Wasserstoffbrücke zu *Lys41* ausgebildet.

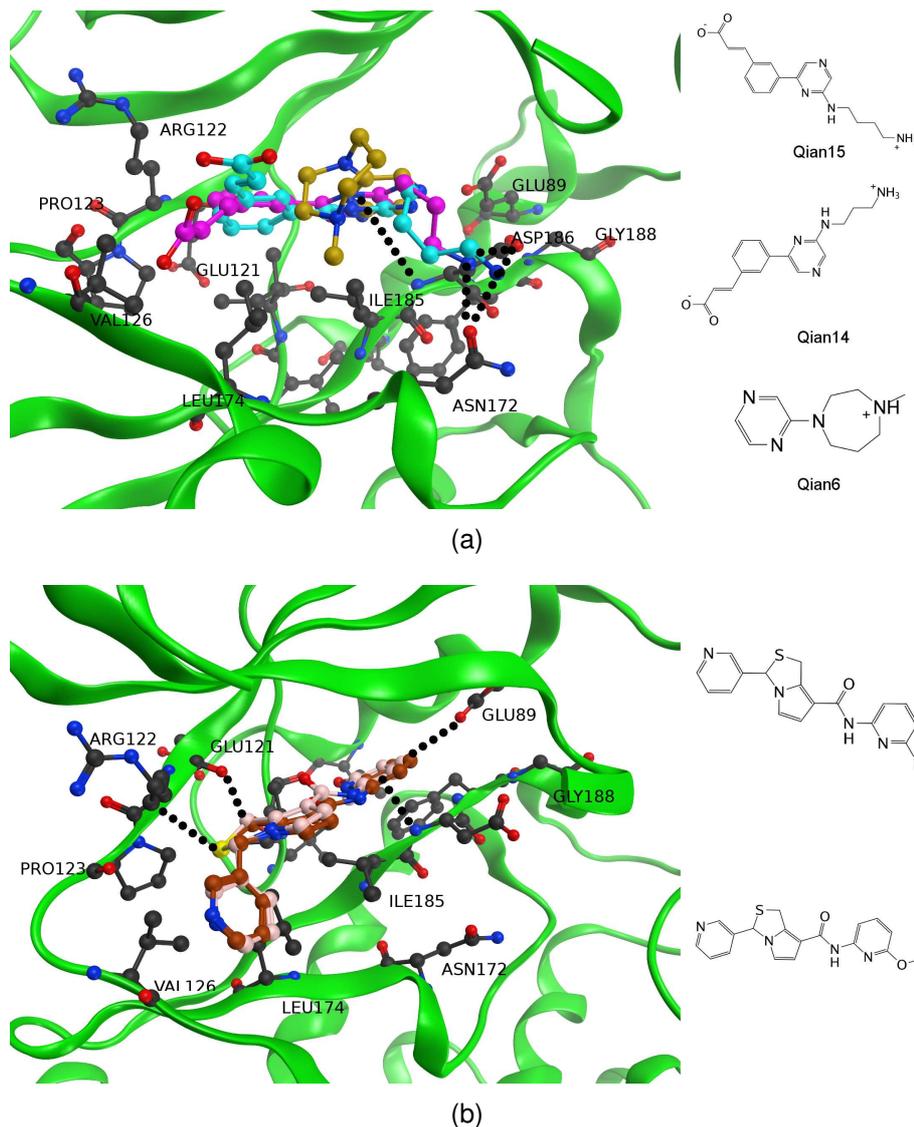


Abbildung 3.2: Aktive und inaktive Inhibitoren PIM gedockt in PDB 2J2I. a) *Qian14* (magenta), *Qian15* (cyan), *Qian6* (gold) b) zwei inaktive Substanzen (braun & rosa); Rezeptoratome schwarz.

Am schlechtesten war die Unterscheidung der aktiven und inaktiven Liganden bei den Substanzen von Popwycz für die CDK2. Die Abbildungen 3.5(a) und 3.5(b) zeigen jeweils die Lösungen für zwei aktive und zwei inaktive Liganden. Die inaktiven Liganden in Abbildung 3.5(b) wurden besser als die beiden aktiven Liganden bewertet. Bei den beiden aktiven Liganden *Purvalanola* und *RoscovitrineR* handelt es sich um zwei bekannte Kinaseinhibitoren. Es fällt aber auf, dass keine der beiden Substanzen eine direkte Wechselwirkung mit der Hinge-Region eingeht. *Purvalanola* bildet nur mit *Asp145* eine Wasserstoffbrückenbindung aus, und mit Pyrimidinanteil des Purin Kation- π - Wechselwirkung aus. *RoscovitrineR* wird durch Kation- π -Wechselwirkungen

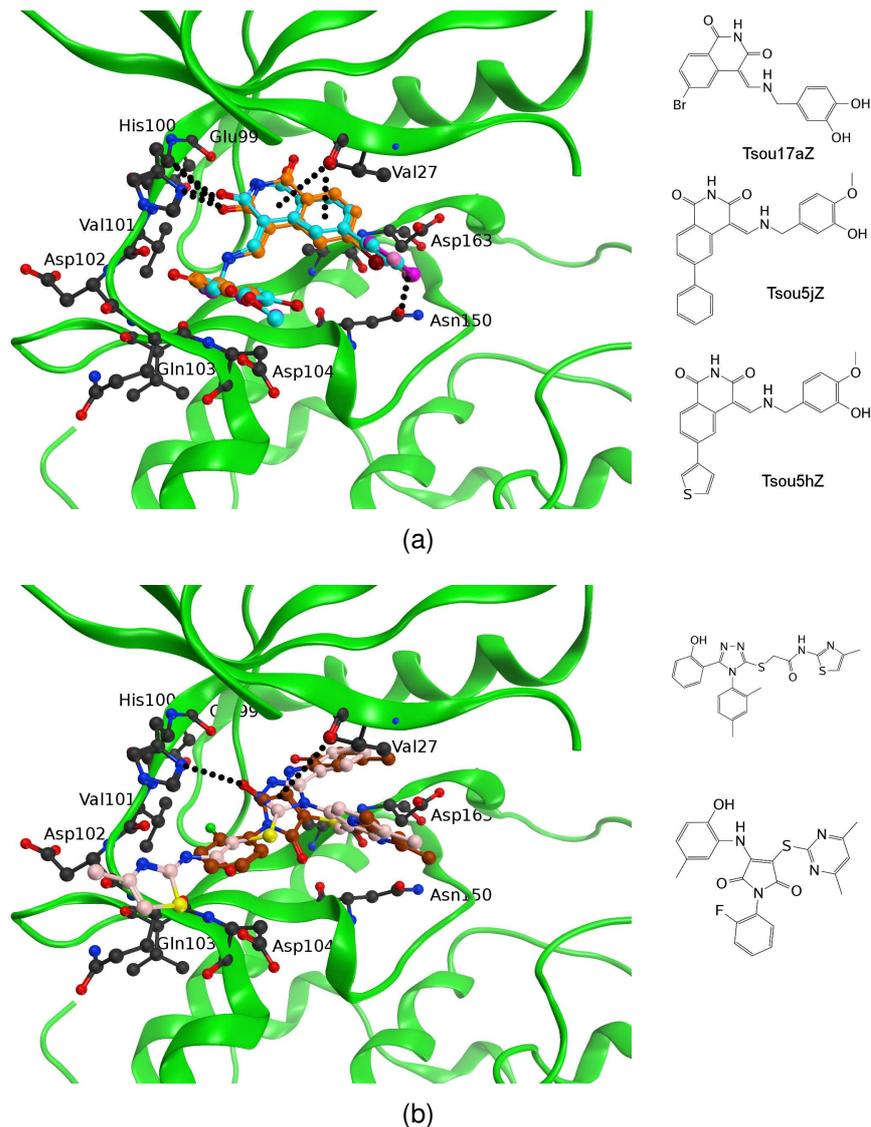


Abbildung 3.3: Aktive und inaktive Inhibitoren CDK4 gedockt in PDB 1XO2. a) *Tsou5jZ* (magenta), *Tsou5hZ* (cyan), *emphTsou17aZ* (gold) b) zwei inaktive Substanzen (braun & rosa); Rezeptoratome schwarz.

zwischen dem Purinanteil und mit den Aminosäuren *Ala144*, *Asp145* und *Val18* in der Binbetasche stabilisiert. Zusätzlich kommen noch zwei Wasserstoffbrückenbindungen zwischen *Asp145* und dem sekundären Amin, sowie der OH-Gruppe und *Asn132*.

Bei den beiden inaktiven Liganden liegen jeweils drei Wasserstoffbrückenbindungen zum Protein vor. Die Bindungen zum *Lys33* und *Asp145* wird von beiden Substanzen ausgebildet. Aber auch hier wird keine Wechselwirkung mit der Hinge-Region eingegangen.

Tabelle 3.5: AUC für GLIDE und GLIDEScore (SP) für die jeweilige Kinase.

CDK2	0,41	0,35	0,24	0,19	0,64	0,42
CK18	0,65	0,17	0,70	0,25	0,61	0,13
P38	0,46	0,08	0,46	0,10	0,46	0,04
CHK1	0,90	0,03	0,90	0,03	–	–
ABL1	0,62	0,15	0,73	0,03	0,47	0,09
P38	0,78	0,07	0,76	0,08	0,83	0,00
CDK7	0,96	–	0,96	–	–	–
Aurora-2	0,69	0,09	0,69	0,10	0,69	–
CDK4	0,66	0,20	0,66	0,20	–	–
PIM	0,61	0,08	0,65	0,05	0,57	0,08
GSK	0,75	0,06	0,71	0,01	0,80	0,00
C-Jun	0,63	0,03	0,63	0,03	–	–
P38	0,87	0,09	0,87	0,11	0,86	0,05
MET	0,55	0,34	0,60	0,37	0,37	–
Aurora	0,57	0,06	0,55	0,04	0,65	–
CDK2	0,43	0,07	0,43	0,07	–	–
CDK4	0,80	0,14	0,80	0,14	–	–
CDK5	0,72	0,04	0,72	0,05	0,71	–
CDK7	0,90	–	0,90	–	–	–
CDK9	0,52	0,52	0,52	0,52	–	–
CDC7	0,68	0,02	0,67	–	0,70	–
VGFR	0,74	0,08	0,76	0,09	0,68	0,01
Aurora	0,75	0,03	0,75	0,03	0,76	–
PIM	0,85	0,02	0,86	0,02	0,84	0,01
Durchschnitt	0,67	0,20	0,67	0,21	0,66	0,19

PARADOCKS/ PMF: Die Sortierung der Ergebnisse führt bei acht Untersuchungen und Verwendung des PMF im Durchschnitt es zu einer AUC von über 0,90. Bei 13 Testungen ist der Wert über 0,50. Die Unterscheidung zwischen aktiven und inaktiven Liganden ist somit gut möglich. Am besten gelingt dies bei den Verbindungen von Popowycz und der CDK7, wobei diese Verbindungen in allen getesteten Kinasen eine AUC von über 0,900 haben. Aber auch hier zeigen die aktiven Liganden keine Interaktionen mit der Hinge-Region, sondern nur zu anderen Aminosäuren innerhalb der ATP-Bindetasche (s. Abbildung 3.6(a)). Die Verbindung *Popowycz7a* bildet nur zwei Interaktionen mit *Lys41*. Diese Wasserstoffbrückenbindung wird auch von *Popowycz7b* ausgebildet, zusätzlich noch aber auch eine Kation- π -Wechselwirkung zwischen der Phenylgruppe und *Thr96*.

Die beiden inaktiven Liganden bilden nur Wechselwirkungen zum *Lys41* aus (s. Abbildung 3.6(b)).

Am schlechtesten war die Unterscheidung bei der C-Jun-Kinase und den Verbindungen von De Surya. Die erste aktive Verbindung ist an Position 84 *Surya8c* der sortieren Ergebnisliste, und die darauf folgende aktive Verbindung ist an Position 94 *Surya3a*.

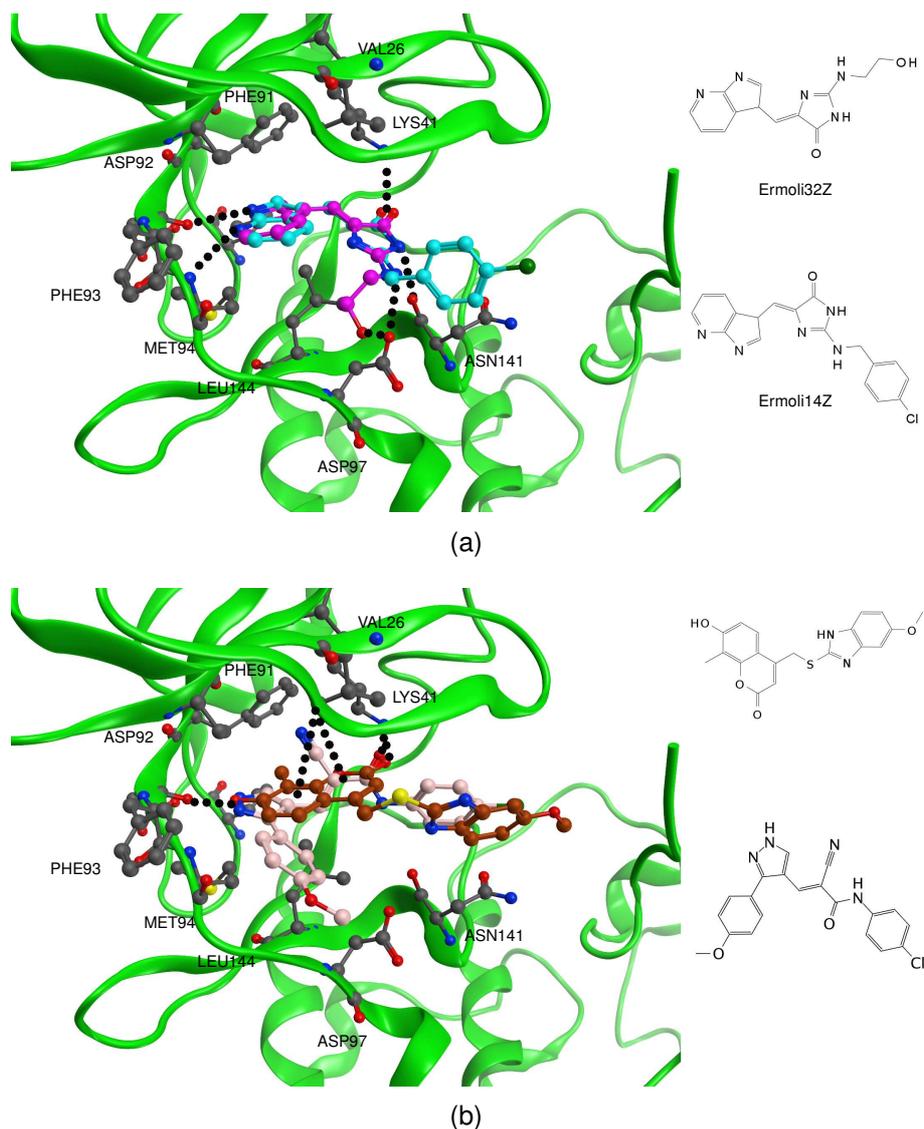


Abbildung 3.4: Aktive und inaktive Inhibitoren CDK7 gedockt in PDB 1UA2. a) *Ermoli32Z* (magenta), *Ermoli14Z* (cyan) b) zwei inaktive Substanzen (braun & rosa); Rezeptoratome schwarz.

In Abbildung 3.7 sind je zwei aktive und zwei inaktive Verbindungen in ihrer jeweiligen gefundenen Position dargestellt. Es ist zu sehen, dass nur eine der aktiven Verbindungen, *Surya8c* innerhalb der ATP-Bindetasche liegt (Abbildung 3.7(a)). Alle anderen geeigneten Lösungen liegen ausserhalb der ATP-Bindetasche. Nur drei der aktiven Liganden werden innerhalb der ATP-Bindetasche platziert, bilden aber nicht das bekannte Muster von Wasserstoffbrücken zur Hinge-Region aus. Sie werden auch so vom PMF bewertet, dass sie erst auf den Rängen ab Position 342 vorkommen.

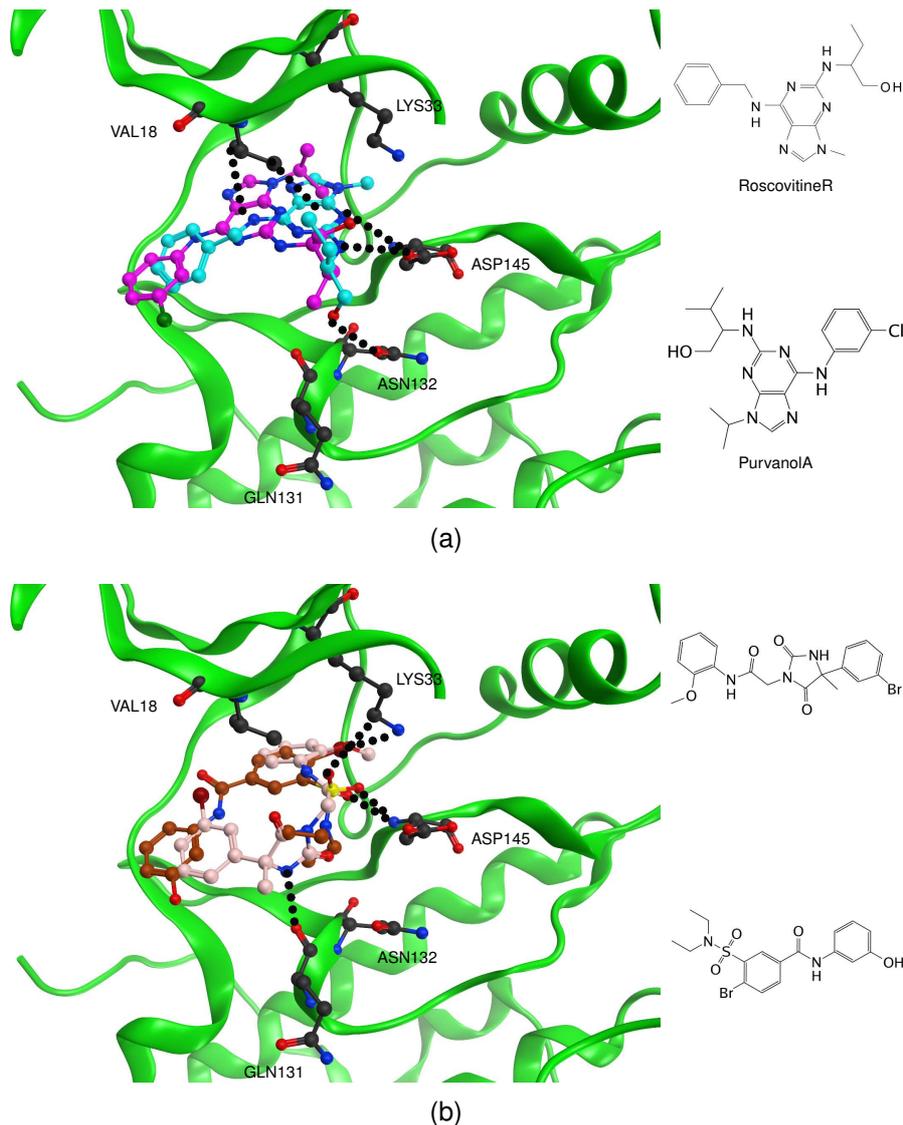


Abbildung 3.5: Aktive und inaktive Inhibitoren CDK2 gedockt in PDB 1E9H. a) *Purvalanola* (magenta), *RoscovitineR* (cyan) b) zwei inaktive Substanzen (braun & rosa);Rezeptoratome schwarz.

PARADOCKS/ PSCORE: Bei den Untersuchungen mit dem PSCORE war die Trennung zwischen den aktiven und inaktiven Liganden am besten bei der MET-Kinase mit den Substanzen von Schröder möglich. Hierbei wurde eine AUC von 0,90 erreicht. Jedoch sind sechs Werte zwischen 0,42 und 0,57. In diesen Fällen war eine gute Unterscheidung nicht möglich, sondern verlief eher zufällig.

In Abbildung 3.8(a) ist zu sehen, wie die Substanz *Schröder11* mit ihrem 2-Aminopyridin zwei Wasserstoffbrücken zu dem Aminosäuren *Pro1158* und *Met1160*. Ebendiese Wechselwirkungen zeigen sich auch bei *Schröder5*. Zusätzlich kommt es hier aber noch zu einer Kation- π -Wechselwirkung mit *Asn1167* und dem 2-Pyridon.

Tabelle 3.6: AUC für PARADOCKS und PMF für die jeweilige Kinase.

	Gesamt		ohne Wasser		mit Wasser	
	Ø	SD	Ø	SD	Ø	SD
CDK2	0,98	0,01	0,98	0,01	0,97	0,01
CK18	0,62	0,21	0,73	0,02	0,49	0,27
P38	0,50	0,14	0,51	0,16	0,47	0,13
CHK1	0,61	0,05	0,61	0,05	–	–
ABL1	0,72	0,09	0,72	0,09	0,73	0,09
P38	0,27	0,15	0,24	0,08	0,33	0,28
CDK7	0,62	–	0,62	–	–	–
Aurora-2	0,49	0,07	0,48	0,07	0,54	–
CDK4	0,36	0,02	0,36	0,02	–	–
PIM	0,91	0,01	0,90	0,00	0,91	0,00
GSK	0,26	0,05	0,30	0,02	0,22	0,01
C-Jun	0,23	0,01	0,23	0,01	–	–
P38	0,66	0,18	0,68	0,17	0,63	0,15
MET	0,50	0,15	0,44	0,09	0,72	–
Aurora	0,92	0,05	0,94	0,03	0,86	–
CDK2	0,99	0,01	0,99	0,01	–	–
CDK4	0,97	0,04	0,97	0,04	–	–
CDK5	0,99	0,01	0,98	0,01	0,99	–
CDK7	1,00	–	1,00	–	–	–
CDK9	0,98	0,02	0,98	0,02	–	–
CDC7	0,34	0,03	0,36	–	0,31	–
VGFR	0,72	0,08	0,73	0,08	0,70	0,09
Aurora	0,76	0,04	0,78	0,01	0,69	–
PIM	0,17	0,02	0,18	0,02	0,16	0,01
Durchschnitt	0,60	0,27	0,65	0,27	0,61	0,28

Diese beiden Substanzen sind auf den Rängen eins und zwei der sortierten Ergebnisliste.

Die beiden inaktiven Liganden bilden jeweils nur eine Wasserstoffbrückenbindung zu *Met1160* bzw. *Gly1163* aus.

Der PSCORE bewertet bei den Untersuchungen an der CDC7-Kinase die als inaktiv bekannten Substanzen deutlich besser, als die aktiven Substanzen. Die R- und S-Enantiomere von Substanz 95 sind die am besten bewerteten aktiven Verbindungen, und kommen in der Rangfolge auf die Plätze 23 und 19.

Rocs: In Tabelle 3.8 sind die Ergebnisse für ROCS aufgeführt. Auch hierbei konnten für die Verbindungen, welche an der MET-Kinase getestet wurden, eine AUC von 1,00 erreicht werden. Alle aktiven Liganden waren am Anfang der sortierten Ergebnisliste. Das gleiche gilt für die Inhibitoren der Aurora-Kinase von Howard. Hier wurden ebenfalls alle aktiven Verbindungen besser als die inaktiven Substanzen bewertet.

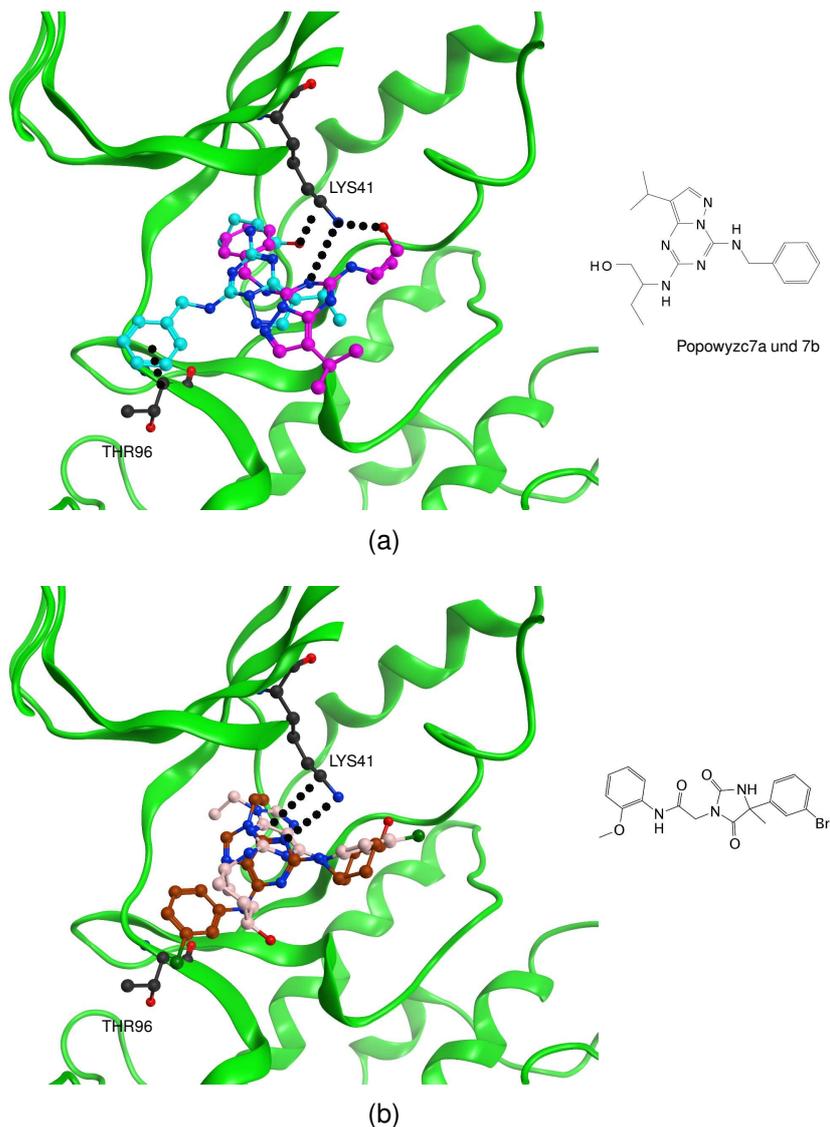


Abbildung 3.6: Aktive und inaktive Inhibitoren CDK7 gedockt in PDB 1UA2. a) *Popowyc7a* (magenta), *Popowyc7b* (cyan). Hierbei handelt es sich um die R- und S-Enantiomere der gleichen Verbindung. b) zwei inaktive Substanzen (braun & rosa). Rezeptoratome schwarz.

Es ist auch zu sehen, dass für die AUC kein Wert schlechter als 0,42 ist. Insgesamt wurden mittels ROCS die inaktiven Verbindungen nicht deutlich besser als die aktiven Verbindungen bewertet.

CONSENSUS-SCORE: Der CONSENSUS-SCORE stellt einen Mittelwert aus verschiedenen Scoring-Funktionen dar. In vier Fällen waren die Werte für die AUC schlechter als eine zufällige Verteilung (s. Tabelle 3.9). Für Verbindungen von Xia und die PIM-Kinase verlief die Untersuchung so ungünstig, dass nahezu alle inaktiven Verbindun-

Tabelle 3.7: AUC für PARADOCKS und PSCORE für die jeweilige Kinase

	Gesamt		ohne Wasser		mit Wasser	
	Ø	SD	Ø	SD	Ø	SD
CDK2	0,88	0,12	0,96	0,02	0,77	0,10
CK18	0,62	0,18	0,71	0,07	0,53	0,25
P38	0,53	0,15	0,54	0,19	0,51	0,09
CHK1	0,53	0,23	0,53	0,23	–	–
ABL1	0,79	0,10	0,84	0,03	0,72	0,14
P38	0,35	0,15	0,32	0,10	0,40	0,27
CDK7	0,49	–	0,49	–	–	–
Aurora-2	0,80	0,12	0,84	0,10	0,66	–
CDK4	0,52	0,10	0,52	0,10	–	–
PIM	0,81	0,07	0,75	0,00	0,87	0,01
GSK	0,85	0,07	0,83	0,06	0,87	0,1
C-Jun	0,42	0,04	0,42	0,04	–	–
P38	0,90	0,06	0,91	0,03	0,86	0,10
MET	0,90	0,53	0,88	0,02	0,99	–
Aurora	0,68	0,05	0,70	0,03	0,60	–
CDK2	0,62	0,34	0,62	0,34	–	–
CDK4	0,61	0,03	0,61	0,03	–	–
CDK5	0,80	0,06	0,81	0,08	0,79	–
CDK7	0,57	–	0,57	–	–	–
CDK9	0,60	0,25	0,60	0,25	–	–
CDC7	0,16	0,03	0,18	–	0,13	–
VGFR	0,64	0,10	0,66	0,11	0,60	0,11
Aurora	0,61	0,18	0,68	0,03	0,30	–
PIM	0,17	0,02	0,17	0,04	0,16	0,01
Durchschnitt	0,66	0,23	0,67	0,21	0,63	0,26

fand keine erneute Konformationsgenerierung statt, und die Lage der Eingabestruktur innerhalb des Proteins wurde nicht verändert.

In Tabelle 3.20 wird die Anzahl aller positiven und negativen Veränderungen auf die AUC dargestellt. Es ist zu sehen, dass durch die CHEMSCORE bei GOLD und PMF die Anzahl der positiven und negativen Veränderungen etwa die Waage halten, nur bei PSCORE überwiegen die Verschlechterungen. Das Rescoring mit GOLDScore führt in den meisten Fällen zu einer Verschlechterung der AUC.

Tabelle 3.11 zeigt die Werte für die AUC nach der Neubewertung der Untersuchungen durch eine weitere Scoring-Funktion. Die Neubewertung wurde mit dem CHEMSCORE und dem GLIDEScore durchgeführt. Vergleicht man diese Werte mit denen aus Tabelle 3.3, so sieht man, dass nur in den Fällen des PMF und PSCORE mit Wasser in der Bindetasche es zu einer Verbesserung kam, wodurch sich auch der Wert der AUC für die gesamte Untersuchung verbessert hat, jedoch nur mit Verwendung des CHEMSCORE.

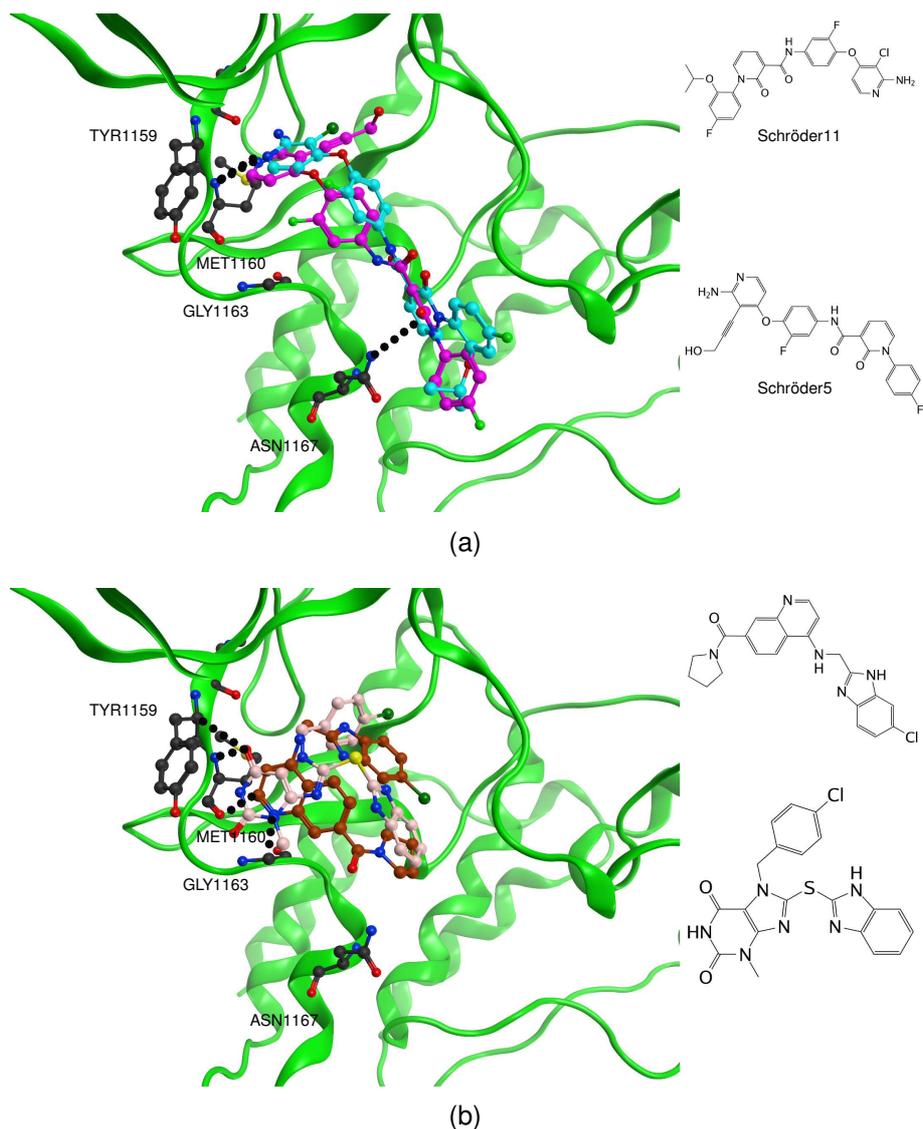


Abbildung 3.8: Aktive und inaktive Inhibitoren MET gedockt in PDB 2WGJ. a) *Schröder5* (magenta), *Schröder11* (cyan) b) zwei inaktive Substanzen (braun & rosa); Rezeptoratome schwarz.

3.1.6 Auswertung der EF

Der Enrichment Factor ist ein Maß dafür, wie gut ein VS zwischen richtig positiven und wahr negativen unterscheiden kann. Diese Methode ist besonders interessant um eine Aussage darüber zu treffen, wie gut in dem oberen Bereich der Liste die Unterscheidung funktioniert. Genauso wie bei dem Wert für die AUC steigt der Wert des EF mit jedem gefundenen aktiven Liganden um einen bestimmten Teilbetrag an. Dieser ist abhängig von der Anzahl der aktiven und inaktiven Liganden (s. Formel 2.32).

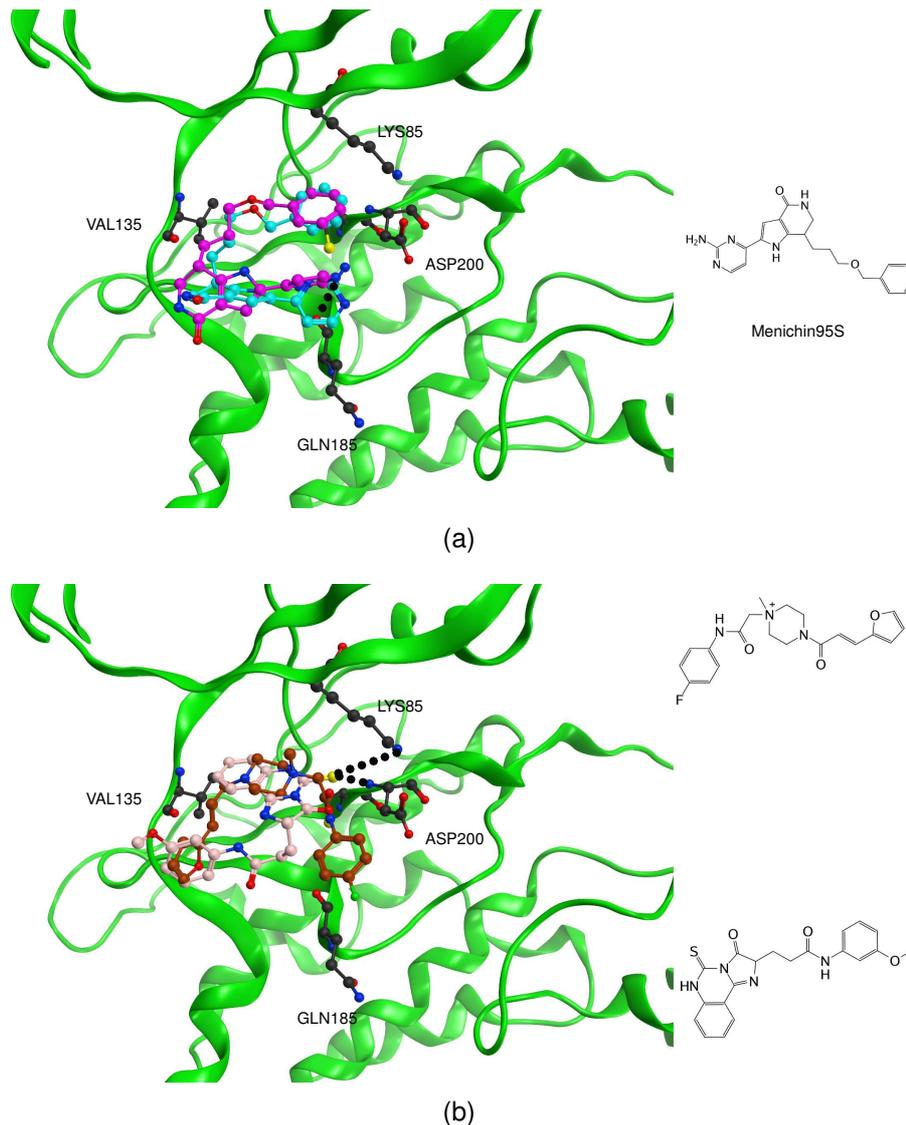


Abbildung 3.9: Aktive und inaktive Inhibitoren CDC7 gedockt in PDB 3DU8. a) *Menichin95S* (magenta), *Menichin95R* (cyan) b) zwei inaktive Substanzen (braun & rosa); Rezeptoratome schwarz.

In diesem Fall ist der jeweilige Maximalwert für den EF 10. Dieser wird dann erreicht, wenn bis zu der jeweiligen Grenze nur aktive Liganden in der Liste stehen.

Hier wurden die Enrichment Factors nach den ersten 5% der Datenbank errechnet. Die entsprechenden Werte sind in Tabelle 3.12 aufgeführt. Auch hier zeigt sich, dass ROCS mit einem EF von 5,42 am besten zwischen den aktiven und inaktiven Liganden unterscheiden kann. Somit sind innerhalb der ersten 5% etwas mehr als die Hälfte aktive Liganden. GOLD und der GOLDScore zeigen hier keine so gute Leistung. Der durchschnittliche Wert liegt bei 2,72. D. h. nicht mal ein Drittel der oberen Ränge in der Datenbank werden von aktiven Liganden belegt. Auch haben alle Methoden Probleme

Tabelle 3.8: AUC für Rocs für die jeweilige Kinase.

	Gesamt	
	Ø	SD
CDK2	0,99	–
CK18	0,50	–
P38	0,50	–
CHK1	0,45	–
ABL1	0,81	–
P38	0,52	–
CDK7	0,89	–
Aurora-2	0,81	–
CDK4	0,58	–
PIM	0,77	–
GSK	0,95	–
C-Jun	0,62	–
P38	0,78	–
MET	1,00	–
Aurora	0,42	–
CDK2	0,75	–
CDK4	0,75	–
CDK5	0,75	–
CDK7	0,75	–
CDK9	0,75	–
CDC7	1,00	–
VGFR	0,58	–
Aurora	1,00	–
PIM	0,71	–
Durchschnitt	0,73	0,18

damit die aktiven Liganden besser als die inaktiven Liganden zu bewerten, wenn noch Wasser in der Bindetasche vorhanden ist.

GOLD/ GOLDScore: In nur fünf Fällen wurde ein durchschnittlicher EF größer oder gleich als 5,00 gefunden. Wurden die Untersuchungen ohne Wasser in der Bindetasche durchgeführt, so war er in nur vier Fällen größer oder gleich 5,00, war zusätzlich noch Wasser in der ATP-Bindetasche so war der Wert nur dreimal größer als 5,00.

Betrachtet man nur den EF innerhalb der ersten 5% der Ergebnisse, so zeigt der GOLDScore das beste Ergebnis bei der CDK2 und den aktiven Liganden von Brasca. Dies ist sowohl im Durchschnitt der Fall, als auch bei der Betrachtung der Dockingergebnisse wenn kein Kristallwasser mehr in der Bindetasche war. Zwischen den aktiven von Popowycz und den Substanzen aus der DUD kann GOLD nicht so gut unterscheiden. Bezieht man noch das Wasser in das Docking mit ein, dann ist die Unterscheidung am besten bei PIM und den aktiven Liganden von Qian. Dies deckt sich mit den Beobachtungen der AUC für GOLD.

Tabelle 3.9: AUC für CONSENSUS-SCORE für die jeweilige Kinase.

	Gesamt		ohne Wasser		mit Wasser	
	Ø	SD	Ø	SD	Ø	SD
CDK2	0,76	0,07	0,77	0,02	0,76	0,12
CK18	0,45	0,18	0,37	0,06	0,52	0,26
P38	0,51	0,11	0,51	0,11	0,50	0,14
CHK1	0,46	0,07	0,46	0,07	–	–
ABL1	0,66	0,12	0,73	0,03	0,57	0,14
P38	0,54	0,11	0,55	0,14	0,52	0,04
CDK7	0,60	–	0,60	–	–	–
Aurora-2	0,70	0,06	0,71	0,07	0,68	–
CDK4	0,53	0,16	0,53	0,16	–	–
PIM	0,72	0,02	0,73	0,03	0,71	0,01
GSK	0,58	0,02	0,59	0,01	0,57	0,03
C-Jun	0,50	0,02	0,50	0,02	–	–
P38	0,79	0,10	0,84	0,05	0,71	0,13
MET	0,74	0,04	0,75	0,04	0,73	–
Aurora	0,63	0,01	0,63	0,01	0,64	–
CDK2	0,72	0,14	0,72	0,14	–	–
CDK4	0,80	0,03	0,80	0,03	–	–
CDK5	0,79	0,02	0,80	0,01	0,76	–
CDK7	0,82	–	0,82	–	–	–
CDK9	0,80	0,05	0,80	0,05	–	–
CDC7	0,40	0,01	0,39	–	0,40	–
VGFR	0,61	0,05	0,62	0,05	0,58	0,05
Aurora	0,76	0,07	0,79	0,01	0,63	–
PIM	0,18	0,04	0,17	0,07	0,19	0,01
Durchschnitt	0,63	0,16	0,63	0,17	0,59	0,15

Tabelle 3.10: Anzahl der Verbesserungen und Verschlechterungen der AUC nach dem Rescoring mit CHEMScore und GLIDEScore (SP)

	CHEMScore		GLIDEScore	
	Besser	Schlechter	Besser	Schlechter
GOLDScore	48	44	31	62
PMF	46	47	24	69
PScore	35	51	23	69

GLIDE/ GLIDEScore (SP): Bei acht der untersuchten Substanzdatenbanken konnte im Durchschnitt ein EF von über 5,00 erreicht werden. Wurde das Wasser aus der ATP-Bindetasche entfernt wurde ein Wert von über 5,000 sechs mal erreicht, mit Wasser in der ATP-Bindetasche ebenso. In keinem Fall wurde ein Wert von 10,00 erzielt, somit wurden vom GLIDEScore nie alle untersuchten aktiven Liganden besser bewertet als die inaktiven Liganden.

Tabelle 3.11: Mittelwerte der AUC und Standardabweichung aller Programme nach dem Rescoring

	Gesamt		ohne Wasser		mit Wasser	
	Ø	SD	Ø	SD	Ø	SD
GOLD mit CHEMScore	0,64	0,19	0,64	0,19	0,63	0,21
PMF mit CHEMScore	0,66	0,19	0,63	0,19	0,73	0,17
PScore mit CHEMScore	0,63	0,21	0,59	0,20	0,71	0,20
GOLD mit GLIDEScore	0,56	0,13	0,51	0,14	0,52	0,10
PMF mit GLIDEScore	0,46	0,13	0,47	0,14	0,42	0,10
PScore mit GLIDEScore	0,48	0,15	0,48	0,16	0,47	0,13

Tabelle 3.12: Mittelwerte des EF5 und Standardabweichung aller getesteten virtueller Screeningmethoden

	Gesamt		ohne Wasser		mit Wasser	
	Ø	Std	Ø	Std	Ø	Std
GOLD/ GOLDScore	2,72	3,17	2,72	3,17	2,80	3,25
GLIDE/ GLIDEScore (SP)	3,67	2,68	3,76	2,63	3,50	2,91
PARADOCKS/ PMF	4,46	3,72	4,72	3,87	3,50	2,91
PARADOCKS/ PScore	4,24	3,33	4,42	3,28	3,93	3,44
ROCS	5,42	2,62	–	–	–	–
CONSENSUS-SCORE	4,74	2,46	4,87	2,51	4,50	2,32

Die beste Unterscheidung war in diesem Fall mit einem EF von 8,13 bei der CHK1 möglich. Wurde das Kristallwasser ignoriert, so wird bei der Aurora-Kinase mit 8,21 der höchste EF erzielt. Mit dem kristallisiertem Wasser im Umfeld der ATP-Bindetasche wurde der beste EF mit 8,13 bei den Untersuchungen an der PIM-Kinase erreicht.

PARADOCKS/ PMF: Bei Verwendung des PMF wurde in drei der untersuchten Fälle nur aktive Liganden so gut bewertet, dass diese ausschließlich in den ersten 5 % der Ergebnisse vorkommen. Insgesamt wurde in 11 von 24 Fällen ein EF von über 5 erzielt. Nur im Falle der C-Jun-Kinase wurden die aktiven Liganden so schlecht bewertet, dass sie nicht innerhalb der Grenze von 5 % liegen.

PARADOCKS/ PScore: Bei sieben Untersuchungen wurde für den EF ein größerer Wert als 5,00 erreicht. Somit waren mindestens 50 % der Verbindungen innerhalb der ersten 5 % der sortierten Ergebnisse aktive Liganden. In den Fällen ohne Wasser in der ATP-Bindetasche wurde für die CDK2 ein EF von 10,00 ermittelt. Somit wurden alle aktiven Liganden besser bewertet als die inaktiven Liganden. Das gleiche gilt für die Untersuchungen an der MET-Kinase mit Wasser in der ATP-Bindetasche. Im Falle der CDK7 wurden die inaktiven Substanzen so gut bewertet, dass innerhalb der obersten 10 % keine aktiven Liganden vorkommen.

Tabelle 3.13: EF bei 5 % für GOLD und GOLDScore für die jeweilige Kinase

	Gesamt		ohne Wasser		mit Wasser	
	Ø	SD	Ø	SD	Ø	SD
CDK2	6,80	4,41	9,06	1,89	3,79	5,42
CK18	5,01	4,61	8,86	1,61	1,16	1,64
P38	1,67	2,15	1,16	1,34	2,71	3,83
CHK1	3,13	0,88	–	–	–	–
ABL1	2,59	2,30	3,57	2,31	1,28	1,90
P38	1,70	0,87	1,64	1,11	1,81	0,23
CDK7	0,91	–	–	–	–	–
Aurora-2	4,00	3,62	4,17	4,19	3,33	–
CDK4	0,26	0,37	–	–	–	–
PIM	6,01	2,86	3,81	1,69	8,22	1,52
GSK	4,38	1,93	2,75	0,35	6,00	0,71
C-Jun	2,00	0,00	–	–	–	–
P38	1,00	1,55	0,75	1,50	1,50	2,12
MET	2,40	1,67	2,00	1,63	4,00	–
Aurora	0,13	0,28	0,16	0,31	0,00	–
CDK2	0,00	0,00	–	–	–	–
CDK4	2,50	3,19	–	–	–	–
CDK5	6,67	3,33	5,00	2,36	6,67	–
CDK7	0,00	–	–	–	–	–
CDK9	5,00	7,07	–	–	–	–
CDC7	0,42	0,59	0,83	–	0,00	–
VGFR	0,07	0,18	0,11	0,22	0,00	0,00
Aurora	4,29	3,64	5,00	3,78	1,43	–
PIM	0,75	0,53	0,93	0,24	0,58	0,82
Durchschnitt	2,72	3,17	2,72	3,17	2,80	3,25

Rocs: In drei Untersuchungen mittels Rocs konnte ein EF von 10,00 erreicht werden, und in elf Fällen wurde ein Wert von größer oder gleich fünf erzielt werden.

Rescoring und EF: Durch das Rescoring der Ergebnisse konnte bis auf eine Ausnahme immer eine Verbesserung des EF erzielt werden, wie in Tabelle 3.18 aufgeführt ist. Bei den Ergebnissen des PMF und dem erneuten bewerten durch den CHEMSCORE halten sich die einzelnen Verbesserungen und Verschlechterungen etwa die Waage. Bei der Neubewertung der Lösungen des PSCORE durch den CHEMSCORE überwiegt die Anzahl der Verschlechterungen der Ergebnisse. Am deutlichsten konnten die Ergebnisse des PMF durch den GLIDEScore verbessert werden. Hier stehen 50 Verbesserungen nur 35 Verschlechterungen gegenüber. In jedem Fall konnten aber die Ergebnisse von mindestens 40 Dockinguntersuchungen verbessert werden.

Tabelle 3.14: EF bei 5% für GLIDE und GLIDEScore (SP) für die jeweilige Kinase

	Gesamt		ohne Wasser		mit Wasser	
	Ø	SD	Ø	SD	Ø	SD
CDK2	3,65	2,45	2,77	1,87	4,84	3,01
CK18	4,07	3,70	6,97	2,17	1,16	1,64
P38	4,64	0,98	5,03	1,00	3,86	0,00
CHK1	8,13	0,88	–	–	–	–
ABL1	1,65	1,44	2,70	1,36	0,26	0,22
P38	2,24	0,32	2,13	0,33	2,46	0,23
CDK7	0,91	–	–	–	–	–
Aurora-2	5,33	1,83	5,00	1,93	6,67	–
CDK4	0,26	0,21	–	–	–	–
PIM	6,37	2,78	4,52	2,70	8,23	1,52
GSK	3,50	0,41	3,25	0,35	3,75	0,35
C-Jun	2,50	0,00	–	–	–	–
P38	5,33	1,37	5,25	1,26	5,50	2,12
MET	0,40	0,90	0,00	0,00	2,00	–
Aurora	5,13	2,98	4,53	3,08	7,50	–
CDK2	3,33	0,00	–	–	–	–
CDK4	4,17	3,19	–	–	–	–
CDK5	5,56	1,93	5,00	2,36	6,67	–
CDK7	0,00	–	–	–	–	–
CDK9	5,00	2,36	–	–	–	–
CDC7	0,47	0,59	0,83	–	0,00	–
VGFR	3,34	0,92	3,45	1,17	3,16	0,00
Aurora	8,00	0,78	8,21	0,71	7,14	–
PIM	0,38	0,31	0,57	0,27	0,19	0,19
Durchschnitt	3,67	2,68	3,76	2,63	3,50	2,91

Tabelle 3.15: EF bei 5% für PARADOCKS und PMF für die jeweilige Kinase

	Gesamt		ohne Wasser		mit Wasser	
	Ø	SD	Ø	SD	Ø	SD
CDK2	10,00	0,00	10,00	0,00	10,00	0,00
CK18	5,22	2,55	6,96	0,00	3,48	2,73
P38	2,71	2,67	2,32	2,96	3,48	2,73
CHK1	6,25	0,00	–	–	–	–
ABL1	2,97	2,17	3,28	2,23	2,57	1,82
P38	0,45	0,55	0,33	0,38	0,68	0,96
CDK7	2,74	–	–	–	–	–
Aurora-2	4,00	2,79	4,17	3,19	3,33	–
CDK4	0,78	0,37	–	–	–	–
PIM	6,25	0,36	6,43	0,00	6,07	0,51
GSK	0,13	0,25	0,25	0,35	0,00	0,00
C-Jun	0,00	0,00	–	–	–	–
P38	2,50	2,74	2,75	3,10	2,00	2,83
MET	1,00	1,41	0,50	1,00	3,00	–
Aurora	8,75	1,17	0,06	1,08	7,50	–
CDK2	10,00	0,00	–	–	–	–
CDK4	8,33	3,33	–	–	–	–
CDK5	8,89	1,92	8,33	2,36	10,00	–
CDK7	10,00	–	–	–	–	–
CDK9	8,33	2,36	–	–	–	–
CDC7	1,04	0,30	0,83	–	1,25	–
VGFR	4,97	2,17	4,78	2,37	5,34	2,52
Aurora	6,71	2,35	7,68	1,07	2,86	–
PIM	0,57	0,22	0,57	0,27	0,57	0,27
Durchschnitt	4,46	3,72	4,72	3,87	3,50	2,91

Tabelle 3.16: EF bei 5% für PARADOCKS und PSCORE für die jeweilige Kinase

	Gesamt		ohne Wasser		mit Wasser	
	Ø	SD	Ø	SD	Ø	SD
CDK2	8,82	2,45	10,00	0,00	7,25	3,39
CK18	6,17	3,93	8,86	1,61	3,48	3,83
P38	3,47	4,30	5,20	4,33	0,00	0,00
CHK1	1,88	0,88	–	–	–	–
ABL1	4,79	2,03	5,88	1,36	3,34	2,57
P38	0,67	1,29	0,19	0,22	1,64	2,32
CDK7	1,37	–	–	–	–	–
Aurora-2	4,67	1,83	5,00	1,93	3,33	–
CDK4	0,78	0,76	–	–	–	–
PIM	4,75	2,06	3,07	0,64	6,43	1,01
GSK	5,50	2,86	4,25	1,77	6,75	3,89
C-Jun	1,00	0,72	–	–	–	–
P38	6,83	3,19	7,25	2,36	6,00	5,66
MET	7,20	2,28	6,50	1,92	10,00	–
Aurora	4,74	2,93	5,30	3,06	2,50	–
CDK2	3,33	4,71	–	–	–	–
CDK4	3,33	2,72	–	–	–	–
CDK5	5,56	1,93	5,00	2,36	6,67	–
CDK7	0,00	–	–	–	–	–
CDK9	3,33	4,71	–	–	–	–
CDC7	0,63	0,29	0,83	–	0,42	–
VGFR	2,30	1,33	2,34	0,92	2,23	2,52
Aurora	5,86	3,37	7,32	0,90	0,00	–
PIM	0,76	0,31	0,95	0,27	0,57	0,27
Durchschnitt	4,24	3,33	4,42	3,28	3,93	3,44

Tabelle 3.17: EF für RocS für die jeweilige Kinase.

	Gesamt	
	Ø	SD
CDK2	10,00	–
CK18	6,18	–
P38	6,18	–
CHK1	3,75	–
ABL1	5,78	–
P38	4,56	–
CDK7	3,19	–
Aurora-2	6,67	–
CDK4	1,82	–
PIM	7,14	–
GSK	4,00	–
C-Jun	5,00	–
P38	5,00	–
MET	10,00	–
Aurora	2,50	–
CDK2	3,33	–
CDK4	3,33	–
CDK5	3,33	–
CDK7	3,33	–
CDK9	3,33	–
CDC7	10,00	–
VGFR	5,34	–
Aurora	7,14	–
PIM	1,52	–
Durchschnitt	5,42	2,62

Tabelle 3.18: Anzahl der Verbesserungen und Verschlechterungen des EF nach dem Rescoring mit CHEMScore und GLIDEScore

	CHEMScore		GLIDEScore (SP)	
	Besser	Schlechter	Besser	Schlechter
GOLD/ GOLDScore	44	26	44	38
PMF	45	44	50	35
PScore	41	43	44	34

Tabelle 3.19: Mittelwerte RIE
und Standardabweichung aller getesteten virtueller Screeningmethoden

	Gesamt		ohne Wasser		mit Wasser	
	Ø	Std	Ø	Std	Ø	Std
GOLD/ GOLDScore	2,28	2,30	2,29	2,30	2,33	2,33
GLIDE/ GLIDEScore	2,32	1,53	2,39	1,63	1,93	1,22
PARADOCKS/ PMF	1,45	1,88	1,39	1,86	1,59	2,00
PARADOCKS/ PSCORE	1,51	1,84	1,47	1,77	1,57	2,05
ROCS	3,05	1,36	3,05	1,36	–	–
CONSENSUS-SCORE	2,82	1,97	3,05	2,17	2,50	1,37

3.1.7 Auswertung RIE

Die durchschnittlichen Ergebnisse für RIE sind in Tabelle 3.19 dargestellt. Da es keine obere Grenze für RIE gibt, lassen sich die Werte nur schwer miteinander vergleichen. Es ist aber zu sehen, dass ROCS wieder die beste Leistung zeigt, während PMF und PSCORE hinter den anderen Scoring-Funktionen zurückliegen. Aus diesem Grund ist hier nur die Übersicht über alle SF aufgeführt. Die Auflistung, welche Werte für RIE jede SF bei jeder Testung erlangt hat, sind im Anhang (s. Kapitel 5.1.3) aufgeführt.

Rescoring und RIE: Für jede SF führte eine Neubewertung der Ergebnisse dazu, dass in der Mehrzahl die Ergebnisse verbessert werden konnten. Am deutlichsten ist dies in der Kombination PMF und CHEMScore zu erkennen. Hier konnten in 73 Fällen der Wert für RIE, und somit die frühe Unterscheidung von aktiven und inaktiven Verbindungen, verbessert werden. Bei keiner Kombination kam es zu dem Fall, dass die Anzahl der Verschlechterungen höher war, als die der Verbesserungen.

Tabelle 3.20: Anzahl der Verbesserungen und Verschlechterungen von RIE nach dem Rescoring mit CHEMScore und GLIDEScore

	CHEMScore		GLIDEScore (SP)	
	Besser	Schlechter	Besser	Schlechter
GOLD/ GOLDScore	55	41	57	42
PARADOCKS/ PMF	73	24	50	48
PARADOCKS/ PScore	55	41	53	44

3.1.8 Diskussion

Es wurde eine retrospektive Studie virtueller Screening Methoden anhand von 19 Datensätzen aus der Literatur entnommener Kinaseinhibitoren durchgeführt. Hierzu wurden verschiedene Dockingprogramme und Scoring-Funktionen verwendet. Die Ergebnisse aus diesen Dockingstudien wurden darauf hin mit anderen Scoring-Funktionen neu bewertet. Zusätzlich wurde noch ein Consensus-Score-Modell verwendet. Drei verschiedene Gütemaße wurden herangezogen, um die Qualität der Ergebnisse zu überprüfen. Das Ziel von allen Dockingprogrammen und SF ist es zwischen aktiven und inaktiven Liganden zu unterscheiden. Das heißt, dass die aktiven Liganden möglichst am Anfang der sortierten Ergebnisliste vorkommen. Die verschiedenen Kombinationen von Dockingprogramm und SF haben dabei unterschiedlich gut funktioniert, und nicht jede Kombination eignet sich gleichermaßen, um zuverlässig zwischen aktiven und inaktiven Verbindungen zu unterscheiden.

Jedes Gütemaß betrachtet einen etwas anderen Aspekt der Ergebnisse. AUC und RIE treffen Aussagen über die Verteilung der aktiven Liganden innerhalb der vollständigen Ergebnisliste, während der EF sich nur auf die ersten 5 % der Liste konzentriert.

AUC: Alle Werte für die AUC liegen über 0,6. Der Wert für das Programm ROCS sogar über 0,7. Diese Werte wurden auch schon in anderen Veröffentlichungen bei anderen Targetproteinen gefunden [217–219]. Somit sind alle untersuchten Methoden insgesamt in der Lage ein überdurchschnittliches Ergebnis zu produzieren. PMF hat in einem Fall und ROCS in drei Fällen eine AUC von 1,0 erzielt, und damit alle aktiven Liganden besser bewertet als die inaktiven. Alle getesteten Methoden liegen in ihrer Leistung dicht beieinander, und es kann kein signifikanter Unterschied zwischen den einzelnen Methoden erkannt werden.

ROCS hat gegenüber den anderen Methoden den Vorteil, dass es eine Ligandenbasierte Methode ist. Sie benutzt die Oberfläche eines Referenzliganden um anhand dieser zu bestimmen, wie gut eine andere Substanz in die Bindetasche passt. In diesem Fall dienten die in den Kristallstrukturen vorhandenen Liganden als Vorlage, um aus einer Konformationsbibliothek der zu untersuchenden Liganden die jeweils beste Konformation zu finden. ROCS zieht in die Bewertung der Pose keine Wechselwirkungen mit dem Protein oder mit vorhandenem Wasser mit ein. Nur die Größe und Form der Verbindungen wird berücksichtigt, somit sind es auch diese Faktoren, die das Ergebnis negativ beeinflussen können. Für die Qualität von ROCS ist es wichtig, welche Liganden als Vorlage für die Erzeugung der Oberfläche verwendet werden [220].

Die Leistung des CONSENSUS-SCORE unterscheidet sich nicht signifikant von der Leistung der einzelnen Scoring-Funktionen.

Wurden die Dockinglösungen mit einer anderen Scoring-Funktion neu bewertet, so zeigte sich, dass bei der Verwendung des CHEMSCORE die Ergebnisse nahezu gleich bleiben, oder sich leicht verbessern. Der GLIDESCORE führt insgesamt jedoch zu einer Verschlechterung der Ergebnisse. Dabei ist der GLIDESCORE (3.1)

$$\begin{aligned} \Delta G_{bind} = & C_{lipo-lipo} \sum f(r_{lr}) + C_{hbond-neut-neut} \sum g(\Delta r)h(\Delta\alpha) + \\ & C_{hbond-neut-charged} \sum g(\Delta r)h(\Delta\alpha) + C_{hbond-charged-charged} \sum g(\Delta r)h(\Delta\alpha) + \\ & C_{max-metal-ion} \sum f(r_{lm}) + C_{rotb}H_{rotb} + C_{polar-phob}V_{polar-phob} + \\ & C_{coul}E_{coul} + C_{vdW}E_{vdW} + solvationterms \end{aligned} \quad (3.1)$$

jedoch eine Erweiterung [221] des CHEMScore (3.2), welcher definiert ist als [222]

$$\begin{aligned} \Delta G_{bind} = & C_0 + C_{lipo} \sum f(r_{lr}) + C_{hbond} \sum g(\Delta r)h(\Delta\alpha) \\ & + C_{metal} \sum f(r_m) + C_{rotb}H_{rotb} \end{aligned} \quad (3.2)$$

Die Einführung von Coulomb- und van-der-Waals-Wechselwirkungen haben wahrscheinlich den grössten Einfluß darauf, dass die Neubewertung schlechter ausfällt als bei der Verwendung des reinen CHEMScore. Eine Verwendung der *extra precision* (XP) Scoringvariante bei GLIDE war hier nicht angebracht. Zum einen wird bei Ihrer Verwendung mehr Konformationen des Liganden erzeugt, die bewertet werden müssen, und zum anderen ist diese SF restriktiver bei der Bewertung, ob eine Konformation in die Bindetasche passt. Um zu verhindern, dass auf diese Weise richtige Konformationen zu schlecht bewertet werden, wird empfohlen mehrere Konformationen des Proteins zu verwenden.

EF: Der EF eignet sich gut, um die frühe Anreicherung von aktiven Liganden innerhalb einer Reihenfolge zu bewerten. Bei dem hier gegebenen Fall konnten für den EF Werte zwischen 0 und 10 erreicht werden. Somit ist, ähnlich wie bei der Bewertung durch die AUC eine gute Vergleichbarkeit der Werte untereinander gegeben. Mit einem EF von 5,42 zeigt ROCS auch im Bereich der frühen Anreicherung die beste Leistung. Mehr als die Hälfte aller Verbindungen innerhalb der ersten 5% der Ergebnisse sind aktive Liganden. GOLD ist am wenigsten dazu in der Lage bei den am besten bewerteten Lösungen zwischen aktiven und inaktiven Verbindungen zu unterscheiden. Weniger als ein Drittel der Lösungen innerhalb der ersten 5% der Ergebnisse sind aktive Liganden Die SF PMF und PSCORE zeigen im Bereich der frühen Anreicherung eine bessere Leistung als GOLDScore und GLIDEScore.

RIE: Bei der Betrachtung der Werte für RIE zeigen PMF und PSCORE die schlechteste Leistung. Bei RIE wird im Gegensatz zu den anderen verwendeten Gütemaßen noch der Score-Wert der jeweiligen Lösung mit betrachtet. Je besser eine Verbindung vom PMF oder PSCORE bewertet wurde, desto mehr liegt der Wert im negativen Bereich. Verbindungen, die nicht in die Bindetasche passen, werden mit bis zu 10.000 bewertet. Dieser große Zahlenbereich führt dazu, dass der Wert für den Divisor $\langle S \rangle$ (s. 2.6.3) ebenfalls einen großen Wert annimmt, und somit der Wert für RIE kleiner wird. Bei ROCS hingegen und dem hier verwendeten TANIMOTOCOMBOScore liegen die Werte zwischen 0 und 2 [223]. Durch den Umstand, dass in die nachträgliche Betrachtung der Dockingstudien durch ein Gütemaß noch der jeweilige Score der

Dockinglösung einbezogen wird, spricht für dieses Gütemaß[224]. Aber da es keine festen Obergrenzen für den Wert von RIE gibt, sind die Werte von verschiedenen Untersuchungen, mit unterschiedlichen SF und abweichender Anzahl von Verbindungen schwer zu vergleichen. Desweiteren ist diese Methode nur ratsam, wenn die Anzahl der getesteten Verbindungen mehr als 300 beträgt. Nur für neun Datensätze wurden einschließlich der Decoys mehr als 300 Verbindungen gefunden. Somit sind die Ergebnisse nicht völlig miteinander zu vergleichen.

CONSENSUS-SCORE: Wurde ein CONSENSUS-SCORE-Modell verwendet, hat dies im Durchschnitt nicht zu einem besseren Wert für das jeweilige Gütemaß geführt. Betrachtet man die jedoch die AUC für den CONSENSUS-SCORE, so sieht man, dass in vier Fällen ein Wert unter 0,500 erzielt wurde. Bei den übrigen SF waren es vier oder fünf Fälle, in denen die AUC unter 0,50 lag. Das heißt, dass in Fällen, mit einer niedrigen AUC sich die Verwendung eines CONSENSUS-SCORE-Modelles lohnen kann, um die Unterscheidung zwischen aktiven und inaktiven Liganden zu verbessern. Ist diese Trennung schon vorher bei der Verwendung nur einer SF gut gelungen, so trägt die Verwendung eines CONSENSUS-SCORE nicht viel zu einer Verbesserung mit bei. Für die übrigen hier verwendeten Gütemaße stellt es sich ähnlich dar. Der Wert für den CONSENSUS-SCORE war nie höher als der Wert für ROCS, aber immer oberhalb der übrigen SF.

Das eine ohnehin schon gute Anreicherung durch ein CONSENSUS-SCORE-Modell nicht weiter verbessert werden konnte, wurde auch schon gezeigt, und stellt somit keine Besonderheit in dieser Untersuchung dar [225]. Zwar konnte der Einfluß von CONSENSUS-SCORE-Modellen auf die bessere Unterscheidung von aktiven und inaktiven Verbindungen schon gezeigt werden (s. [140, 226]), jedoch wurden die inaktiven Verbindungen danach ausgewählt, ob diese überhaupt als mögliche Wirkstoffe in Frage kommen, und nicht noch danach, wie ähnlich die inaktiven Verbindungen den aktiven Verbindungen sind.

Zusammenfassend lässt sich sagen, dass keine der hier getesteten Methoden in allen Fällen zwischen aktiven und inaktiven Verbindungen unterscheiden konnte. Ebenso hat keine Variante immer falsch unterschieden. Betrachtet man alle getesteten Protein-Ligand-Komplexe, so ist die durchschnittliche Unterscheidung von aktiven und inaktiven Liganden besser als eine zufällige Verteilung. ROCS zeigt bei allen betrachteten Gütemaßen die beste Unterscheidung. Dies ist darauf zurückzuführen, dass ROCS bei der Bewertung der Substanzen nur auf die Oberfläche und Form der Liganden achtet, und nicht auf die Wechselwirkungen mit dem Protein und mit eventuell noch vorhandenem Wasser. Als Vorlage für die Oberfläche dienen schon kristallisierte Liganden, und somit trainiert man die Vorlage schon in die richtige Richtung. Den Einfluß des Proteins hätte man noch durch die Verwendung eines Pharmacophor-Modelles testen können. Die getesteten Docking-Programme erstellen selbst eine Vielzahl von Konformationen der Liganden, und bewerten diese dann im Protein. Sowohl GOLD als auch GLIDE sind in der Lage verschiedene Constraints zu verwenden, um eine bessere Lage des Liganden in der Bindetasche zu erhalten. Als Beispiele für Constraints sind Wasserstoffbrückenbindungen zum Protein oder die Lage von hydrophoben Bereichen zu nennen, die eine Konformation erfüllen muss, um weiter untersucht oder bewertet zu werden.

Diese Möglichkeit wurde in dieser Studie nicht verwendet, da PARADOCKS bisher nicht über diese Technik verfügt. GOLD und GLIDE sollten keinen besonderen Vorteil haben. Allerdings ist es sehr wahrscheinlich, dass sich durch diese Constraints die Genauigkeit der Programme verbessert hätten.

Das Neubewerten der Lösungen durch andere Scoring-Funktionen hat insgesamt zu einer Verbesserung der Ergebnisse geführt, wobei die Verwendung des CHEMSCORE besser funktioniert hat als der GLIDESCORE. Die Scoring-Funktionen PMF und PS-CORE zeigen vor allem im Bereich der frühen Anreicherung eine bessere Leistung, als die Kraftfeldbasierten Methoden des GOLDScore, sowie die empirischen SF CHEMS-CORE und GLIDESCORE.

3.2. PRK1

Als Grundlage für die Untersuchungen an der PRK1 diente die *Biomol Kinase and Phosphatase inhibitor library* mit 84 Substanzen. Zu Beginn wurde mit diesen Substanzen eine biologische Testung in der Arbeitsgruppe von Prof. Jung (Freiburg) durchgeführt [227]. Bei einer eingesetzten Inhibitorkonzentration von 100 nM konnten nur die beiden Substanzen Ro-318220 und Staurosporin als Inhibitoren identifiziert werden. 200 weitere eigene Substanzen wurden zunächst getestet.

Abbildung 3.10 zeigt die vier als aktiv identifizierten PRK1 Inhibitoren. Die zugehörigen IC_{50} -Werte sind in Tabelle 3.21 aufgeführt.

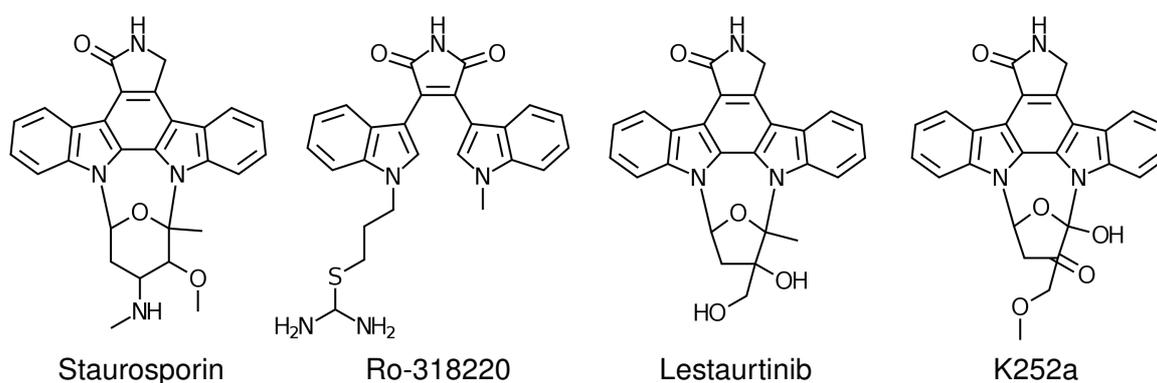


Abbildung 3.10: Aktive Liganden der PRK1

3.2.1 Dockingstudien

Es wurden Dockingstudien mit den Substanzen der Biomol-Datenbank und den als aktiv erkannten Substanzen durchgeführt. Von der PRK1 gibt es keine Kristallstruktur, jedoch ist die Sequenz bekannt [228]. Deshalb konnte auf Grundlage der Sequenz ein Homologiemodell für die folgenden Dockingstudien erstellt werden. Als Vorlage für die Modellierung wurde wegen der hohen Ähnlichkeit (Identität = 49%) die Protein Kinase C Theta (*PKC- θ* , PDB 2JED) ausgewählt. Die Modellierung selber erfolgte über SWISS-MODEL [229–231]. Das erhaltene Modell wurde mittels *prosa* und *procheck* auf Konsistenz geprüft. Bei dem Modell liegt das DFG-Motiv in seiner inaktiven Konformation vor.

Die Untersuchungen wurde mit den 84 Substanzen der Biomol-Datenbank, sowie mit den oben erwähnten ausgewählten Kinaseinhibitoren durchgeführt. Das Zentrum des Suchraums lag bei Glutamat (GLU696) der Hinge-Region und der Radius betrug 15 Å. Hierzu wurden die vier Programme GOLD, GLIDE, PARADOCKS und DOCK, sowie fünf verschiedene SF (GOLDScore, CHEMScore, GLIDEScore, PScore und AMBERScore) verwendet. Der AMBERScore und der CHEMScore wurden hier nur zum Rescoring der von GOLD und GOLDScore erhaltenen Ergebnisse verwendet.

63 Substanzen aus der Biomol-Datenbank und die sieben zusätzlich ausgewählten Substanzen konnten mit den passenden Interaktionen in das Homologiemodell gedockt werden. Tabelle 3.21 zeigt die normalisierten Bewertungen für jede SF und den CONSENSUS-SCORE. Der jeweils höchste Score wurde hierbei auf 10 festgelegt. Es ist zu sehen, dass die einzelnen SF eine ähnliche Rangfolge produzieren, auch wenn es einige wenige Ausnahmen gibt. Es fällt auf, dass Staurosporin und Ro-318220 von allen SF, bis auf PSCORE am besten bewertet werden. PSCORE bewertet die Substanz 84 am besten.

Tabelle 3.21: Dockinglösungen an der PRK1 sortiert nach dem CONSENSUS-SCORE. Die aktiven Substanzen sind markiert. Scoringwerte sind normalisiert dargestellt.

Substanz	CHEM- SCORE	GOLD- SCORE	AMBER- SCORE	GLIDE- SCORE	P- SCORE	CONSENSUS- SCORE	IC ₅₀ [nM]
Ro-318820	10,00	10,00	8,30	10,00	9,27	9,51	28,6 ±6,5
Staurosporin	9,76	7,14	10,00	9,27	9,98	9,23	0,8±0,2
31	9,93	8,13	7,45	9,56	9,14	8,84	
K252A	7,34	8,68	6,93	7,27	9,91	8,03	3,2±0,6
Lestaurtinib	7,67	8,08	6,51	7,45	9,69	7,88	8,6±0,9
SB216763	8,74	7,24	6,11	8,28	7,62	7,60	
34	7,02	7,00	7,06	6,38	7,25	6,94	
48	6,58	7,80	6,22	5,62	7,83	6,81	
35	6,67	7,46	5,53	5,86	7,84	6,67	
58	5,63	6,11	4,96	6,46	9,80	6,59	
63	6,14	5,45	6,18	6,33	8,63	6,54	
Lapatinib	6,50	6,79	6,70	3,19	9,30	6,50	
41	5,61	4,91	5,66	6,55	9,51	6,45	
9	7,79	6,24	3,94	5,39	8,50	6,37	
8	7,53	3,86	5,42	5,87	8,16	6,17	
69	7,30	5,66	4,58	6,67	6,57	6,16	
45	5,81	7,38	5,30	5,01	7,15	6,13	
62	5,61	6,64	5,28	6,38	7,14	6,12	
22	7,47	5,50	4,51	6,42	6,63	6,11	
84	4,46	7,39	3,52	5,16	10,00	6,10	
Erlotinib	5,91	6,06	5,89	4,53	8,04	6,08	
68	7,83	4,89	4,35	7,85	5,49	6,08	
36	5,39	6,45	5,76	7,23	5,58	6,08	
2	6,23	7,16	5,06	3,97	7,83	6,05	
52	5,88	5,84	5,50	6,46	6,35	6,01	
Vatalanib	6,14	4,96	6,79	4,70	7,41	6,00	
35	7,16	6,87	7,78	0,00	8,11	5,98	
47	5,64	6,96	5,43	5,25	6,63	5,98	

Fortsetzung nächste Seite ...

Substanz	CHEM- SCORE	GOLD- SCORE	AMBER- SCORE	GLIDE- SCORE	P- SCORE	CONSENSUS- SCORE	IC ₅₀ [nM]
24	5,75	5,88	5,32	6,45	6,48	5,98	
38	4,59	5,62	6,90	6,48	6,19	5,96	
56	4,97	6,26	6,63	5,56	6,30	5,94	
76	6,50	6,70	5,78	2,35	8,10	5,89	
51	5,29	5,53	5,21	5,26	7,46	5,75	
46	5,00	7,13	5,11	4,95	6,51	5,74	
23	5,49	5,45	4,83	6,62	6,22	5,72	
32	7,16	6,48	6,40	0,00	8,49	5,71	
42	5,55	5,10	6,20	5,24	6,44	5,71	
39	6,49	6,45	3,30	5,60	6,53	5,67	
37	4,36	6,33	4,52	6,84	6,17	5,64	
3	5,39	4,45	4,41	5,30	8,16	5,54	
18	6,10	4,63	4,88	5,54	6,50	5,53	
25	6,04	5,74	2,76	6,13	6,93	5,52	
SB203580	5,67	4,63	5,14	4,07	7,97	5,50	
4	4,65	4,71	4,94	6,95	6,16	5,48	
5	5,02	5,00	4,57	7,39	5,09	5,41	
1	4,94	4,44	4,60	6,56	6,40	5,39	
77	3,88	7,61	4,18	4,65	6,58	5,38	
67	4,39	5,98	3,96	4,91	7,14	5,28	
28	5,94	4,37	4,56	4,68	6,81	5,27	
43	4,36	4,40	5,04	6,01	6,47	5,25	
60	3,78	4,04	3,53	8,39	6,29	5,21	
54	4,54	3,64	4,14	7,27	6,35	5,19	
66	5,09	5,13	2,83	6,30	6,53	5,18	
14	5,53	5,42	3,58	6,11	5,00	5,13	
70	6,90	7,30	4,70	0,00	6,72	5,12	
49	3,56	5,88	3,12	6,40	6,54	5,10	
15	4,00	6,14	4,50	4,67	6,18	5,10	
53	4,82	3,72	4,13	4,66	8,02	5,07	
27	4,10	5,84	4,32	4,83	6,21	5,06	
13	5,83	4,74	3,99	5,29	4,98	4,97	
26	5,07	4,52	3,55	4,99	5,14	4,65	
21	5,42	5,72	0,90	5,62	5,20	4,59	
17	5,32	4,25	3,70	4,03	5,44	4,55	
61	4,47	5,27	3,00	4,92	5,14	4,51	
83	3,11	7,74	2,80	4,50	4,32	4,49	
20	3,60	2,80	4,52	4,54	6,49	4,39	
11	4,70	4,39	2,83	4,70	4,56	4,24	
12	4,14	4,32	2,53	5,07	4,95	4,20	
44	3,35	4,30	2,58	6,53	3,35	4,02	
50	3,68	4,15	3,27	0,00	5,21	3,26	

Abbildung 3.11 zeigt Dockinglösungen der in dem Assay als aktiv erkannten Substanzen. Bei Staurosporin und Ro-318220 bilden die Lactam- und die Maleimidgruppe jeweils eine Wasserstoffbrückenbindungen zur den Aminosäuren GLU696 und SER698 der Hinge-Region aus. Zusätzlich bildet die protonierte Aminogruppe des Staurosporins und die Thioharnstoffgruppe von Ro-318220 eine Wasserstoffbrückenbindung zum ASP744 in der Aktivierungsschleife aus. Wie erwartet zeigen Lestaurtinib und K252a die gleichen Interaktionen mit der Hinge-Region wie Staurosporine oder Ro-318220.

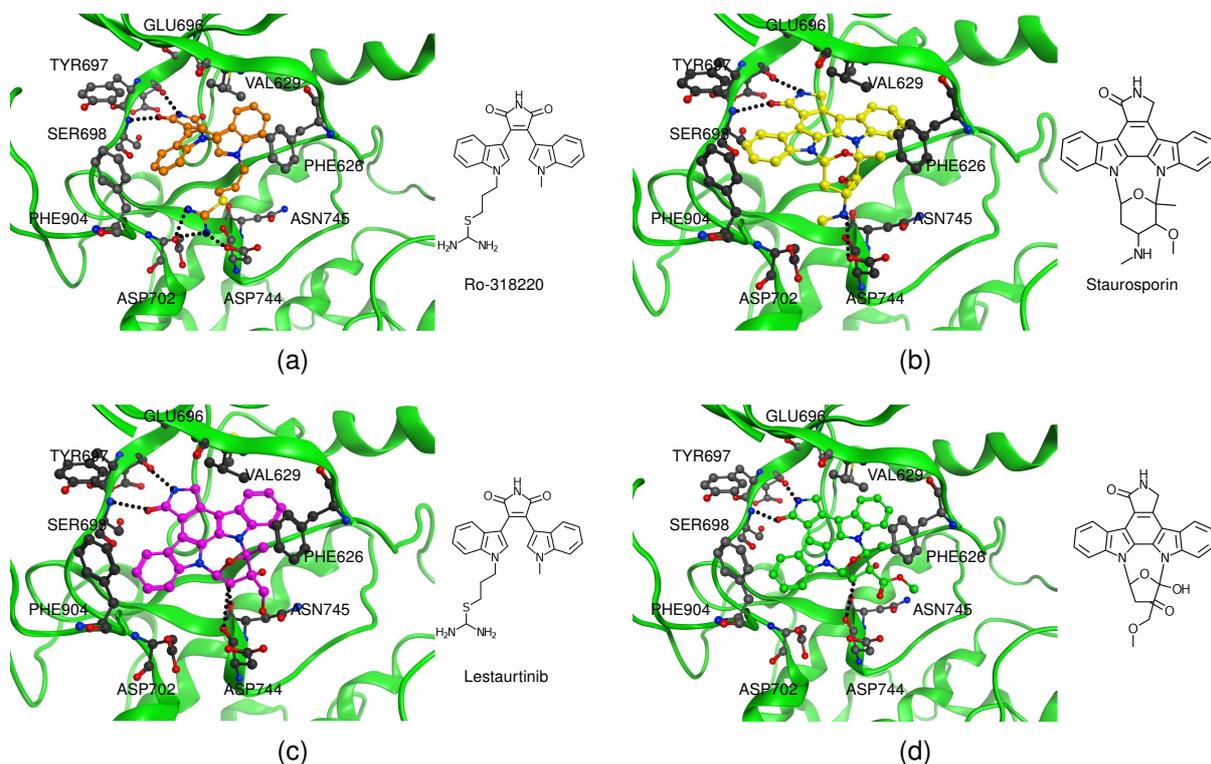


Abbildung 3.11: Dockinglösungen in PRK1

a) Ro-318220 (orange) b) Staurosporin (gelb) c) Lestaurtinib (magenta) d) K252 (grün); Rezeptoratome in schwarz.

Die Abbildung 3.12 zeigt Lösungen von vier als inaktiv ermittelten Substanzen. 31 wurde von allen SF als gut bewertet, und steht bei der Bewertung durch den CONSENSUS-SCORE an dritter Stelle. Bei diesem Stoff handelt es sich um ein Bisindolylmaleimiddesivat, und bildet zusätzlich auch van-der-Waals-Wechselwirkungen mit den Seitenketten von PHE626 und PHE904 aus. Dies zeigt sich in den hohen Bewertungen dieser Substanz bei allen SF.

Substanz 84, zu sehen in Abbildung 3.12(b) fällt besonders auf. Bei dem Docking mit PARADOCKS und Bewertung durch den PSCORE wird diese als Beste bewertet. Alle anderen Programme und SF bewerten diese nur mittelmäßig. Durch das gute abschneiden bei PARADOCKS erhält sie beim CONSENSUS-SCORE noch einen Platz im oberen drittel der Rangfolge. Es ist auch zu sehen, dass nur Wasserstoffbrücken zur Hinge-Region ausgebildet werden, und es sonst keine Interaktionen mit der ATP-

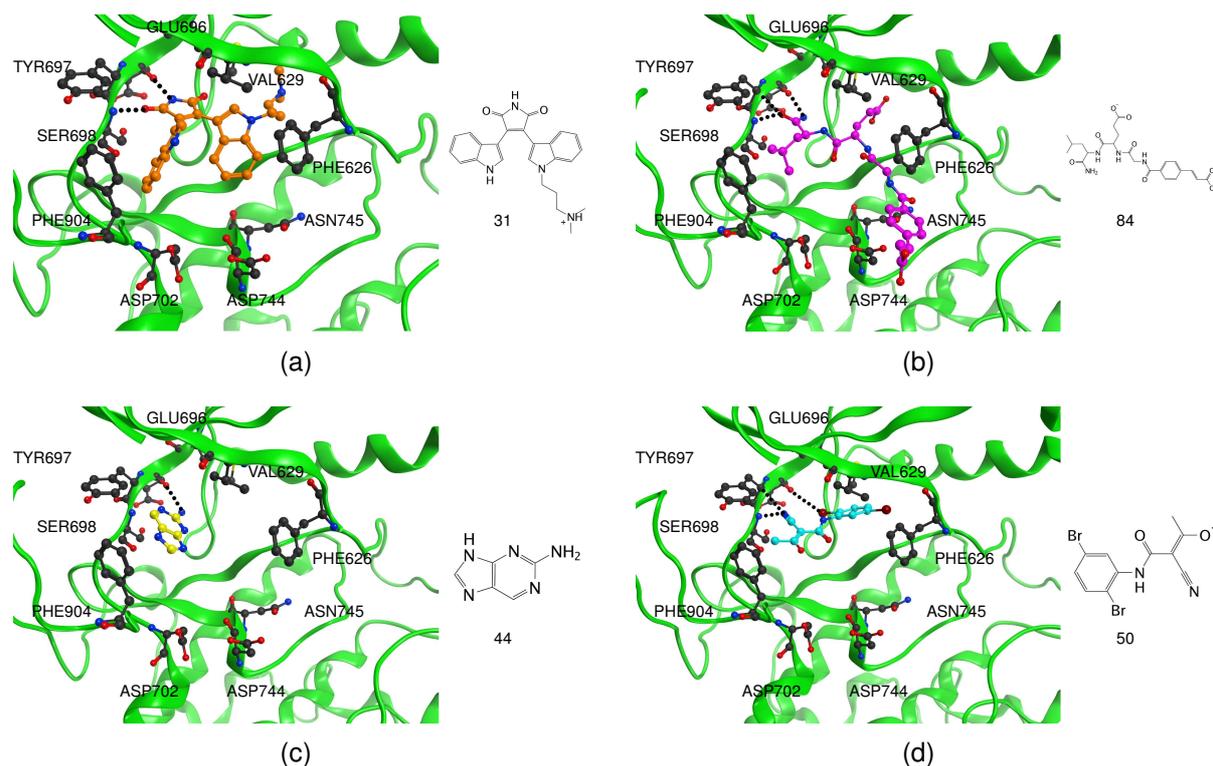


Abbildung 3.12: Dockinglösungen in PRK1

a) 31 (orange); b) 84 (magenta); c) 44 (gelb); d) 50 (magenta); Rezeptoratome in schwarz.

Bindetasche gibt.

In Abbildung 3.12(c) ist Substanz 44 dargestellt, die den vorletzten Platz in der Rangfolge einnimmt. Die Aminogruppe bildet nur eine Wasserstoffbrücke zum GLU696 aus. In diesem Docking wurde die Substanz 50 am schlechtesten bewertet. Allein der PS-CORE hat diese noch so bewertet, dass sie in der oberen Hälfte aller getesteten Substanzen lag. Der Cyanstickstoff bildet zwei Wasserstoffbrücken aus. Jeweils eine zu SER698 und TYR697.

3.2.2 Diskussion

Identifizierte aktive Substanzen: Es sollte ein neuer möglicher PRK1-Inhibitor gefunden werden. Die biologische Testung ergab, dass von den getesteten Substanzen nur Bisindolymaleimidderivate in der Lage waren die PRK1 zu hemmen. Neben den schon bekannten Inhibitoren Ro-318220 [92] und Staurosporin [232] konnten Lestaurtinib und K252a als potentielle Inhibitoren identifiziert werden. Lestaurtinib ist seit 2009 in der klinischen Phase-II für akute myelotische Leukämie sowie für myeloproliferative Erkrankungen.

K252a ist ebenfalls schon als selektiver Inhibitor für Tyrosinkinase aus der Familie der Tropomyosin-Rezeptor-Kinasen (TRK) bekannt [233, 234].

Homologie Modell: In dieser Untersuchung wurde ein Homologie Modell der PRK1-Kinase verwendet. Es hat sich gezeigt, dass dies allen Anforderungen entsprach und dazu geführt hat, dass in den Dockingstudien alle aktiven Substanzen im oberen Bereich der Rangfolge waren. Die Prüfung mittel prosa und procheck hat ergeben, dass die Qualität des Modells ausreichend ist, um mit diesem weitere Untersuchungen vorzunehmen.

Dockingstudien: Nach dem Docking zeigten nur 63 der 284 hier getesteten Substanzen die erwarteten Wasserstoffbrückenbindungen zur Hinge-Region auf. Alle verwendeten Scoring-Methoden bewerten die getesteten Verbindungen ähnlich. Es konnte gezeigt werden, dass Lestaurtinib die gleichen Interaktionen mit der Hinge-Region von PRK1 aufweist, wie die schon bekannten Inhibitoren. Alle aktiven Substanzen waren bei jeder SF innerhalb der ersten 10 % der Rangfolge zu finden. Es waren jedoch bei allen SF zwei inaktive Substanzen innerhalb dieser Gruppierung zu finden. Beide Stoffe, 31 und SB216763 sind Bisindolymaleimide. SB216763 ist jedoch schon als selektiver Glykogensynthase-Kinase 3 Inhibitor beschrieben [235, 236].

3.2.2.3.1 Unterschiede in den Scoring-Funktionen Der CHEMScore und AMBERScore wurden nur zum rescoren der Dockingergebnisse von GOLD verwendet. Somit wurde der Konformationsgenerator von DOCK nicht benutzt. Es kann also nicht gesagt werden, ob der eigene Konformationsgenerator bessere Lösungen gefunden hätte.

Auch GLIDE hat mit dem GLIDEScore – vor allem im Bereich der oberen 10 % – fast die gleiche Reihenfolge wie der CHEMScore oder GOLDScore. GLIDE verwendet jedoch einen anderen Suchalgorithmus als GOLD. Bei GLIDE werden zu Beginn die zu testenden Stoffe in Kern- und Rotamergruppen unterteilt. Die Kerngruppe bleibt starr und für die Rotamergruppen werden viele Kombinationen erzeugt. So werden weniger als 500 Konformationen pro Substanz erzeugt, die dann jeweils in verschiedenen Positionen und Orientierungen in die Bindetasche des Protein eingepasst werden. Diese Auswahl wird dann anhand verschiedener Filter (z. B. sterische Beeinflussung, Abstoßende Ladungen) verkleinert. Diese übriggebliebenen Konformationen werden darauf

mit einer eigenen Variante des CHEMScore bewertet, wobei hydrophobe Wechselwirkungen und Wasserstoffbrückenbindungen besonders favorisiert werden. Die hiernach übrigen Konformationen werden dann noch in ihrer Position minimiert. Zum Schluß werden die besten Konformationen nochmals mit dem GLIDEScore bewertet, der seinerseits auch auf dem CHEMScore beruht [110]. Dies erklärt weshalb sich die Bewertung der Liganden durch den CHEMScore und GLIDEScore so sehr gleichen.

Bei der Bewertung durch den PScore kommt es zu einer Besonderheit: Substanz 84 wird am besten bewertet. Eine Möglichkeit ist, dass von der SF die Wechselwirkung der in das Solvent ragende Carboxylgruppe im Verhältnis zu den anderen Wechselwirkungen überbewertet wird. Und die Position der anderen Carboxylgruppe in der hydrophoben Tasche II wird vielleicht mit einem zu kleinen Strafterm belegt.

3.3. MYT1-Kinase

Es wurden verschiedene neue mögliche Inhibitoren (s. Abbildung 3.13) für die MYT1 im Rahmen der Diplomarbeit von Benjamin Sauer synthetisiert. Grundlage für die Auswahl dieser Substanzen war, dass mittels Click-Chemie [237] ein hydrophober- und ein Zuckerrest miteinander verbunden werden kann, um so eine große Strukturvielfalt zu erhalten. Der hydrophobe Teil sollte dabei mit der Hinge-Region interagieren, und der Zuckerrest Richtung Aktivierungsschleife zeigen. Als Ausgangssubstanz wurde das 6-Azido-6-desoxy- α -methylglucopyranosid verwendet, welches mit verschiedenen Alkylkomponenten zu entsprechenden Triazolen umgesetzt wurde. Bevor die Substanzen in einem LanthaScreen™ Assay (s. 2.10.2) getestet wurden, erfolgte die Testung mittels verschiedener Docking-Methoden, ob und welche dieser Verbindungen als potentielle Inhibitoren in Frage kommen [238]. Sowohl die Dockingstudien, als auch die Moleküldynamik-Simulationen wurden an der Kristallstruktur der aktiven Konformation der MYT1 (PDB 3P1A) durchgeführt, und die Struktur wurde wie in Kapitel 2.8 beschrieben vorbereitet.

3.3.1 Docking

Die durchgeführten Dockingstudien mit GOLD und GLIDE hatten die gleichen Einstellungen für die Untersuchungen. Dazu wurde das Zentrum des Suchraums auf das

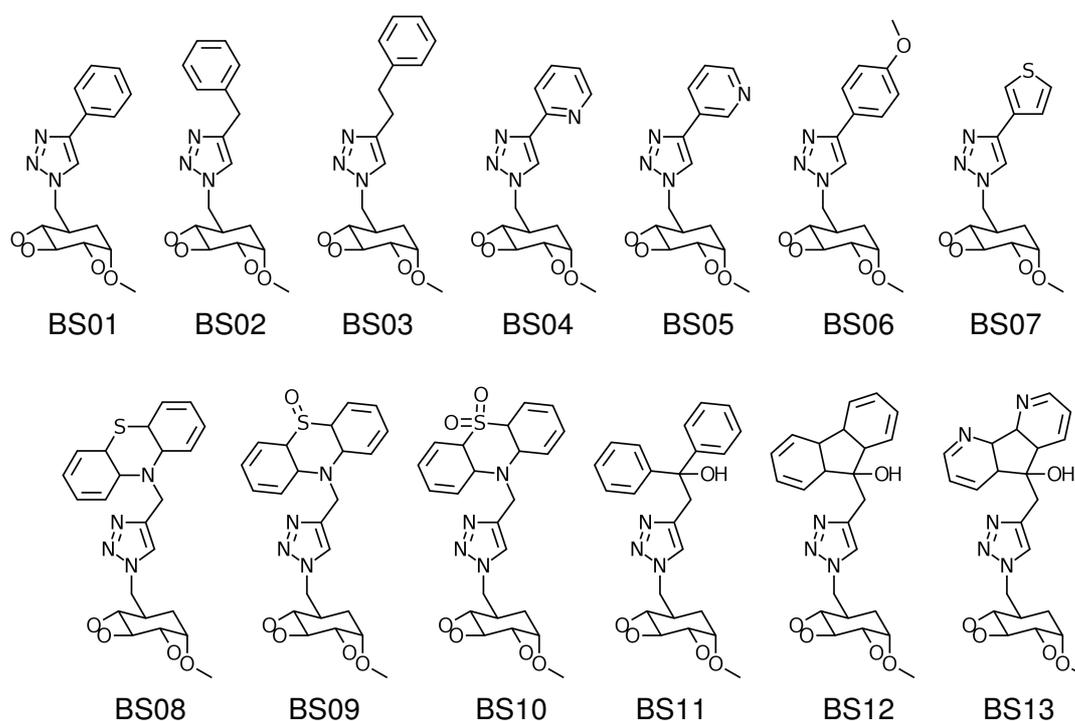


Abbildung 3.13: Design von Inhibitoren der MYT1
Diese Moleküle wurden mit verschiedenen Docking-Programmen getestet.

Tabelle 3.22: Ergebnisse der Dockings mit GOLD und GLIDE

Die lt. erzieltem Score beste Lösung ist hervorgehoben. Zusätzlich ist die Summe der CONSENSUS-SCORES angegeben. Der CONSENSUS-SCORE wurde anhand des jeweiligen Top-Scores berechnet, wobei der höchste CONSENSUS-SCORE auf 10.00 festgesetzt wurde.

Name	GOLDScore	GLIDEScore	Σ CONSENSUS-SCORE
BS01	53,75	-7,89	16,14
BS02	57,25	-7,37	16,05
BS03	55,98	-7,44	15,95
BS04	54,82	-7,42	15,77
BS05	53,12	-7,52	15,65
BS06	52,23	-7,20	15,17
BS07	55,70	-7,32	15,78
BS08	54,98	-6,90	15,22
BS09	57,76	-6,23	14,86
BS10	59,46	-8,68	17,80
BS11	60,90	-9,08	18,44
BS12	72,12	-8,04	18,86
BS13	62,02	-7,93	17,33
Staurosporine	53,37	-7,78	15,97

CYS190 in der Hinge-Region gesetzt und der Radius betrug 15 Å. Zusätzlich wurde ein *Constraint* zur Hinge-Region gesetzt, sodass eine Wasserstoffbrücke zwischen Substanz und Hinge-Region bestehen muss. Die Ergebnisse sind in Tabelle 3.22 aufgeführt.

Es wurde zudem geprüft, ob die gefundenen Lösungen auch das bekannte Bindungsmuster zur Hinge-Region aufweisen (s. 1.9). Dabei fiel auf, dass für *BS12*, obwohl diese einen guten GOLDScore erhielt, keine sinnvolle Docking-Lösung gefunden wurde. Es wurde keine Wasserstoffbrückenbindung mit der Hinge-Region gefunden, obwohl dies als Bedingung gesetzt war. Zudem zeigt der Zuckerrest in die hydrophobe Tasche, und nicht in Richtung der Aktivierungsschleife (s. Abbildung 3.14(a)). Das gleiche gilt für die am besten bewertete Pose von GLIDE (*BS11*). Auch hier sagt der Score alleine nichts über die Güte der Lösung aus, und der Ligand liegt in einer falschen Orientierung vor.

Nur für die Strukturen *BS09*, *BS10* und *BS13* liegen Dockingergebnisse vor, die eine sinnvolle Orientierung zeigen (s. Abbildung 3.14). Es gibt die Wasserstoffbrückenbinden zur Hinge-Region, und der Zuckerrest zeigt in Richtung der Aktivierungsschleife. Bei den übrigen Verbindungen liegt wie schon bei *BS12* der Zuckerrest immer in der hydrophoben Tasche, und es ist keine Interaktion mit der Hinge-Region zu sehen. Deshalb schieden diese Substanzen für die weiteren Betrachtungen aus. Aufgrund der Posen, und nicht aufgrund der erzielten Scores wurden die Strukturen *BS09*, *BS10* und *BS13* für die weitere Untersuchung ausgewählt.

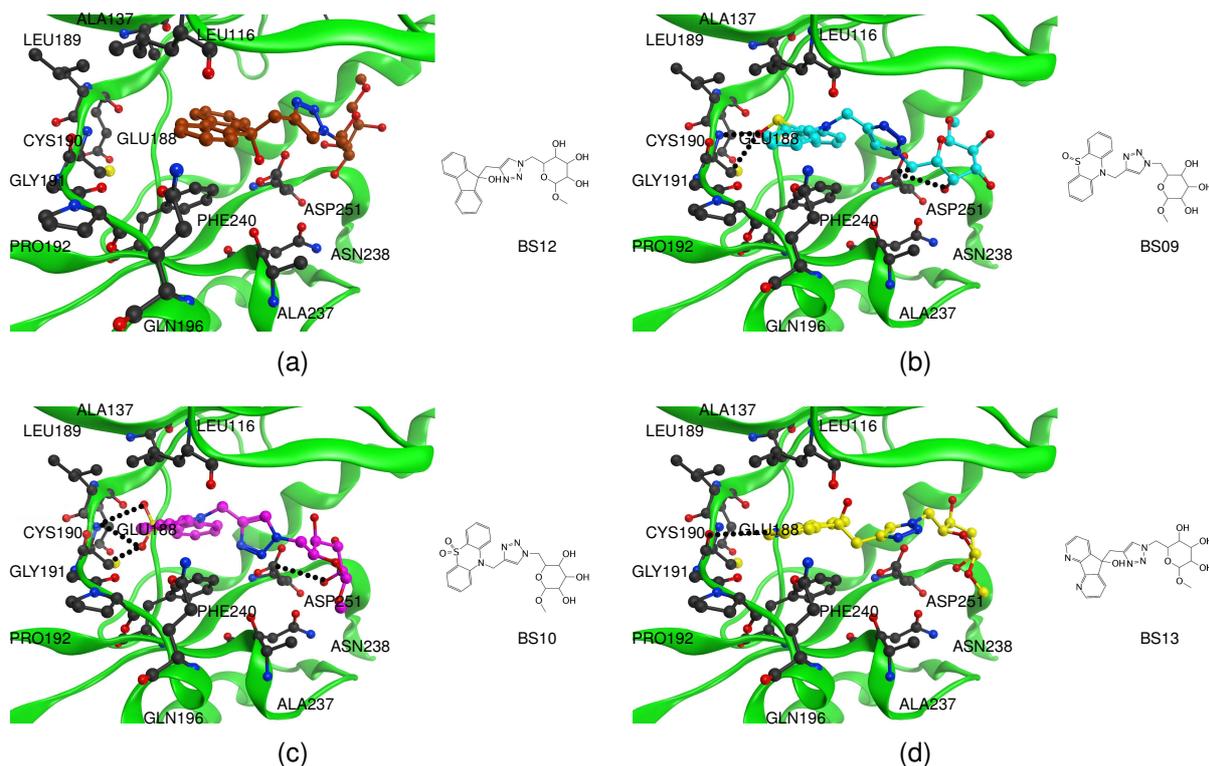


Abbildung 3.14: Dockinglösungen in MYT1

a) *BS12* (braun) hat keine Interaktion mit der Hinge-Region, und der Zuckerrest ragt in die falsche Tasche. b) *BS09* (cyan) GOLD hat diesen Liganden als durchschnittlich bewertet, GLIDE als schlecht. Bei Betrachtung des Consensus-Scores wird dieser Ligand als der Schlechteste angesehen. Die Orientierung stimmt jedoch, und es gibt Interaktionen zur Hinge-Region. c) *BS10* (magenta) gehört zu den besser bewerteten Strukturen, und weist zudem noch die passenden Interaktionsmuster auf. d) *BS13* (gelb) ebenfalls gut bewertet, und zeigt die richtigen Interaktionen zur Hinge-Region; Rezeptoratome in schwarz.

3.3.2 Moleküldynamik-Simulation an der MYT1

Da die in Kapitel 3.3.1 gezeigten Ergebnisse von einem starren System ausgehen, und somit die Wirklichkeit nur bedingt wiedergeben, und zudem keine Aussage über die Stabilität der Komplexe machen können, wurde mit den Liganden, welche Interaktionsmuster mit der Hinge-Region aufweisen, Moleküldynamik-Simulationen durchgeführt, und die Berechnungen für die freie Energie mit expliziten und impliziten Solvent gemacht.

Abbildung 3.15 zeigt die RMSD-Plots der untersuchten Verbindungen, und Tabelle 3.23 die daraus errechneten Energien. Die Ergebnisse der MD und des Dockings korrelieren miteinander. Während der gesamten Simulation liegt *BS09* nicht stabil in der Bindetasche, sondern verändert über den gesamten Verlauf der Simulation seine Orientierung. Dies wird auch durch die positiven ΔG -Werte deutlich, die man sowohl bei der MM-PBSA als auch bei der MM-GBSA-Berechnung bekommt. Ein positiver Wert bei der freien Energie deutet auf eine instabile Bindung hin. *BS10* zeigt ein ganz ande-

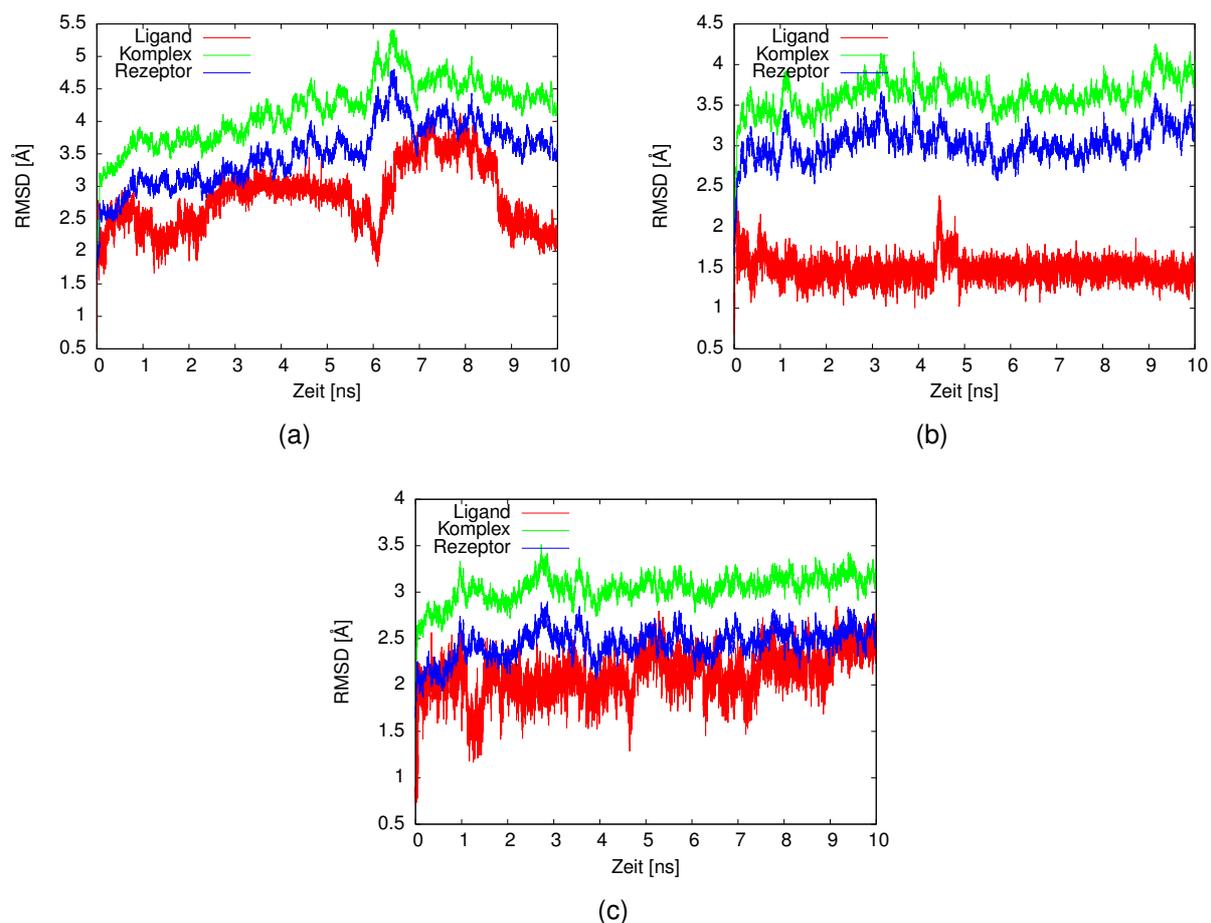


Abbildung 3.15: RMSD-Plots der ausgewählten Komplexe
 a) *BS09* b) *BS10* c) *BS13* jeweils für 10 ns.

res Bild. Der Ligand bleibt sehr stabil, und bewegt sich nur in einem kleinen Rahmen, wie die RMSD-Werte um 1,5 Å anzeigen. Zwischen 4 und 5 ns gibt es zwar eine Lageänderung, aber diese beträgt auch nur 0,5 Å verglichen mit dem Rest. Hier erhielten wir bei beiden Energieberechnungen auch einen negativen Wert. Ebenso war die Lage des Rezeptors stabil und hat über den gesamten Verlauf der Simulation RMSD-Werte von unter 4 Å (s. Abbildung 3.15 b). *BS13* ist anhand der Energien nicht eindeutig. Es hat einen positiven Wert bei der MM-PBSA Berechnung, aber einen deutlich negativeren als *BS10* bei der MM-GBSA Berechnung. Der RMSD-Plot zeigt jedoch keine ganz stabile Position für den Liganden an, aber dennoch ist der Rezeptor sehr stabil.

Tabelle 3.24 zeigt die Ergebnisse aus der biologischen Testung. Es fällt auf, dass bis auf die bekannten Inhibitoren Dasatinib, Tyrphostin AG 1478 und PD0166285 keine andere getestete Substanz die Kinase inhibiert. PD0166285 vermag dies auch nur bei einer eingesetzten Konzentration vom 10 μM, wobei die Kinase aber zu 100 % gehemmt wird. Die IC_{50} liegt bei 7,1 nM. Keine der in den Dockingstudien und Moleküldynamik-

Tabelle 3.23: ΔG -Werte von BS09, BS10 und BS13 in [kcal/mol], SD - Standardabweichung

MM-PBSA						
	BS09		BS10		BS13	
	\emptyset	SD	\emptyset	SD	\emptyset	SD
PBTOT	-17,51	4,69	-29,97	2,71	-20,33	4,11
TSTOT	-25,77	24,43	-23,33	24,64	-21,38	25,33
freie Energie	8,26		-6,64		1,05	
MM-GBSA						
	ΔG	BS09	ΔG	BS10	Δ	BS13
	\emptyset	SD	\emptyset	SD	\emptyset	SD
GBTOT	-22.01	2.34	-29,66	2,53	29,60	3,19
TSTOT	-25.77	24.43	-23,33	24,64	-21,38	25,33
freie Energie	3.76		-6,64		-8,22	

Tabelle 3.24: Ergebnisse der biologischen Testung nach einem LanthaScreen Binding Assay ohne Glycoglycerolipide n. i.: Keine Inhibierung (< 10%) IC_{50} basieren auf zwei Werten in einem linearen Bereich

Substanz	Inhibition@5 μ M	Inhibition@10 μ M	IC_{50}
BS01	n. i.	n. i.	-
BS02	n. i.	n. i.	-
BS03	n. i.	n. i.	-
BS04	n. i.	n. i.	-
BS05	n. i.	n. i.	-
BS06	n. i.	n. i.	-
BS07	n. i.	n. i.	-
BS08	n. i.	n. i.	-
BS09	n. i.	n. i.	-
BS10	n. i.	n. i.	-
BS11	n. i.	n. i.	-
BS12	n. i.	n. i.	-
BS13	n. i.	n. i.	-
Dasatinib	96%	97%	63 nM
Tyrphostin AG 1478	25%	39%	16 μ M
Staurosporine	n. i.	n. i.	-
PD0166285	-	100%	7,1 nM

Simulationen als potentielle aktive bestimmte Substanz zeigte in dem Assay eine inhibierende Wirkung.

3.3.3 Diskussion

Aufgrund der Dockingergebnisse und den anschließenden MDs wurde davon ausgegangen, dass die Substanzen *BS310* und *BS313* potentielle MYT1-Inhibitoren darstellen. Die experimentellen Testung zeigte jedoch keine Inhibition der MYT1.

Allerdings darf nicht vergessen werden, dass ohnehin für nur drei der zu testenden Substanzen überhaupt möglich war eine geeignete Dockinglösung zu finden, die dem bekannten Bindungsmodus entsprach. Die berechneten freien Energien sind ebenfalls nicht eindeutig. Nur für *BS10* wurde überhaupt eine negative freie Energie sowohl mittels MM-GBSA und MM-PBSA berechnet. Die erhaltenen Werte um 6 kcal/mol liegen eher am Rand der Aktivität. Werte ab etwa 10 kcal/mol deuten mit einer höheren Wahrscheinlichkeit auf Aktivität hin [239, 183]. Stattdessen führte keine vorhergesagte Substanz zu einer Inhibition der MYT1. Drei weitere als Kinaseinhibitoren bekannte Substanzen waren in der Lage die Kinase zu hemmen.

In mehreren Tests konnte gezeigt werden, dass Staurosporin unspezifisch viele Kinasefamilien schwach hemmt [46, 240, 241]. 2011 wurde gezeigt, dass Staurosporin die MYT1 nicht hemmt, aber auf Kinasen aus der AGC-Gruppe hemmend wirkt [238, 242]. Insofern ist das Ergebnis für Staurosporin nicht überraschend.

Dasatinib ist seit 2006 auf dem Markt und wird in Deutschland bei chronischer myeloischer Leukämie und bei Philadelphia-Chromosom-positiver akuter lymphatischer Leukämie eingesetzt, jedoch bisher nur bei Patienten die eine Imatinib-Resistenz oder -Unverträglichkeit aufweisen. 2007 konnte in einem Test mit *kinobeads* (Kinaseinhibitor beads) gezeigt werden, dass Dasatinib an weitere Kinasen bindet, darunter MYT1. Wobei der IC_{50} -Wert an dieser im einstelligen μM -Bereich liegt [243], und schon 2004 wurde in einem Tumorzellassay gezeigt, dass 14 von 15 klinisch relevanten Kinasen in einen niedrigen nanomolaren Bereich gehemmt werden [244]. Deswegen ist es nicht überraschend, dass Dasatinib im Assay ebenfalls aktiv ist, wobei hier eine höhere Aktivität gemessen wurde. Abbildung 3.16 zeigt die Kristallstruktur von Dasatinib in ABL, sowie die Dockinglösung von Dasatinib in der MYT1. In beiden Abbildungen sieht man den gleichen Bindungsmodus, was die inhibitorischen Eigenschaften zusätzlich bestätigt.

PD0166285 ist ebenfalls als Kinasehemmer bekannt. 2001 und 2011 konnte in einem Zellassay gezeigt werden, dass die WEE-1 und MYT1-Kinase durch PD0166285 im nanomolaren Bereich gehemmt werden [246, 247]. PD0166285 inhibiert nicht nur diese beiden Kinasen, sondern viele weitere Tyrosinkinasen [248].

Tyrphostin AG1478 besteht aus einem Quinazolingrundgerüst, und ist ein spezifischer Inhibitor für die EGFR-Kinase [249, 250]. Der Bindungsmodus von Quinazolinderivaten an Kinasen wurde 1999 aufgeklärt [251]. In Abbildung 3.17 sind die Bindungsmodi von Tyrphostin AG1478 und einem Quinazolin gezeigt. Das Quinazolingrundgerüst, mit den beiden Methoxygruppen, liegt in beiden Kinasen an der gleichen Stelle, lediglich der Benzolring liegt um etwa 90° gedreht vor. Aufgrund der Ähnlichkeit zwischen EGFR und der MYT1-Kinase ist es verständlich, weshalb Tyrphostin AG1478 ebenfalls die MYT1-Kinase hemmt, jedoch mit deutlich schwächeren Werten im μM Bereich.

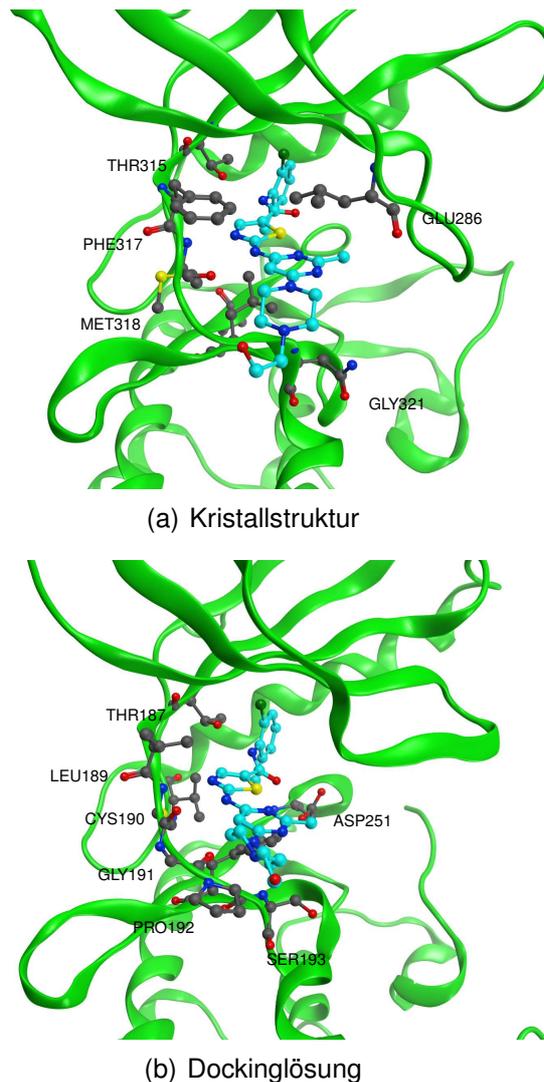
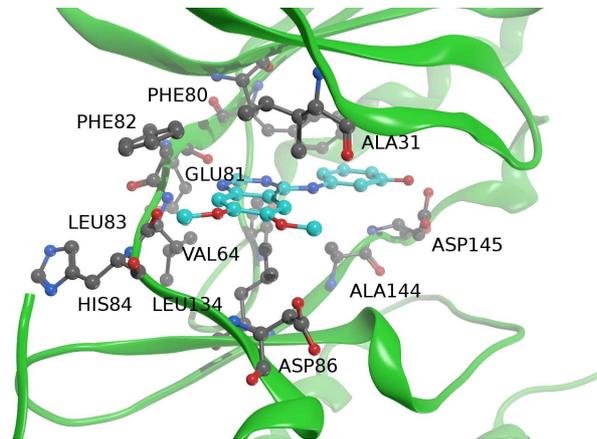


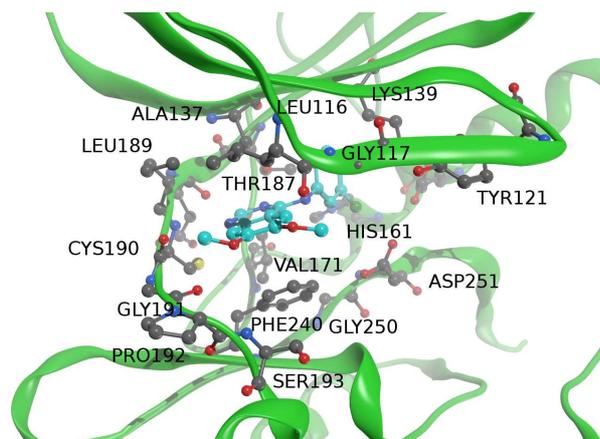
Abbildung 3.16: Lage von Dasatinib

Dasatinib (cyan); a) PDB 2GQG b) Dasatinib in PDB 3P1A; Rezeptoratome schwarz; die bezeichneten Aminosäuren stammen aus [245].

Mit dem hier verwendeten Assay werden sowohl ATP-Kompetitive als auch allosterisch hemmende Substanzen identifiziert. Jedoch kann der Test nicht unterscheiden, ob der Ligand an die aktive oder die inaktive Konformation der Kinase bindet. Zudem kann nicht gesagt werden, ob eine Substanz ATP-Kompetitiv oder allosterisch hemmt. Da nur drei synthetisierte Substanzen in den Dockinstudien einen Bindungsmodus gezeigt haben, der dem von ATP entspricht, diese aber nicht aktiv sind, kann dies auch darauf hindeuten, dass die Stoffklasse – wenn überhaupt – nur eine deutlich schwächere Affinität zu den Kinasen aufweist als ATP, und somit bei der Testung keine Aktivität zeigt. Weitere Modifikationen der aromatischen Reste, die in der Nähe der Hinge-Region liegen, sind notwendig, um zu aktiven MYT1-Inhibitoren aus der Klasse der Glyko-Triazole zu gelangen.



(a) Kristallstruktur



(b) Dockinglösung

Abbildung 3.17: Lage von Tyrphostin AG1478

a) Bindung von einem Quinazolin in die CDK2 (PDB 1DI8); b) Gedockt in die MYT1; Tyrphostin jeweils cyan; Rezeptoratome schwarz.

3.4. VEGFR, EGFR, IGFR

In der Arbeitsgruppe von PD Dr. Hilgeroth, MLU Halle-Wittenberg, wurden unterschiedliche mögliche 4-Benzylamino-substituierte Inhibitoren für VEGFR, EGFR und IGFR synthetisiert. Als Grundgerüst wurden dafür zum einen 1-aza-9-oxafluorene, und zum anderen α -carboline verwendet. (s. Abbildung 3.18). Um die Wirksamkeit dieser Verbindungen zu testen wurde ein Kinase Affinitäts Assay durchgeführt. [252, 197]. Bis auf die Substanz 11g zeigten alle Verbindungen bei wenigstens einer getesteten Kinase inhibitorische Aktivität im niedrigen mikro- bis hin zum nanomolaren Bereich (s. Tabelle 3.25). Vor allem die Wirksamkeit der Klasse der 1-aza-9-oxafluorene konnte schon mehrfach an verschiedenen Kinasen gezeigt werden [253, 254]. Nach den Dockinguntersuchungen wurden verschiedene MD verwendet, um zu testen, ob es möglich ist, anhand der errechneten freien Bindungsenergie Rückschlüsse auf die gemessenen K_i Werte zu ziehen.

Sowohl bei den Dockingstudien als auch bei den MD wurden jeweils die gleichen Kristallstrukturen aus der PDB verwendet, namentlich 2ITY (EGFR), 2P2H (VEGFR) und 2ZM3 (IGFR).

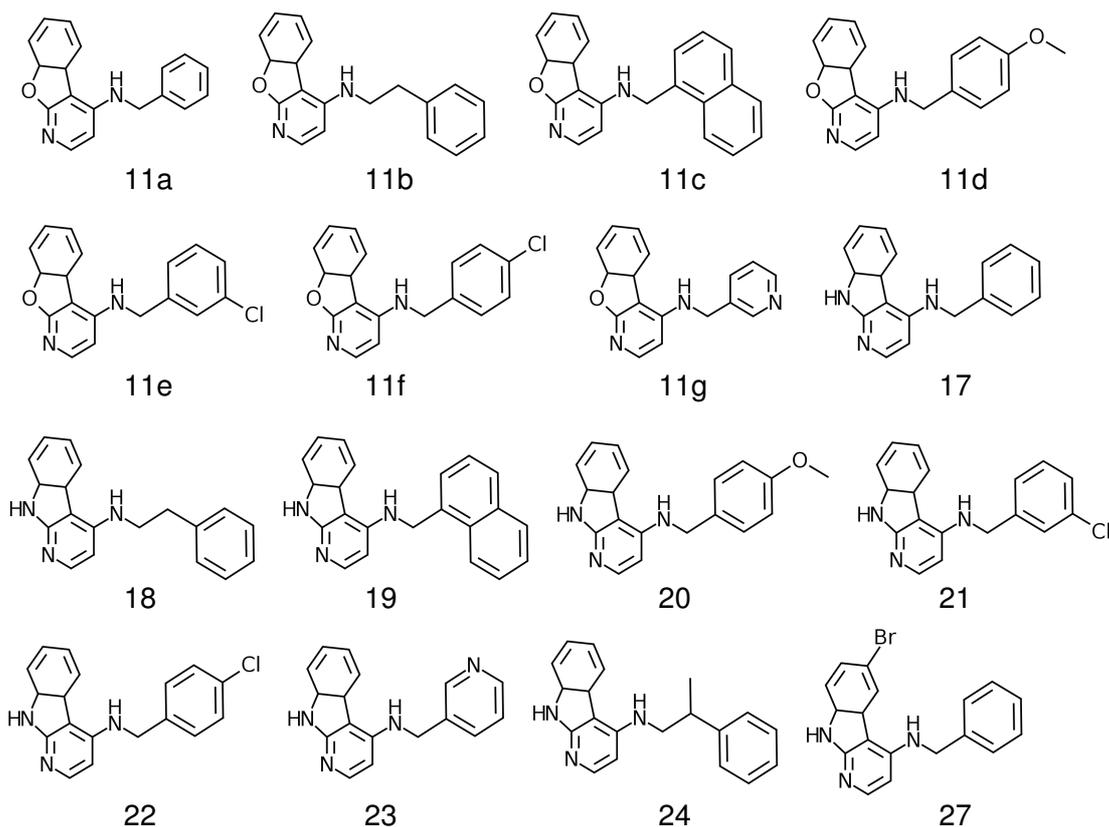


Abbildung 3.18: Getestete Liganden für VEGFR, EGFR und IGFR. 11a - 11g 1-aza-9-oxafluorene, 17 - 25 α -carboline, 27 6-bromo Derivat.

Tabelle 3.25: K_i -Werte aus der biologischen Testung
11g zeigt keine Aktivität. n.i.: $K_i > 100\mu\text{M}$

Name	K_i [nM]		
	EGFR	VEGFR	IGFR
11a	590	470	370
11b	3450	2010	2110
11c	680	300	570
11d	400	n.i.	150
11e	n.i.	n.i.	580
11f	500	180	110
11g	n.i.	n.i.	n.i.
17	85	50	7840
18	1040	640	180
19	360	410	150
20	450	390	130
21	99	100	118
22	1042	n.i.	430
23	3830	n.i.	3430
24	840	2620	420
27	43	500	33

Tabelle 3.26: Dockingergebnisse
Angabe ist der GOLDScore. Die besten Werte sind hervorgehoben.

Name	EGFR	VEGFR	IGFR
11a	43,92	35,79	26,07
11b	42,01	38,48	29,17
11c	49,74	47,19	29,77
11d	41,68	42,74	26,01
11e	48,36	36,66	28,37
11f	44,89	38,66	27,93
11g	44,56	34,99	25,05
17	47,68	51,22	44,29
18	34,09	51,12	50,02
19	39,37	58,09	41,52
20	39,16	52,79	49,32
21	45,04	54,55	48,15
22	49,49	52,77	50,02
23	46,47	49,05	41,29
24	40,49	47,14	50,79
27	51,08	53,98	39,99

3.4.1 Docking

Bei den hier durchgeführten Dockingstudien wurden alle Wassermoleküle und alle kleine Moleküle aus der Kristallstruktur entfernt. Desweiteren wurden die Grundeinstellungen von GOLD benutzt, und als Zielregion wurde ein Bereich von 15 Å um den Gatekeeper (THR916 bei VEGFR, THR790 bei EGFR, MET1082 bei IGFR) festgelegt. Mit den gleichen Einstellungen wurde zuvor getestet, ob der Bindungsmodus von dem in der jeweiligen Kristallstruktur co-kristallisierten Liganden vom Programm wiedergefunden wurde. Der RMSD-Wert der am besten bewerteten Lösung für alle Liganden betrug < 1 Å, damit wurde gezeigt, dass die Einstellungen den gestellten Anforderungen entsprachen, und für die Dockingstudien verwendet werden konnten. Tabelle 3.26 zeigt die GOLDScores der besten Dockingpose für jeden Ligand.

Abbildung 3.19 zeigt für verschiedene getestete Substanzen in EGFR die am besten bewerteten Posen. Es ist zu erkennen, dass alle Substanzen einen ähnlichen Bindungsmodus aufweisen. Die Trizyklen sind alle zur Hinge-Region orientiert und die Heteroatome bilden Wasserstoffbrücken zum LEU792 und MET793 aus. Die Phenylreste von Substanz 11b, 17 und 27, sowie der Pyridinrest von Substanz 23 zeigen alle in die gleiche Richtung zum LYS745. Obwohl 11b einen um eine Methylgruppe verlängerte Kette zwischen Trizyklus und Phenylring hat, als die anderen abgebildeten Substanzen, passt der Phenylring an die gleiche Stelle. Substanz 27 wird beim Docking in EGFR am besten bewertet, und zeigt ebenfalls dort die höchste Aktivität. Auffallend ist, dass sowohl 11b, als auch 23 bei allen Aktivitätsbestimmungen einen

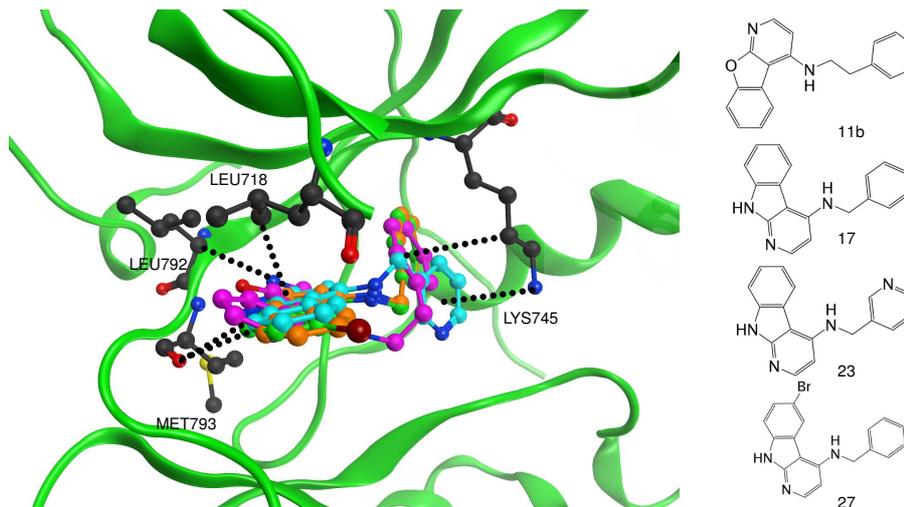


Abbildung 3.19: EGFR Vergleich Dockinglösungen

Dargestellt sind die Dockinglösungen verschiedener Substanzen. 23 (cyan); 11b (magenta), 17 (orange), 27 (grün); Rezeptoratome in schwarz. Substanzen 27 und 17 zeigten die höchste Aktivität, 11b und 23 die Niedrigste.

K_i -Wert von über $2 \mu\text{M}$ aufweisen. Vergleicht man den GOLDScore nun mit den in Tabelle 3.25 angegebenen K_i -Werten, so ist kein Zusammenhang zu erkennen, und es lassen sich anhand der erzielten Scores keine Aussagen über die Aktivität treffen.

Bei den Kinasen IGFR und VEGFR wurde ebenfalls der in Abbildung 3.19 gezeigte Bindungsmodus gefunden. Jedoch waren bei diesen Kinasen andere Substanzen aktiver. Auch in diesen Fällen ist kein Zusammenhang zwischen gefundener Pose und gemessener Aktivität zu erkennen.

In Abbildung 3.20 ist für jede Kinase die jeweils best-bewertete Pose dargestellt. Alle drei Liganden bilden zwei Wasserstoffbrücken zur Hinge-Region aus, und die Phenylringe zeigen in Richtung des Lysinrestes.

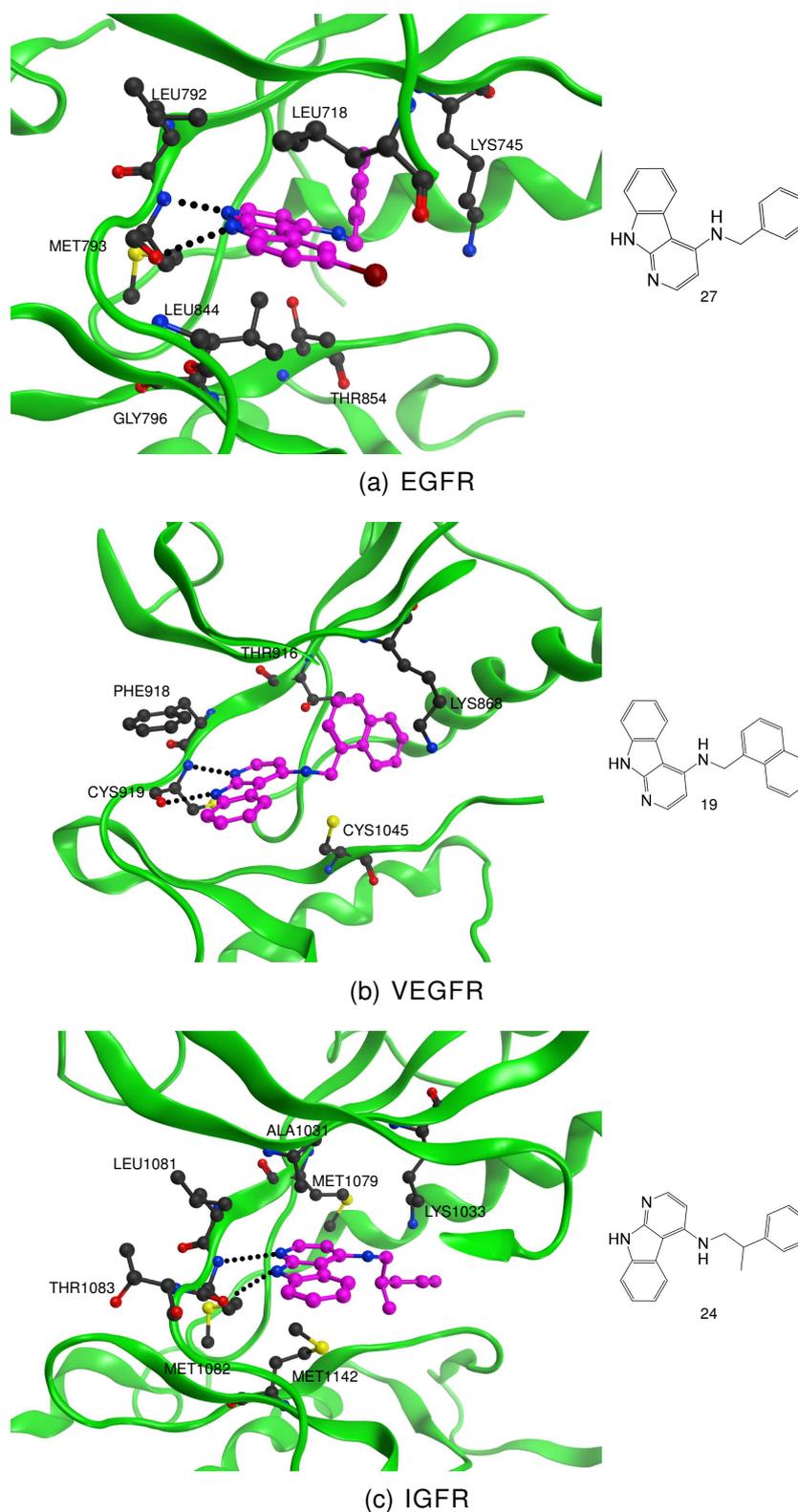


Abbildung 3.20: Ausgewählte Dockingposen

Die abgebildeten Liganden wurden in der jeweiligen Kinase am besten bewertet. EGFR Substanz 27, VEGFR Substanz 19, IGFR Substanz 24; Substanzen immer magenta, Rezeptortome schwarz

3.4.2 MM-PB/SA

Mit den aus den Dockingstudien erhaltenen Lösungen für die einzelnen Liganden wurden MDs durchgeführt, um die Stabilität des Komplexes zu bestätigen, und um die freien Energien zu bestimmen. Alle Simulationen wurden für einen Zeitraum von 10 Nanosekunden, einschließlich Equilibrierung und Aufheizen mit expliziten Solvent durchgeführt. Als Wassermodell wurde *TIP3P* verwendet. Die simulierte Wasserbox war bei der Initialisierung maximal 10 Å vom Rand des Komplexes entfernt und hatte die Form eines Oktaeders. Damit der Komplex sich über den Verlauf der Simulation nicht aus der Wasserbox heraus bewegt, wurden periodische Randbedingungen verwendet. Über den gesamten Simulationszeitraum von jeweils 10 ns blieben dabei alle Komplexe stabil.

Die Substanzen 19, 24 und 27 haben in den vorangegangenen Dockingstudien die besten Interaktionen gezeigt, weshalb diese Verbindungen für weitere Untersuchungen durch eine MD ausgewählt wurden. Verbindung 19 wurde dabei mit VEGFR simuliert, Substanz 24 mit IGFR und Verbindung 27 mit EGFR.

Abbildung 3.21 zeigt den Komplex von EGFR und Substanz 27 vor und nach der MD. Die Hinge-Region hat sich kaum verändert, genauso wie die Lage des Liganden ebenfalls nur sehr wenig. Der Trizyklus bleibt in der gleichen Position, jedoch wird eine andere Wasserstoffbrücke zur Hinge-Region ausgebildet. Sie besteht zwar noch weiterhin zum MET793, aber sie besteht nicht mehr zwischen dem Der Phenylrest kippt um ca. 90°, behält dennoch die gleiche Lage zum LYS745.

Abbildung 3.22 zeigt die Sicht vom N-terminalen Bereich auf die Hinge-Region von VEGFR im Komplex mit Substanz 19. Die Wasserstoffbrücken des Trizyklus zum CYS919 bleiben über den Verlauf der MD bestehen. Die deutlichste Veränderung ist an der Kette zwischen Alpha-Carbolin (9*H*-Pyrido[2,3-*b*]indol) und Naphthalin zu beobachten. An der Bindung zur letzten Methylgruppe klappt der Naphthalinrest jedoch um ca. 90° aus der hydrophoben Tasche II raus, und entfernt sich so vom LYS868.

Substanz 24 zeigt bei der MD im Komplex mit IGFR etwas Ähnliches wie Substanz 19 (s. Abbildung 3.23). Der Trizyklus hat sich ein wenig innerhalb der Bindetasche in Richtung der hydrophoben Tasche II bewegt, und ist weiterhin zur Hinge-Region ausgerichtet. Die Anzahl der Wasserstoffbrücken hat sich auf eine reduziert. Die größte Veränderung ist hier an der Kette zwischen Trizyklus und Phenylring aufgetreten. Am sekundären Amin findet die Drehung statt, der Phenylring kippt hier ebenfalls um ca. 180° und zeigt nicht mehr zum LYS1033, sondern aus der Tasche raus.

Der Verlauf der RMSD-Werte dieser Substanzen ist in Abbildung 3.24 dargestellt. Die Bewegung von Substanz 27 in EGFR (Abbildung 3.24 a) ist nicht stark. Der RMSD-Wert schwankt zwischen 0,5 und 1,5 Å. Nach 4 ns steigt dieser auf 2,5 Å, fällt aber wieder auf den vorherigen Wert zurück. Zum Schluß steigt er wieder auf 2,5 Å.

Substanz 19 mit den besten Interaktionen bei VEGFR (Abbildung 3.24 b) zeigt nur sehr kurz einen maximalen RMSD-Wert von etwa 3 Å, und nach 4 ns Simulationszeit bewegt sich der RMSD-Wert zwischen 2 und 2,5 Å. Obwohl es auch hier kurz vor Ende der Simulation zu einer Umorientierung kommt, die später wieder in die Ausgangslage übergeht, liegt dieser Ligand stabil in der Bindetasche.

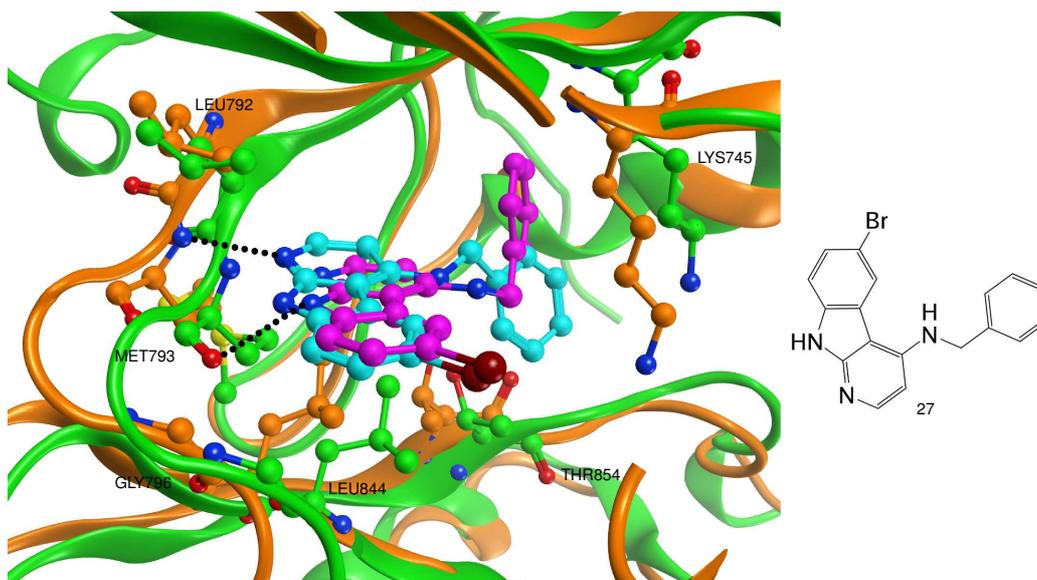


Abbildung 3.21: Vergleich der Lage nach MD von Substanz 27 in EGFR Kristallstruktur (grün), Dockingpose (magenta), orange Protein nach der MD, cyan Ligand nach MD.

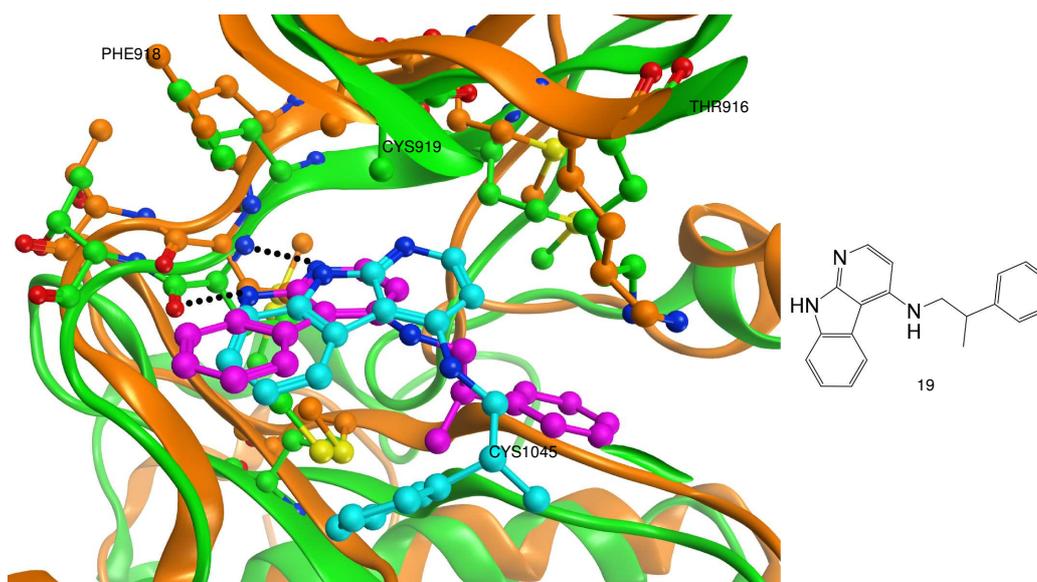


Abbildung 3.22: Vergleich der Lage nach MD von Substanz 19 in VEGFR Kristallstruktur (grün), Dockingpose (magenta), Protein nach MD (orange), Ligand nach MD (cyan).

Bei der Simulation von Substanz 24 bei IGFR (Abbildung 3.24 c) findet direkt zu Beginn eine Neuausrichtung innerhalb der Bindetasche statt. Danach ändert sich die Position des Liganden kaum noch.

In Tabelle 3.27 sind die errechneten Werte für die Enthalpie (ΔH), die Entropie (ΔS) und die freie Energie (ΔG) aufgeführt. Die Energien wurden anhand von 100 Snapshots

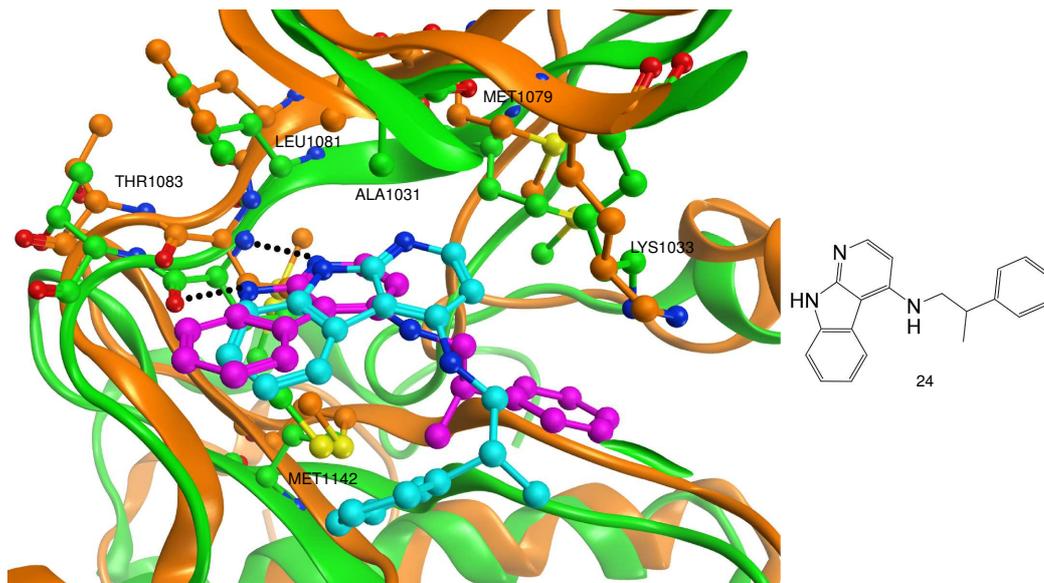


Abbildung 3.23: Vergleich der Lage nach MD von Substanz 24 in IGFR Kristallstruktur (grün), Dockingpose (magenta), Protein nach der MD (orange), Ligand nach MD (cyan).

erstellt. Diese Snapshots wurden aus einem Teil der Simulation entnommen, wo diese stabil verlief. In allen Fällen zwischen der achten und neunten Nanosekunde. Trägt man die pK_i -Werte gegen die errechneten Werte auf, so erhält man einen Graphen, aus denen man das r^2 ermitteln kann. Der Wert r^2 ist ein Maß für den Grad der Korrelation zwischen zwei Zahlenreihen, und kann Werte $0 \leq 1$ annehmen, wobei 0 für keinen Zusammenhang steht, und 1 für einen 100%igen Zusammenhang. Die Graphen dazu sind in den Abbildungen 3.25(a) – 3.25(c).

Die Standardabweichung der Energiewerte für die Entropie (ΔS) ist deutlich größer als bei der Enthalpie (ΔH). Es ist zu sehen, dass es keine Korrelation zwischen den errechneten Werten für die freie Energie und den gemessenen K_i -Werten gibt. Aus diesem Grund sind keine r^2 -Werte aufgeführt. Wie an den Werten in Tabelle 3.25 zu sehen ist, haben bei jeder getesteten Kinase mindestens zwei Substanzen einen K_i -Wert von über 1000 μM . Diese zeigen eine deutlich schwächere Wirkung als die anderen getesteten Substanzen.

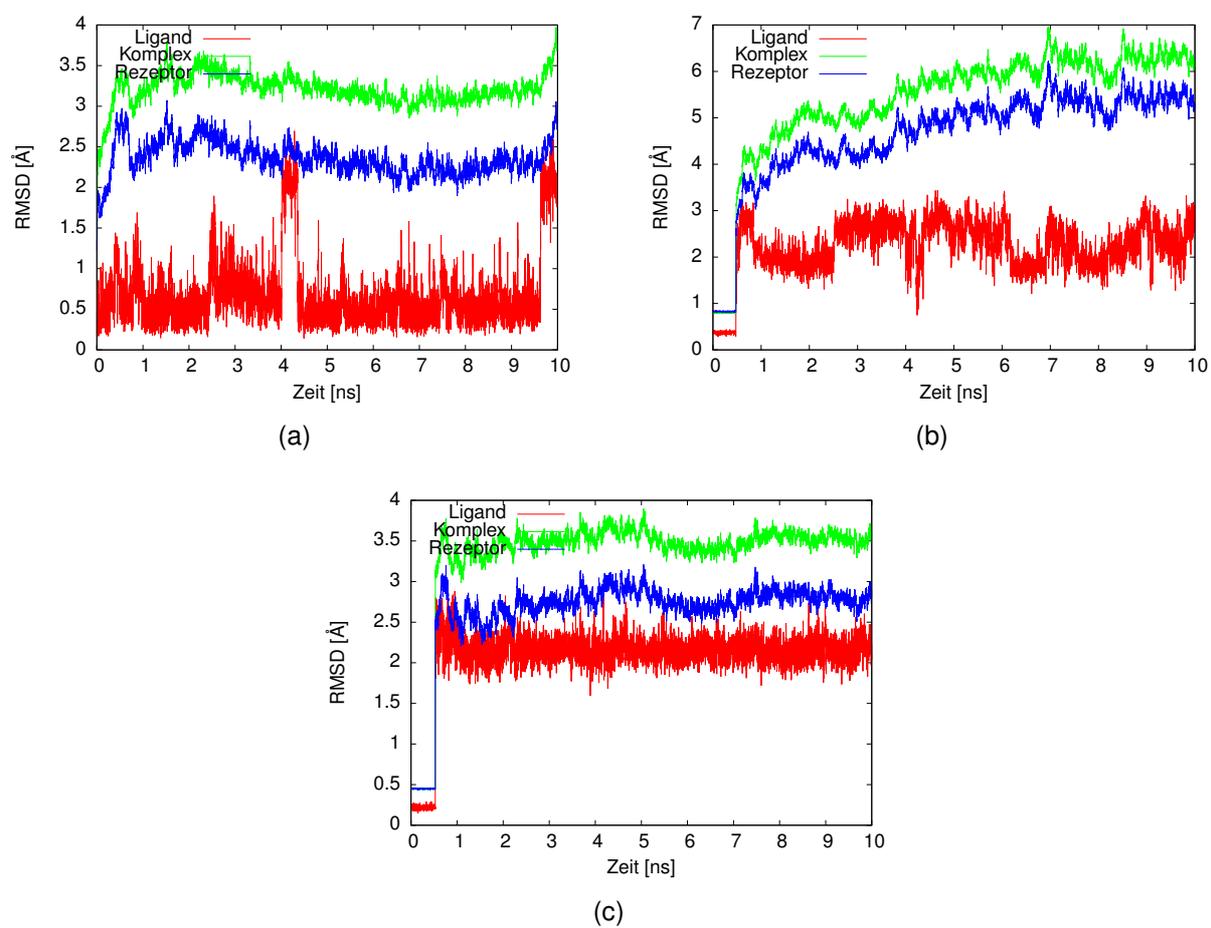


Abbildung 3.24: RMSD-Plots der Komplexe
Jeweils mit den besten Dockinglösungen. a) EGFR Substanz 27, b) VEGFR Substanz 19, c)
IGFR Substanz 24.

Tabelle 3.27: Berechnete freien Energien aller getesteten Substanzen nach der MD in allen Kinasen. Werte in [kcal/mol]

EGFR					
Name	ΔH (PB)	ΔH (GB)	ΔS	ΔG (PB)	ΔG (GB)
11a	-22,49	-28,90	-18,96	-3,53	-9,94
11b	-20,87	-33,81	-19,76	-1,11	-14,05
11c	-32,58	-39,70	-20,91	-11,67	-18,79
11d	-27,74	-34,65	-18,91	-8,83	-15,74
11e	-27,82	-30,84	-21,98	-5,84	-8,86
11f	-25,82	-37,06	-18,47	-7,35	-18,59
11g	-22,81	-34,20	-20,47	-2,34	-13,73
17	-25,49	-33,20	-19,41	-6,08	-13,79
18	-17,31	-28,74	-17,60	0,29	-11,14
19	-32,68	-41,83	-20,11	-12,57	-21,72
20	-21,29	-30,66	-21,40	0,11	-9,26
21	-34,32	-44,37	-19,98	-14,34	-24,39
22	-23,99	-37,93	-20,54	-3,45	-17,39
23	-24,24	-33,16	-18,33	-5,91	-14,83
24	-25,57	-27,84	-17,28	-8,29	-10,56
27	-25,08	-36,99	-20,69	-4,39	-16,30
VEGFR					
11a	-22,72	-18,73	-18,14	-4,58	-0,59
11b	-20,41	-32,39	-20,62	0,21	-11,84
11c	-28,49	-32,48	-21,53	-6,96	-10,95
11d	-24,34	-29,37	-17,02	-7,32	-12,35
11e	-24,48	-34,46	-18,03	-6,45	-16,43
11f	-25,31	-37,26	-20,10	-5,21	-17,16
11g	-25,57	-34,26	-16,56	-9,01	-17,70
17	-21,95	-28,26	-17,78	-4,17	-10,48
18	-21,62	-27,01	-19,07	-2,55	-7,94
19	-29,57	-33,52	-19,83	-9,74	-13,69
20	-25,27	-45,28	-20,88	-4,38	-24,40
21	-26,37	-40,35	-19,04	-7,33	-21,31
22	-22,59	-37,53	-20,60	-1,99	-16,93
23	-22,20	-29,31	-17,93	-4,27	-11,38
24	-27,20	-25,20	-18,40	-8,80	-6,80
27	-20,24	-35,23	-18,54	-1,86	-16,69
IGFR					
11a	-15,30	-18,04	-18,60	3,30	0,56
11b	-23,42	-32,85	-21,01	-2,41	-11,84
11c	-24,77	-32,25	-21,78	-2,99	-10,47
11d	-20,80	-30,17	-17,90	-2,90	-12,27
11e	-14,32	-31,00	-18,38	4,06	-12,62
11f	-20,75	-30,33	-18,98	-1,77	-11,35
11g	-13,59	-22,65	-16,11	2,52	-6,54
17	-13,52	-18,89	-18,51	4,99	-0,38
18	-15,24	-22,09	-19,59	4,35	-2,50
19	-31,90	-41,54	-19,64	-12,26	-21,90
20	-34,87	-45,38	-18,07	-16,80	-27,31
21	-17,66	-26,52	-17,00	-0,66	-9,52
22	-22,45	-27,09	-19,22	-3,23	-7,87
23	-21,80	-29,05	-19,42	-2,38	-9,63
24	-6,06	-22,03	-17,44	11,38	-4,59
27	-20,27	-25,83	-18,27	-2,00	-7,56

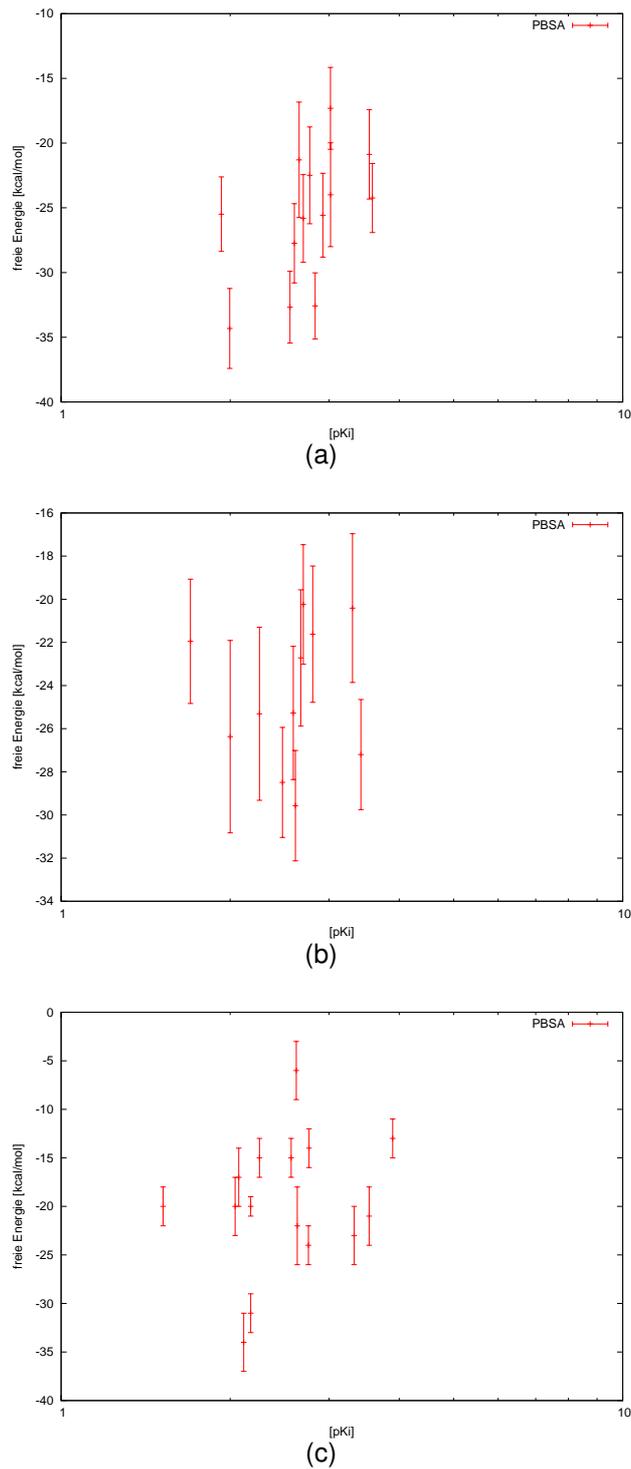


Abbildung 3.25: Korrelationen zwischen errechneter freier Energie und pK_i -Werten. Die Werte für die freien Energien stammen aus den MM-PB/SA Berechnungen (s. Tabelle 3.27) und die pK_i -Werte aus Tabelle 3.25 a) EGFR b) VEGFR c) IGFR es liegt keine Korrelation zwischen den gemessenen pK_i -Werten und den errechneten freien Energien vor.

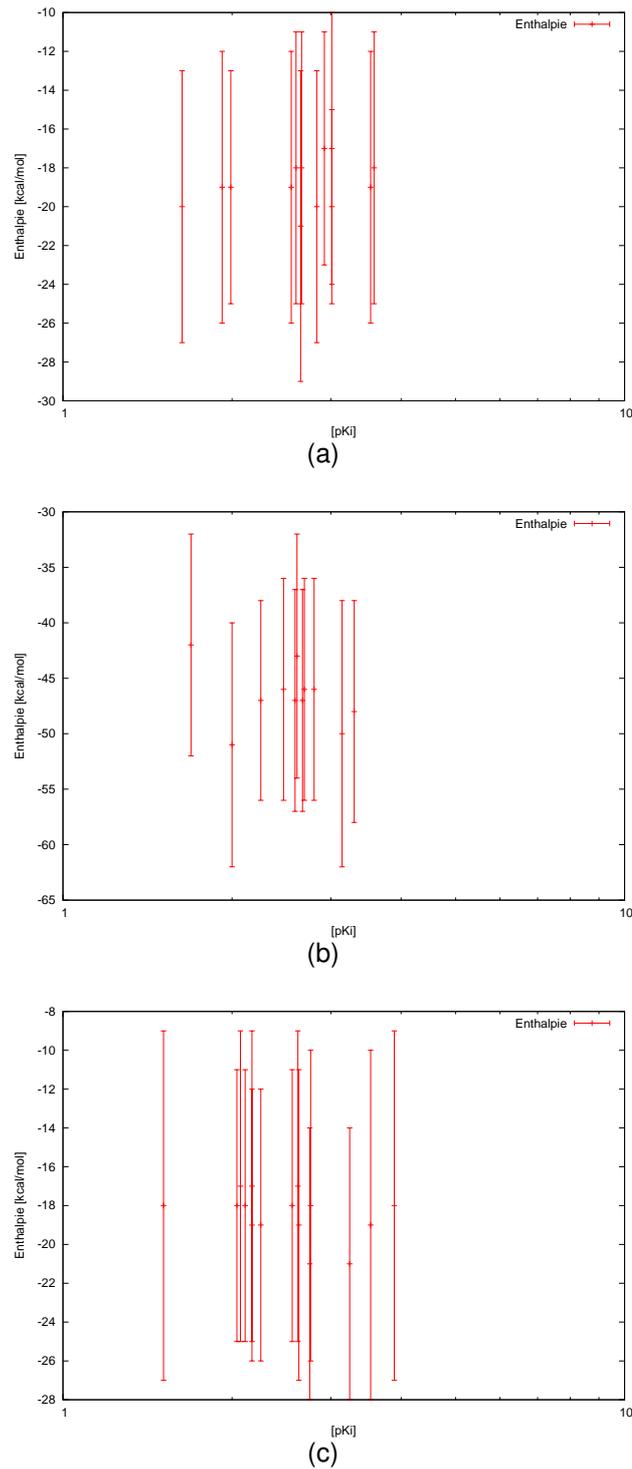


Abbildung 3.26: Korrelationen zwischen errechneten ΔH -Werten pK_i -Werten. Die Werte für ΔH stammen aus den MM-PB/SA Berechnungen (s. Tabelle 3.27) und die pK_i -Werte aus Tabelle 3.25 a) EGFR b) VEGFR c) IGFR es liegt keine Korrelation zwischen den gemessenen pK_i -Werten und den errechneten ΔH -Werten.

3.4.3 LR-MMPBSA

Zusätzlich zu den schon bekannten MDs wurde hier auch die in Kapitel 2.5.3 beschriebene Methode LR-MMPBSA verwendet, um zu prüfen, ob die Ergebnisse eine bessere Korrelation mit den gemessenen K_i -Werten haben. In Tabelle 3.28 sind die errechneten Werte für PBTOT und GBTOT aufgeführt.

Tabelle 3.28: LR-MMPBSA

Aufgeführt sind die ermittelten freien Energien nach der MD in [kcal/mol]

Name	EGFR		VEGFR		IGFR	
	PBTOT	GBTOT	PBTOT	GBTOT	PBTOT	GBTOT
11a	-24,25	-38,46	-13,95	-33,05	-5,92	-27,41
11b	-27,08	-38,46	-14,21	-34,66	-19,55	-31,14
11c	-32,28	-43,28	-25,09	-39,14	-19,05	-34,28
11d	-24,05	-38,83	-24,82	-37,31	-16,30	-30,94
11e	-23,35	-38,38	-14,34	-35,33	-9,92	-31,39
11f	-26,27	-39,99	-14,57	-35,03	-10,40	-30,21
11g	-20,44	-34,80	-6,45	-25,37	-19,71	-26,90
17	-30,69	-42,05	-19,85	-33,39	-25,15	-32,20
18	-22,96	-37,79	-14,41	-31,07	-17,42	-35,46
19	-38,23	-48,20	-29,89	-40,02	-16,71	-35,32
20	-16,54	-31,83	-22,05	-35,43	-19,98	-34,96
21	-28,50	-42,73	-22,35	-36,31	-11,35	-32,42
22	-32,06	-45,16	-21,66	-35,59	-28,04	-35,76
23	-25,72	-34,08	-19,42	-34,28	-23,23	-32,77
24	-21,91	-30,55	-19,19	-33,73	-13,83	-37,54
27	-33,69	-45,37	-21,79	-35,80	-19,95	-35,84

Trägt man auch hier die errechneten Werte für die freie Energie gegen die gemessenen K_i -Werte auf, sieht man, dass es ebenfalls keine Korrelation zwischen diesen Werten gibt.

3.4.4 Diskussion

Die in den Dockingstudien gefundenen Lösungen der getesteten Substanzen decken sich mit den bekannten Interaktionsmustern von ATP-Kompetitiven Inhibitoren. Alle Trizyklen bildeten die typischen Wasserstoffbrücken zur Hinge-Region aus. Die Verbindungen blieben über den Simulationszeitraum in der ATP-Bindetasche, und verändern kaum ihre Position. Dies deutet darauf hin, dass die gefundenen Posen richtig sind. Das für diese Posen jeweils ein negativer ΔG -Wert bestimmt wurde, unterstützt die Vermutung. Leider war keine der verwendeten Methoden in der Lage einen linearen Zusammenhang zwischen errechneten freien Energien und K_i -Werten herzustellen. In früheren Studien konnten sowohl MM-PB/SA als auch LR-MMPBSA zeigen, dass sich diese Methoden eignen, um Aktivitäten von Substanzen vorherzusagen [255, 183].

Einer der Gründe das diese Methoden in diesen Untersuchungen nicht funktioniert hat, ist wahrscheinlich, dass die Aktivität der getesteten Substanzen sich über einen zu kleinen Bereich verteilt. In früheren Veröffentlichungen, die sich mit dieser Frage befasst haben, ging der Bereich der Aktivität über mindestens vier Größenordnungen [183, 256–261]. Ein weiteres Problem ist, dass der Datensatz mit nur 16 Substanzen recht klein ist, vorallem da nicht alle Substanzen aktiv sind. Bei den Untersuchungen an VEGFR ist dies ein Problem, da fünf von diesen 16 Substanzen nicht aktiv sind. Eine Datenbank mit 30 Substanzen und mehr ist besser geeignet, um eine statistisch signifikante Aussage zu treffen. Ein weiteres Problem hierbei ist, dass die meisten aktiven Substanzen in einem engen Aktivitätsbereich liegen, und die Aktivitätswerte nicht über den gesamten Bereich verteilt sind. Entfernt man die weniger aktiven Substanzen aus der Betrachtung, wird das Bestimmtheitsmaß R^2 nicht besser. Vor allem ergibt sich dann das Problem, dass der Bereich der Aktivität viel zu klein ist. Substanz 11b und 23 haben in allen getesteten Fällen K_i -Werte von über 2000 μM , und 11g ist bei allen Kinasen inaktiv. Diese beiden Substanzen unterscheiden sich strukturell wenig von den übrigen. Substanz 23 und 11g haben sowohl den Pyridinring als auch das α -Carbolin gemeinsam. Da dieses seine Lage während der Simulation bei keiner der getesteten Substanzen deutlich verändert, kann dieser Trizyklus nicht der Grund für die schlechte Aktivität sein. Der Pyridinrest jedoch ist hydrophil [262]. Somit ist es möglich, dass es für die Substanzen 11g und 23 energetisch günstiger ist, wenn sich das Pyridin aus der hydrophoben Tasche II in Richtung Lösungsmittel bewegt. Dies erklärt leider nicht, weshalb sich das lipophile Naphtalin [263] aus dem hydrophoben Bereich in das Lösungsmittel bewegt. In der hydrophoben Tasche gibt es zwar ein hydrophiles Threonin, aber dieser abstoßende Effekt dürfte nicht so stark sein. Um den genauen Effekt dieses Threonins zu bestimmen, könnte man Mutationsstudien an diesen Kinasen durchführen, und es durch hydrophobe Aminosäuren an dieser Stelle ersetzen.

Die Berechnungen der freien Energie wurden mittels Molecular Mechanics generalized Born/Surface Area (MM-GB/SA) gemacht. Diese Methode ist nicht gut geeignet, um in hydrophoben Bereichen die Energien zuverlässig zu bestimmen [264]. Da die Bindetasche über mehrere hydrophobe Taschen verfügt, die für die Bindung mit einem Liganden wichtig sind, könnte dies eine Erklärung dafür sein, dass MM-GB/SA-Berechnungen hier nicht funktionieren.

4. ZUSAMMENFASSUNG

Ziel dieser Arbeit war zum einen die Untersuchung von verschiedenen virtuellen Screening-Methoden anhand von Proteinkinasen.

Dazu wurde eine retrospektive Studie anhand von veröffentlichten Datensätzen durchgeführt, und die Methoden wurden darauf hin untersucht, wie gut sie zwischen aktiven und inaktiven Liganden unterscheiden können. Die inaktiven Verbindungen wurden dahingehend ausgewählt, dass sie den aktiven Substanzen ähnlich sind. So sollte verhindert werden, dass schon allein durch die strukturellen Unterschiede eine Differenzierung möglich ist. Bei der Testung wurden nicht nur die einzelnen Programme mit ihren jeweiligen Konformationsgeneratoren getestet und Scoring-Funktionen untersucht, sondern es wurden auch die von einem Programm gefundenen Lösungen durch die Scoring-Funktionen der anderen Programme neu bewertet. Zusätzlich wurde ein CONSENSUS-SCORE-Modell untersucht. Dazu wurden die für jede Lösung erhaltenen Bewertungen normalisiert, und für jede getestete Verbindung wurden die von jeder Scoring-Funktion errechneten Werte addiert, und die Lösungen anhand des CONSENSUS-SCORE neu sortiert. Um zu prüfen, wie gut die unterschiedlichen Kombinationen zwischen aktiven und inaktiven Substanzen differenzieren konnten, wurden drei verschiedene Gütemaße verwendet, wobei jedes dieser Gütemaße einen anderen Aspekt näher beleuchtete.

Es stellte sich raus, dass es keine signifikanten Unterschiede bei den verwendeten Methoden gibt, und alle eine ähnliche Leistung zeigen, um zwischen aktiven und inaktiven Liganden zu unterscheiden. ROCS hat jedoch eine etwas bessere Leistung gezeigt. Vor allem im Bereich der frühen Anreicherung, das heißt innerhalb der besten fünf Prozent einer Datenbank zeigt ROCS die beste Leistung bei der Unterscheidung zwischen den inaktiven und aktiven Substanzen.

Desweiteren wurden mögliche neue Inhibitoren für die MYT1-Kinase gedockt. Als Struktur in welche die Verbindungen gedockt werden sollten diente hier ein Homologie-Modell. Die Prüfung des Modells ergab, dass dieses den Ansprüchen für ein vertrauenswürdigen Modell genügt. Es wurden Dockingstudien und Moleküldynamik-Simulationen durchgeführt, um eine Vorhersage über die Aktivität der Substanzen zu treffen. Sowohl nach den Dockingstudien, und den Simulationen wurden zwei Verbindungen als potentiell aktiv vorhergesagt. Dies konnte leider nicht in der biologischen Testung bestätigt werden.

Für die Kinasen, IGFR, EGFR und VEGFR wurde untersucht, wie der Bindemodus einer neuen Substanzklasse ist. Untersucht wurden 1-aza-9-oxafluorene und α -carboline. Die gemessene inhibitorische Wirkung konnte durch die gefundenen Dockinglösungen und Moleküldynamik-Simulationen bestätigt werden. Jedoch war es nicht möglich eine Korrelation zwischen den errechneten ΔG -Werten und den im Labor ermittelten K_i -Werten festgestellt werden.

5. LITERATUR

- [1] J. M. Rollinger, H. Stuppner, and T. Langer. Virtual screening for the discovery of bioactive natural products. *Progress in drug research.*, 2008, 65.
- [2] U. Rester. From virtuality to reality - Virtual screening in lead discovery and lead optimization: a medicinal chemistry perspective. *Curr. Opin. Drug Discov. Dev.*, 2008, 11(4):559–568.
- [3] R. S. Bohacek, C. McMartin, and W. C. Guida. The art and practice of structure-based drug design: A molecular modeling perspective. *Med. Res. Rev.*, 1996, 16(1):3–50.
- [4] G. Manning, D. B. Whyte, R. Martinez, T. Hunter, and S. Sudarsanam. The Protein Kinase Complement of the Human Genome. *Science*, 2002, 298(5600):1912–1934.
- [5] T. Tuccinardi, M. Botta, A. Giordano, and A. Martinelli. Protein Kinases: Docking and Homology Modeling Reliability. *J. Chem. Inf. Model.*, 2010, 50(8):1432–1441.
- [6] K. Anamika, N. Garnier, and N. Srinivasan. Functional diversity of human protein kinase splice variants marks significant expansion of human kinome. *BMC genomics*, 2009, 10(1):622–629.
- [7] P. Blume-Jensen and T. Hunter. Oncogenic kinase signalling. *Nature*, 2001, 411(6835):355–365.
- [8] P. Cohen. Protein kinases the major drug targets of the twenty-first century? *Nature Rev Drug Disc.*, 2002, 1(4):309–315.
- [9] L. R. Pearce, D. Komander, and D. R. Alessi. The nuts and bolts of AGC protein kinases. *Nat Rev Mol Cell Biol*, 2010, 11(1):9–22.
- [10] N. Kannan, N. Haste, S. S. Taylor, and A. F. Neuwald. The hallmark of AGC kinase functional divergence is its C-terminal tail, a cis-acting regulatory module. *PNAS*, 2007, 104(4):1272–1277.
- [11] F. C. Stevens. Calmodulin: an introduction. *Canadian journal of biochemistry and cell biology = Revue canadienne de biochimie et biologie cellulaire*, 1983, 61(8):906–910.
- [12] D. Chin and A. R. Means. Calmodulin: a prototypical calcium sensor. *Trends Cell Biol*, 2000, 10(8):322–328.
- [13] E. J. Eide and D. M. Virshup. Casein kinase i: Another cog in the circadian clockworks. *Chronobiol Int*, 2001, 18(3):389–398.
- [14] G. Manning. Genomic overview of protein kinases. *WormBook : the online review of C. elegans biology*, 2005, pages 1–19.
- [15] N. Kannan and A. F. Neuwald. Evolutionary constraints associated with functional specificity of the CMGC protein kinases MAPK, CDK, GSK, SRPK, DYRK, and CK2 α . *Protein Sci.*, 2004, 13(8):2059–2077.

- [16] S. L. Pelech. Kinase connections on the cellular intranet. signalling pathways. *Curr Biol*, 1996, 6(5):551–554.
- [17] B. N. Lee and E. A. Elion. The MAPKKK Ste11 regulates vegetative growth through a kinase cascade of shared signaling components. *PNAS*, 1999, 96(22):12679–12684.
- [18] J. M. Goldberg, G. Manning, A. Liu, P. Fey, K. E. Pilcher, Y. Xu, and J. L. Smith. The Dictyostelium Kinome-Analysis of the Protein Kinases from a Simple Model Organism. *PLoS Genet*, 2006, 2(3):291–303.
- [19] D. R. Robinson, Y. M. Wu, and S. F. Lin. The protein tyrosine kinase family of the human genome. *Oncogene*, 2000, 19(49):5548–5557.
- [20] C. A. Bradham, K. R. Foltz, W. S. Beane, M. I. Arnone, F. Rizzo, J. A. Coffman, A. Mushegian, M. Goel, J. Morales, A.-M. Genevriere, F. Lapraz, A. J. Robertson, H. Kelkar, M. Loza-Coll, I. K. Townley, M. Raisch, M. M. Roux, T. Lepage, C. Gache, D. R. McClay, and G. Manning. The sea urchin kinome: A first look. *Dev. Biol.*, 2006, 300(1):180–193.
- [21] Parkinson Study Group PRECEPT Investigators. Mixed lineage kinase inhibitor CEP-1347 fails to delay disability in early Parkinson disease. *Neurology*, 2007, 69(15):1480–1490.
- [22] M. Scharpfenecker, M. van Dinther, Z. Liu, R. L. van Bezooijen, Q. Zhao, L. Pukac, C. W. Löwik, and P. ten Dijke. BMP-9 signals via ALK1 and inhibits bFGF-induced endothelial cell proliferation and VEGF-stimulated angiogenesis. *J. Cell Sci.*, 2007, 120 (Pt 6):964–972.
- [23] S. K. Hanks and T. Hunter. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *The FASEB Journal*, 1995, 9(8):576–596.
- [24] F. Zuccotto, E. Ardini, E. Casale, and M. Angiolini. Through the Gatekeeper Door: Exploiting the Active Kinase Conformation. *J. Med. Chem.*, 2009, 53(7):2681–2694.
- [25] D. Huang, T. Zhou, K. Lafleur, C. Nevado, and A. Caflisch. Kinase selectivity potential for inhibitors targeting the ATP binding site: a network analysis. *Bioinformatics*, 2010, 26 (2):198–204.
- [26] P. J. Alaimo, Z. A. Knight, and K. M. Shokat. Targeting the gatekeeper residue in phosphoinositide 3-kinases. *Bioorg. Med. Chem.*, 2005, 13(8):2825–2836.
- [27] J. Zhang, P. L. Yang, and N. S. Gray. Targeting cancer with small molecule kinase inhibitors. *Nature Reviews Cancer*, 2009, 9(1):28–39.
- [28] B. Nolen, S. Taylor, and G. Ghosh. Regulation of Protein Kinases Controlling Activity through Activation Segment Conformation. *Mol Cell*, 2004, 15(5):661–675.
- [29] G. Burnett and E. P. Kennedy. The enzymatic phosphorylation of proteins. *J Biol Chem*, 1954, 211(2):969–980.
- [30] E. H. Fischer, D. J. Graves, E. R. Crittenden, and E. G. Krebs. Structure of the Site Phosphorylated in the Phosphorylase b to a Reaction. *J Biol Chem*, 1959, 234(7):1698–1704.

- [31] E. H. Fischer and E. G. Krebs. Conversion of Phosphorylase B to Phosphorylase A in Muscle Extracts. *J Biol Chem*, 1955, 216(1):121–132.
- [32] C. C. Lee, Y. Jia, N. Li, X. Sun, K. Ng, E. Ambing, M.-Y. Gao, S. Hua, C. Chen, S. Kim, P.-Y. Michellys, S. A. Lesley, J. L. Harris, and G. Spraggon. Crystal structure of the alk (anaplastic lymphoma kinase) catalytic domain. *Biochem. J.*, 2010, 430(3):425–437.
- [33] L. N. Johnson, M. E. M. Noble, and D. J. Owen. Active and Inactive Protein Kinases: Structural Basis for Regulation. *Cell*, 1996, 85(2):149–158.
- [34] V. H. E. Stehelin, D., J. M. Bishop, and P. K. Vogt. Dna related to the transforming gene(s) of avian sarcoma viruses is present in normal avian dna. *Nature*, 1976, 260 (5547):170–173.
- [35] S. Grant. Therapeutic protein kinase inhibitors. *Cell Mol Life Sci*, 2009, 66:1163–1177.
- [36] E. Sala, L. Mologni, S. Truffa, C. Gaetano, G. E. Bollag, and C. Gambacorti-Passerini. BRAF Silencing by Short Hairpin RNA or Chemical Blockade by PLX4032 Leads to Different Responses in Melanoma and Thyroid Carcinoma Cells. *Mol Cancer Res*, 2008, 6(5):751–759.
- [37] D. Schrag. The price tag on progress - chemotherapy for colorectal cancer. *New Engl J Med*, 2004, 351(4):317–319.
- [38] E. W. M. Ng, D. T. Shima, P. Calias, E. T. Cunningham, D. R. Guyer, and A. P. Adamis. Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. *Nat Rev Drug Discov*, 2006, 5(2):123–132.
- [39] H. L. De Bondt, J. Rosenblatt, J. Jancarik, H. D. Jones, D. O. Morgant, and S.-H. Kim. Crystal structure of cyclin-dependent kinase 2. *Nature*, 1993, 363(6430):595–602.
- [40] Z. A. Knight and K. M. Shokat. Features of Selective Kinase Inhibitors. *Chem Biol*, 2005, 12(6):621–637.
- [41] Z. Wang, B. J. Canagarajah, J. C. Boehm, S. Kassisà, M. H. Cobb, P. R. Young, S. Abdel-Meguid, J. L. Adams, and E. J. Goldsmith. Structural basis of inhibitor selectivity in MAP kinases. *Structure*, 1998, 6(9):1117–1128.
- [42] K. Wilson. The structural basis for the specificity of pyridinylimidazole inhibitors of p38 MAP kinase. *Chem Biol*, 1997, 4(6):423–431.
- [43] W. Steglich. Slime moulds (Myxomycetes) as a source of new biologically active metabolites. *Pure Appl. Chem.*, 1989, 61(3):281–288.
- [44] S. Omura, Y. Iwai, A. Hirano, A. Nakagawa, J. Awaya, H. Tsuchya, Y. Takahashi, and R. Masuma. A new alkaloid AM-2282 OF *Streptomyces* origin. Taxonomy, fermentation, isolation and preliminary characterization. *J. Antibiot.*, 1977, 30(4):275–282.
- [45] N. Funato, H. Takayanagi, Y. Konda, Y. Toda, Y. Harigaya, Y. Iwai, and S. Omura. Absolute Configuration of Staurosporine By X-Ray Analysis. *Tetrahedron Letters*, 1994, 35(8):1251–1254.
- [46] M. W. Karaman, S. Herrgard, D. K. Treiber, P. Gallant, C. E. Atteridge, B. T. Campbell, K. W. Chan, P. Ciceri, M. I. Davis, P. T. Edeen, R. Faraoni, M. Floyd, J. P. Hunt, D. J. Lockhart, Z. V. Milanov, M. J. Morrison, G. Pallares, H. K. Patel, S. Pritchard, L. M.

- Wodicka, and P. P. Zarrinkar. A quantitative analysis of kinase inhibitor selectivity. *Nat Biotech*, 2008, 26(1):127–132.
- [47] W. D. Schmitt. *Expression und Regulation der MAPK-Phosphatasen in Ovarialkarzinomen*.
- [48] U. T. Rüegg and G. M. Burgess. Staurosporine, K-252 and UCN-01: potent but nonspecific inhibitors of protein kinases. *Trends Pharmacol Sci*, 1989, 10(6):218–220.
- [49] T. Schindler, W. Bornmann, P. Pellicena, W. T. Miller, B. Clarkson, and J. Kuriyan. Structural Mechanism for STI-571 Inhibition of Abelson Tyrosine Kinase. *Science*, 2000, 289(5486):1938–1942.
- [50] M. S. Cohen, C. Zhang, K. M. Shokat, and J. Taunton. Structural Bioinformatics-Based Design of Selective, Irreversible Kinase Inhibitors. *Science*, 2005, 308(5726): 1318–1321.
- [51] E. L. Kwak, R. Sordella, D. W. Bell, N. Godin-Heymann, R. A. Okimoto, B. W. Brannigan, P. L. Harris, D. R. Driscoll, P. Fidias, T. J. Lynch, S. K. Rabindran, J. P. McGinnis, A. Wissner, S. V. Sharma, K. J. Isselbacher, J. Settleman, and D. A. Haber. Irreversible inhibitors of the EGF receptor may circumvent acquired resistance to gefitinib. *PNAS*, 2005, 102(21):7665–7670.
- [52] S. K. Rabindran, C. M. Discafani, E. C. Rosfjord, M. Baxter, M. B. Floyd, J. Golas, W. A. Hallett, B. D. Johnson, R. Nilakantan, E. Overbeek, M. F. Reich, R. Shen, X. Shi, H.-R. Tsou, Y.-F. Wang, and A. Wissner. Antitumor Activity of HKI-272, an Orally Active, Irreversible Inhibitor of the HER-2 Tyrosine Kinase. *Cancer Research*, 2004, 64(11): 3958–3965.
- [53] S. Kobayashi, H. Ji, Y. Yuza, M. Meyerson, K.-K. Wong, D. G. Tenen, and B. Halmos. An Alternative Inhibitor Overcomes Resistance Caused by a Mutation of the Epidermal Growth Factor Receptor. *Cancer Research*, 2005, 65(16):7096–7101.
- [54] D. W. Fry, A. J. Kraker, A. McMichael, L. A. Ambroso, J. M. Nelson, W. R. Leopold, R. W. Connors, and A. J. Bridges. A specific inhibitor of the epidermal growth factor receptor tyrosine kinase. *Science*, 1994, 265(5175):1093–1095.
- [55] T. Tokoroyama. Discovery of the Michael Reaction. *Eur. J. Org. Chem.*, 2010, 2010(10): 2009–2016.
- [56] J. V. Heymach, M. Nilsson, G. Blumenschein, V. Papadimitrakopoulou, and R. Herbst. *Clin. Cancer Res*.
- [57] E. Felip, M. Santarpia, and R. Rosell. Emerging drugs for non-small-cell lung cancer. *Expert Opin. Emerging Drugs*, 2007, 12(3):449–460.
- [58] J. F. Ohren, H. Chen, A. Pavlovsky, C. Whitehead, E. Zhang, P. Kuffa, C. Yan, P. McConnell, C. Spessard, C. Banotai, W. T. Mueller, A. Delaney, C. Omer, J. Sebolt-Leopold, D. T. Dudley, I. K. Leung, C. Flamme, J. Warmus, M. Kaufman, S. Barrett, H. Tecle, and C. A. Hasemann. Structures of human MAP kinase kinase 1 (MEK1) and MEK2 describe novel noncompetitive kinase inhibition. *Nature Structural & Molecular Biology*, 2004, 11(12):1192–1197.
- [59] P. S. Klein and D. A. Melton. A molecular mechanism for the effect of lithium on development. *PNAS*, 1996, 93(16):8455–8459.

- [60] E. Zwick, J. Bange, and A. Ullrich. Receptor tyrosine kinase signalling as a target for cancer intervention strategies. *Endocrine-related cancer*, 2001, 8(3):161–173.
- [61] J. Schlessinger. Cell Signaling by Receptor Tyrosine Kinases. *Cell*, 2000, 103(2): 211–225.
- [62] S. Koch, S. Tugues, X. Li, L. Gualandi, and L. Claesson-Welsh. Signal transduction by vascular endothelial growth factor receptors. *Biochem. J.*, 2011, 437(2):169–183.
- [63] R. L. Kendall and K. A. Thomas. Inhibition of vascular endothelial cell growth factor activity by an endogenously encoded soluble receptor. *PNAS*, 1993, 90(22): 10705–10709.
- [64] R. J. Albuquerque, T. Hayashi, W. G. G. Cho, M. E. Kleinman, S. Dridi, A. Takeda, J. Z. Baffi, K. Yamada, H. Kaneko, M. G. Green, J. Chappell, J. Wilting, H. A. Weich, S. Yamagami, S. Amano, N. Mizuki, J. S. Alexander, M. L. Peterson, R. A. Brekken, M. Hirashima, S. Capoor, T. Usui, B. K. Ambati, and J. Ambati. Alternatively spliced vascular endothelial growth factor receptor-2 is an essential endogenous inhibitor of lymphatic vessel growth. *Nature medicine*, 2009, 15(9):1023–1030.
- [65] O. Aprelikova, K. Pajusola, J. Partanen, E. Armstrong, R. Alitalo, S. K. Bailey, J. McMahon, J. Wasmuth, K. Huebner, and K. Alitalo. FLT4, a novel class III receptor tyrosine kinase in chromosome 5q33-qter. *Cancer research*, 1992, 52(3):746–748.
- [66] I. Nilsson, F. Bahram, X. Li, L. Gualandi, S. Koch, M. Jarvius, O. Söderberg, A. Anisimov, I. Kholová, B. Pytowski, M. Baldwin, S. Ylä-Herttuala, K. Alitalo, J. Kreuger, and L. Claesson-Welsh. VEGF receptor 2/3 heterodimers detected in situ by proximity ligation on angiogenic sprouts. *The EMBO journal*, 2010, 29(8):1377–1388.
- [67] Y. Yang, P. Xie, Y. Opatowsky, and J. Schlessinger. Direct contacts between extracellular membrane-proximal domains are required for VEGF receptor activation and cell signaling. *PNAS*, 2010, 107(5):1906–1911.
- [68] C. Ruch, G. Skiniotis, M. O. Steinmetz, T. Walz, and K. Ballmer-Hofer. Structure of a VEGF-VEGF receptor complex determined by electron microscopy. *Nat Struct Mol Biol*, 2007, 14(3):249–250.
- [69] D. Dell’Era Dosch and K. Ballmer-Hofer. Transmembrane domain-mediated orientation of receptor monomers in active VEGFR-2 dimers. *The FASEB Journal*, 2010, 24(1): 32–38.
- [70] T. T. Batchelor, D. G. Duda, E. di Tomaso, M. Ancukiewicz, S. R. Plotkin, E. Gerstner, A. F. Eichler, J. Drappatz, F. H. Hochberg, T. Benner, D. N. Louis, K. S. Cohen, H. Chea, A. Exarhopoulos, J. S. Loeffler, M. A. Moses, P. Ivy, A. G. Sorensen, P. Y. Wen, and R. K. Jain. Phase II study of cediranib, an oral pan-vascular endothelial growth factor receptor tyrosine kinase inhibitor, in patients with recurrent glioblastoma. *J. Clin. Oncol.*, 2010, 28(17):2817–2823.
- [71] D. A. Reardon, S. Turner, K. B. Peters, A. Desjardins, S. Gururangan, J. H. Sampson, R. E. McLendon, J. E. Herndon, L. W. Jones, J. P. Kirkpatrick, A. H. Friedman, J. J. Vredenburgh, D. D. Bigner, and H. S. Friedman. A review of VEGF/VEGFR-targeted therapeutics for recurrent glioblastoma. *Journal of the National Comprehensive Cancer Network : JNCCN*, 2011, 9(4):414–427.

- [72] R. Herbst. Review of epidermal growth factor receptor biology. *Int. J. Radiat. Oncol. Biol. Phys.*, 2004, 59(2):S21–S26.
- [73] H. Zhang, A. Berezov, Q. Wang, G. Zhang, J. Drebin, R. Murali, and M. I. Greene. ErbB receptors: from oncogenes to targeted cancer therapies. *J. Clin. Invest.*, 2007, 117(8): 2051–2058.
- [74] S. Goel, M. Hidalgo, and R. Perez-Soler. EGFR inhibitor-mediated apoptosis in solid tumors. *JETO*, 2007, 6(4):305–320.
- [75] J. Fallon, S. Reid, R. Kinyamu, I. Opole, R. Opole, J. Baratta, M. Korc, T. L. Endo, A. Duong, G. Nguyen, M. Karkehabadhi, D. Twardzik, S. Patel, and S. Loughlin. In vivo induction of massive proliferation, directed migration, and differentiation of neural cells in the adult mammalian brain. *PNAS*, 2000, 97(26):14686–14691.
- [76] J. Downward, P. Parker, and M. D. Waterfield. Autophosphorylation sites on the epidermal growth factor receptor. *Nature*, 1984, 311(5985):483–485.
- [77] K. Oda, Y. Matsuoka, A. Funahashi, and H. Kitano. A comprehensive pathway map of epidermal growth factor receptor signaling. *Mol. Sys. Biol.*, 2005, 1(1):1–17.
- [78] K. Umekawa, T. Kimura, S. Kudoh, T. Suzumura, M. Nagata, S. Mitsuoka, K. Matsuura, T. Oka, N. Yoshimura, Y. Kira, and K. Hirata. Reaction of plasma adiponectin level in non-small cell lung cancer patients treated with EGFR-TKIs. *Osaka city medical journal*, 2013, 59(1):53–60.
- [79] Y. Shi, L. Zhang, X. Liu, C. Zhou, L. Zhang, S. Zhang, D. Wang, Q. Li, S. Qin, C. Hu, Y. Zhang, J. Chen, Y. Cheng, J. Feng, H. Zhang, Y. Song, Y.-L. L. Wu, N. Xu, J. Zhou, R. Luo, C. Bai, Y. Jin, W. Liu, Z. Wei, F. Tan, Y. Wang, L. Ding, H. Dai, S. Jiao, J. Wang, L. Liang, W. Zhang, and Y. Sun. Icotinib versus gefitinib in previously treated advanced non-small-cell lung cancer (ICOGEN): a randomised, double-blind phase 3 non-inferiority trial. *The lancet oncology*, 2013, 14(10):953–961.
- [80] D. S. Meinbach and B. L. Lokeshwar. Insulin-like growth factors and their binding proteins in prostate cancer: Cause or consequence? *Urologic Oncology: Seminars and Original Investigations*, 2006, 24(4):294–306.
- [81] R. S. Garofalo and O. M. Rosen. Insulin and insulinlike growth factor 1 (IGF-1) receptors during central nervous system development: expression of two immunologically distinct IGF-1 receptor beta subunits. *Mol. Cell. Biol.*, 1989, 9(7):2806–2817.
- [82] D. LeRoith, H. Werner, D. Beitner-Johnson, and C. T. Roberts Jr. Molecular and Cellular Aspects of the Insulin-Like Growth Factor I Receptor. *Endocrine Reviews*, 1995, 16(2): 143–163.
- [83] C. W. Gregory, A. Begeorges, and R. A. Sikes. *Recent research developments in cancer*, volume 3.
- [84] J. I. Jones and D. R. Clemmons. Insulin-like growth factors and their binding proteins: biological actions. *Endocrine reviews*, 1995, 16(1):3–34.
- [85] L. Mohanraj and Y. Oh. Targeting IGF-I, IGF-BPs and IGF-I receptor system in cancer: the current and future in breast cancer therapy. *Recent patents on anti-cancer drug discovery*, 2011, 6(2):166–177.

- [86] E. Himpe, S. Potikanond, P. Verdood, and R. Kooijman. Attenuation of IGF-I receptor signaling inhibits serum-induced proliferation of prostate cancer cells. *Growth hormone & IGF research*, 2011, 21(5):252–259.
- [87] K. Gately, I. Collins, L. Forde, B. Al-Alao, V. Young, M. Gerg, F. Feuerhake, and K. O'Byrne. A Role for IGF-1R-Targeted Therapies in Small-Cell Lung Cancer? *Clinical Lung Cancer*, 2011, 12(1):38–42.
- [88] F. Liu, J. J. Stanton, Z. Wu, and H. Piwnica-Worms. The human Myt1 kinase preferentially phosphorylates Cdc2 on threonine 14 and localizes to the endoplasmic reticulum and Golgi complex. *Mol. Cell. Biol.*, 1997, 17(2):571–583.
- [89] N. Watanabe, M. Broome, and T. Hunter. Regulation of the human WEE1Hu CDK tyrosine 15-kinase during the cell cycle. *The EMBO journal*, 1995, 14(9):1878–1891.
- [90] J. Gautier, M. J. Solomon, R. N. Booher, Bazan, and M. W. Kirschner. cdc25 is a specific tyrosine phosphatase that directly activates p34cdc2. *Cell*, 1991, 67(1):197–211.
- [91] G. Cooper. *The Cell: A Molecular Approach*. Sunderland (MA) Sinauer Associates,
- [92] E. Metzger, N. Yin, M. Wissmann, N. Kunowska, K. Fischer, N. Friedrichs, D. Patnaik, J. M. Higgins, N. Potier, K.-H. H. Scheidtmann, R. Buettner, and R. Schüle. Phosphorylation of histone H3 at threonine 11 establishes a novel chromatin mark for transcriptional regulation. *Nature Cell Biol.*, 2008, 10(1):53–60.
- [93] A. Schmidt, J. Durgan, A. Magalhaes, and A. Hall. Rho GTPases regulate PRK2/PKN2 to control entry into mitosis and exit from cytokinesis. *The EMBO Journal*, 2007, 26(6):1624–1636.
- [94] N. E. Torbett, A. Casamassima, and P. J. Parker. Hyperosmotic-induced Protein Kinase N 1 Activation in a Vesicular Compartment Is Dependent upon Rac1 and 3-Phosphoinositide-dependent Kinase 1. *J. Biol. Chem.*, 2003, 278(34):32344–32351.
- [95] H. Takagi, C.-P. Hsu, K. Kajimoto, D. Shao, Y. Yang, Y. Maejima, P. Zhai, G. Yehia, C. Yamada, D. Zablocki, and J. Sadoshima. Activation of PKN Mediates Survival of Cardiac Myocytes in the Heart During Ischemia/Reperfusion / Novelty and Significance. *Circulation Research*, 2010, 107(5):642–649.
- [96] P. J. Hurd, A. J. Bannister, K. Halls, M. A. Dawson, M. Vermeulen, J. V. Olsen, H. Ismail, J. Somers, M. Mann, T. Owen-Hughes, I. Gout, and T. Kouzarides. Phosphorylation of histone H3 Thr-45 is linked to apoptosis. *J. Biol. Chem.*, 2009, 284(24):16575–16583.
- [97] A. Fischer, H. Stuckas, M. Gluth, T. D. Russell, M. C. Rudolph, N. E. Beeman, S. Bachmann, S. Umemura, Y. Ohashi, M. C. Neville, and F. Theuring. Impaired tight junction sealing and precocious involution in mammary glands of PKN1 transgenic mice. *J. Cell Sci.*, 2007, 120(13):2272–2283.
- [98] H. Mukai. The Structure and Function of PKN, a Protein Kinase Having a Catalytic Domain Homologous to That of PKC. *Journal of Biochemistry*, 2003, 133(1):17–27.
- [99] E. Smythe and K. R. Ayscough. The Ark1/Prk1 family of protein kinases. Regulators of endocytosis and the actin skeleton. *EMBO reports*, 2003, 4(3):246–251.
- [100] P. Ripphausen, B. Nisius, L. Peltason, and J. Bajorath. Quo Vadis, Virtual Screening? A Comprehensive Survey of Prospective Applications. *J. Med. Chem.*, 2010, 53(24):8461–8467.

- [101] P. D. Lyne. Structure-based virtual screening: an overview. *Drug Discovery Today*, 2002, 7(20):1047–1055.
- [102] I. Kuntz. A geometric approach to macromolecule-ligand interactions*1. *J. Mol. Biol.*, 1982, 161(2):269–288.
- [103] B. K. Shoichet and I. D. Kuntz. Matching chemistry and shape in molecular docking. *Protein Engineering*, 1993, 6(7):723–732.
- [104] D. Whitley. A genetic algorithm tutorial. *Statistics and Computing*, 1994, 4(2):65–85.
- [105] J. Kennedy and R. Eberhart. Particle swarm optimization. In *Neural Networks, 1995. Proceedings., IEEE International Conference on*, volume 4, pages 1942–1948 vol.4. ISBN 0-7803-2768-3.
- [106] T. N. Hart and R. J. Read. A multiple-start Monte Carlo docking method. *Proteins*, 1992, 13(3):206–222.
- [107] M. Dorigo, V. Maniezzo, and A. Colorni. Ant system: optimization by a colony of cooperating agents. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 1996, 26(1):29–41.
- [108] G. Jones. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.*, 1997, 267(3):727–748.
- [109] H.-M. Chen, B.-F. Liu, H.-L. Huang, S.-F. Hwang, and S.-Y. Ho. SODOCK: Swarm optimization for highly flexible protein-ligand docking. *J. Comput. Chem.*, 2007, 28(2): 612–623.
- [110] *Glide, version 5.6*. Schrödinger, LLC, 2010.
- [111] M. K. Gilson, J. A. Given, B. L. Bush, and J. A. McCammon. The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophysical journal*, 1997, 72(3):1047–1069.
- [112] P. R. Andrews, D. J. Craik, and J. L. Martin. Functional group contributions to drug-receptor interactions. *J. Med. Chem.*, 1984, 27(12):1648–1657.
- [113] R. Rajamani and A. C. Good. Ranking poses in structure-based lead discovery and optimization: Current trends in scoring function development. *Curr. Opin. Drug Discov. Dev.*, 2007, 10(3):308–315.
- [114] M. H. Seifert, J. Kraus, and B. Kramer. Virtual high-throughput screening of molecular databases. *Curr. Opin. Drug Discov. Dev.*, 2007, 10(3):298–307.
- [115] M. Stahl. Structure-Based Library Design. 2008, pages 229–264.
- [116] A. Warshel, P. K. Sharma, M. Kato, and W. W. Parson. Modeling electrostatic effects in proteins. *Biochimica et biophysica acta*, 2006, 1764(11):1647–1676.
- [117] J. W. Ponder and D. A. Case. *Force Fields for Protein Simulations*, volume 66 of *Advances in Protein Chemistry*, pages 27–85. ISBN 9780120342662.
- [118] A. D. MacKerell, J. Wiorkiewicz-Kuczera, and M. Karplus. An all-atom empirical energy function for the simulation of nucleic acids. *J. Am. Chem. Soc.*, 1995, 117(48): 11946–11975.

-
- [119] A. D. MacKerell, D. Bashford, Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins†. *J. Phys. Chem. B*, 1998, 102(18):3586–3616.
- [120] S. N. Ha, A. Giammona, M. Field, and J. W. Brady. A revised potential-energy surface for molecular mechanics studies of carbohydrates. *Carbohydrate Research*, 1988, 180(2):207–221.
- [121] S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, and P. Weiner. A new force field for molecular mechanical simulation of nucleic acids and proteins. *Journal of the American Chemical Society*, 1984, 106(3):765–784.
- [122] S. J. Weiner, P. A. Kollman, D. T. Nguyen, and D. A. Case. An all atom force field for simulations of proteins and nucleic acids. *J. Comput. Chem.*, 1986, 7(2):230–252.
- [123] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.*, 1983, 4(2):187–217.
- [124] E. C. Meng, B. K. Shoichet, and I. D. Kuntz. Automated docking with grid-based energy evaluation. *J. Comput. Chem.*, 1992, 13(4):505–524.
- [125] H.-J. Böhm. Prediction of binding constants of protein ligands: A fast method for the prioritization of hits obtained from de novo design or 3D database search programs. *Journal of Computer-Aided Molecular Design*, 1998, 12(4):309.
- [126] H.-J. Böhm. The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure. *Journal of Computer-Aided Molecular Design*, 1994, 8(3):243–256.
- [127] D. A. Mcquarrie. *Statistical Mechanics*.
- [128] S.-Y. Huang and X. Zou. *Mean-Force Scoring Functions for Protein-Ligand Binding*, volume 6, pages 280–296. ISBN 9780444535528.
- [129] S. Tanaka and H. A. Scheraga. Medium- and long-range interaction parameters between amino acids for predicting three-dimensional structures of proteins. *Macromolecules*, 1976, 9(6):945–950.
- [130] S. Miyazawa and R. L. Jernigan. Estimation of effective interresidue contact energies from protein crystal structures: quasi-chemical approximation. *Macromolecules*, 1985, 18(3):534–552.
- [131] M. J. Sippl. Calculation of conformational ensembles from potentials of mean force: An approach to the knowledge-based prediction of local structures in globular proteins. *Journal of Molecular Biology*, 1990, 213(4):859–883.
- [132] M. Hendlich, P. Lackner, S. Weitkus, H. Floeckner, R. Froschauer, K. Gottsbacher, G. Casari, and M. Sippl. Identification of native protein folds amongst a large number of incorrect models: The calculation of low energy conformations from potentials of mean force. *Journal of Molecular Biology*, 1990, 216(1):167–180.

- [133] D. T. Jones, W. R. Taylor, and J. M. Thornton. A new approach to protein fold recognition. *Nature*, 1992, 358(6381):86–89.
- [134] P. D. Thomas and K. A. Dill. An iterative method for extracting energy-like quantities from protein structures. *Proceedings of the National Academy of Sciences of the United States of America*, 1996, 93(21):11628–11633.
- [135] I. Muegge. PMF Scoring Revisited. *Journal of Medicinal Chemistry*, 2006, 49(20):5895–5902.
- [136] H. Gohlke, M. Hendlich, and G. Klebe. Knowledge-based scoring function to predict protein-ligand interactions. *Journal of Molecular Biology*, 2000, 295(2):337–356.
- [137] H. F. G. Velec, H. Gohlke, and G. Klebe. DrugScoreCSDKnowledge-Based Scoring Function Derived from Small Molecule Crystal Data with Superior Recognition Rate of Near-Native Ligand Poses and Better Affinity Prediction. *J. Med. Chem.*, 2005, 48(20):6296–6303.
- [138] R. Clark. Consensus scoring for ligand/protein interactions. *Journal of Molecular Graphics and Modelling*, 2002, 20(4):281–295.
- [139] R. Teramoto and H. Fukunishi. Supervised Consensus Scoring for Docking and Virtual Screening. *J. Chem. Inf. Model.*, 2007, 47(2):526–534.
- [140] C. Bissantz, G. Folkers, and D. Rognan. Protein-Based Virtual Screening of Chemical Databases. 1. Evaluation of Different Docking/Scoring Combinations. *J. Med. Chem.*, 2000, 43(25):4759–4767.
- [141] R. Wang and S. Wang. How Does Consensus Scoring Work for Virtual Library Screening? An Idealized Computer Experiment. *J. Chem. Inf. Comput. Sci.*, 2001, 41(5):1422–1426.
- [142] C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. PII of original article: S0169-409X(96)00423-1. The article was originally published in *Advanced Drug Delivery Reviews* 23 (1997) 3-25.1. *Advanced Drug Delivery Reviews*, 2001, 46(1-3):3–26.
- [143] A. K. Ghose, V. N. Viswanadhan, and J. J. Wendoloski. A Knowledge-Based Approach in Designing Combinatorial or Medicinal Chemistry Libraries for Drug Discovery. 1. A Qualitative and Quantitative Characterization of Known Drug Databases. *J. Com. Chem.*, 1999, 1(1):55–68.
- [144] H.-D. Höltje, W. Sippl, D. Rognan, and G. Folkers. *Molecular Modeling - Basic Principles and Applications*. WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim,
- [145] G. Klebe. *Wirkstoffdesign*. Spektrum Akademischer Verlag Heidelberg,
- [146] A. Bender and R. C. Glen. Molecular similarity: a key technique in molecular informatics. *Org. Biomol. Chem.*, 2004, 2(22):3204–3218.
- [147] T. R. Hagadone. Molecular substructure similarity searching: efficient retrieval in two-dimensional structure databases. *J. Chem. Inf.*, 1992, 32(5):515–521.

- [148] D. J. Wild and P. Willett. Similarity Searching in Files of Three-Dimensional Chemical Structures. Alignment of Molecular Electrostatic Potential Fields with a Genetic Algorithm. *J. Chem. Inf.*, 1996, 36(2):159–167.
- [149] J. Mestres, D. C. Rohrer, and G. M. Maggiora. MIMIC: A molecular-field matching program. Exploiting applicability of molecular similarity approaches. *J. Comput. Chem.*, 1997, 18(7):934–954.
- [150] S. Kearsley and G. Smith. An alternative method for the alignment of molecular structures: Maximizing electrostatic and steric overlap. *Tetrahedron Computer Methodology*, 1990, 3(6):615–633.
- [151] J. Mount, J. Ruppert, W. Welch, and A. N. Jain. IcePick: A Flexible Surface-Based System for Molecular Diversity. *J. Med. Chem.*, 1999, 42(1):60–66.
- [152] P. G. Mezey. *Shape in Chemistry*. WILEY-VCH Verlag,
- [153] J. A. Grant and B. T. Pickup. A Gaussian Description of Molecular Shape. *J. Phys. Chem.*, 1995, 99(11):3503–3510.
- [154] D. J. Livingstone. ChemInform Abstract: The Characterization of Chemical Structures Using Molecular Properties. A Survey. *ChemInform*, 2000, 31(23).
- [155] B. Cuissart, F. Touffet, B. Crémilleux, R. Bureau, and S. Rault. The Maximum Common Substructure as a Molecular Depiction in a Supervised Classification Context: Experiments in Quantitative Structure/Biodegradability Relationships. *J. Chem. Inf. Comput. Sci.*, 2002, 42(5):1043–1052.
- [156] J. Gasteiger, H. Lanig, and T. Clark. *Cheminformatics: a textbook*, chapter Calculation of Physical and Chemical Data, pages 319–400. Wiley-VCH Verlag GmbH & Co. KGaA,
- [157] N. E. Shemetulskis, D. Weininger, C. J. Blankley, J. J. Yang, and C. Humblet. Stigmata: An Algorithm To Determine Structural Commonalities in Diverse Datasets. *J. Chem. Inf. Comput. Sci.*, 1996, 36(4):862–871.
- [158] D. J. Rogers and T. T. Tanimoto. A Computer Program for Classifying Plants. *Science*, 1960, 132(3434):1115–1118.
- [159] P. Willett, J. M. Barnard, and G. M. Downs. Chemical Similarity Searching. *J. Chem. Inf. Comput. Sci.*, 1998, 38(6):983–996.
- [160] A. Maldonado, J. Doucet, M. Petitjean, and B.-T. Fan. Molecular similarity and diversity in chemoinformatics: From theory to applications. *Molecular Diversity*, 2006, 10(1): 39–79.
- [161] M. J. McGregor and P. V. Pallai. Clustering of Large Databases of Compounds: Using the MDL Keys as Structural Descriptors. *J. Chem. Inf. Comp. Sci.*, 1997, 37(3):443–448.
- [162] *ROCS, version 3.1.2*. OpenEye Scientific Software Inc., 2010.
- [163] N. Huang, C. Kalyanaraman, K. Bernacki, and M. P. Jacobson. Molecular mechanics methods for predicting protein-ligand binding. *Phys. Chem. Chem. Phys.*, 2006, 8(44): 5166–5177.

- [164] W. L. Jorgensen and J. Tirado-Rives. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *PNAS*, 2005, 102(19): 6665–6670.
- [165] A. Nicholls and B. Honig. A rapid finite difference algorithm, utilizing successive over-relaxation to solve the Poisson-Boltzmann equation. *J. Comput. Chem.*, 1991, 12(4):435–445.
- [166] V. Tsui and D. A. Case. Theory and applications of the generalized born solvation model in macromolecular simulations. *Biopolymers*, 2000, 56(4):275–291.
- [167] A. Ghosh, C. S. Rapp, and R. A. Friesner. Generalized Born Model Based on a Surface Integral Formulation. *J. Phys. Chem.*, 1998, 102(52):10983–10990.
- [168] I. D. Kuntz, E. C. Meng, and B. K. Shoichet. Structure-Based Molecular Design. *Acc. Chem. Res.*, 1994, 27(5):117–123.
- [169] D. A. Pearlman and P. S. Charifson. Are Free Energy Calculations Useful in Practice? A Comparison with Rapid Scoring Functions for the p38 MAP Kinase Protein System†. *J. Med. Chem.*, 2001, 44(21):3417–3423.
- [170] J. Wang, P. Morin, W. Wang, and P. A. Kollman. Use of MM-PBSA in Reproducing the Binding Free Energies to HIV-1 RT of TIBO Derivatives and Predicting the Binding Mode to HIV-1 RT of Efavirenz by Docking and MM-PBSA. *Journal of the American Chemical Society*, 2001, 123(22):5221–5230.
- [171] M. R. Reddy and M. D. Erion. Calculation of Relative Binding Free Energy Differences for Fructose 1,6-Bisphosphatase Inhibitors Using the Thermodynamic Cycle Perturbation Approach. *JACS*, 2001, 123(26):6246–6252.
- [172] S. Homans. Water, water everywhere - except where it matters? *Drug Discovery Today*, 2007, 12(13-14):534–539.
- [173] P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case, and T. E. Cheatham. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.*, 2000, 33(12):889–897.
- [174] H. Gohlke and G. Klebe. Approaches to the Description and Prediction of the Binding Affinity of Small-Molecule Ligands to Macromolecular Receptors. *Angewandte Chemie International Edition*, 2002, 41(15):2644–2676.
- [175] D. Sitkoff, K. A. Sharp, and B. Honig. Accurate Calculation of Hydration Free Energies Using Macroscopic Solvent Models. *J. Phys. Chem.*, 1994, 98(7):1978–1988.
- [176] J. Srinivasan, T. E. Cheatham, P. Cieplak, P. A. Kollman, and D. A. Case. Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate-DNA Helices. *JACS*, 1998, 120(37):9401–9409.
- [177] W. C. Still, A. Tempczyk, R. C. Hawley, and T. Hendrickson. Semianalytical treatment of solvation for molecular mechanics and dynamics. *JACS*, 1990, 112(16):6127–6129.
- [178] A. Onufriev, D. A. Case, and D. Bashford. Effective Born radii in the generalized Born approximation: The importance of being perfect. *J. Comput. Chem.*, 2002, 23(14): 1297–1304.

- [179] B. K. Ho and K. A. Dill. Folding Very Short Peptides Using Molecular Dynamics. *PLoS Comput Biol*, 2006, 2(4):228–237.
- [180] R. Zhou. Free energy landscape of protein folding in water: Explicit vs. implicit solvent. *Proteins*, 2003, 53(2):148–161.
- [181] Z. Zhou and J. D. Madura. Relative free energy of binding and binding mode calculations of HIV-1 RT inhibitors based on dock-MM-PB/GS. *Proteins*, 2004, 57(3):493–503.
- [182] Z. Zhou, M. Bates, and J. D. Madura. Structure modeling, ligand binding, and binding affinity calculation (LR-MM-PBSA) of human heparanase for inhibition and drug design. *Proteins*, 2006, 65(3):580–592.
- [183] K. Wichapong, M. Lawson, S. Pianwanit, S. Kokpol, and W. Sippl. Postprocessing of Protein-Ligand Docking Poses Using Linear Response MM-PB/SA: Application to Wee1 Kinase Inhibitors. *J. Chem. Inf. Model.*, 2010, 50(9):1574–1588.
- [184] N. Huang, B. K. Shoichet, and J. J. Irwin. Benchmarking Sets for Molecular Docking. *Journal of Medicinal Chemistry*, 2006, 49(23):6789–6801.
- [185] D. J. Hand and R. J. Till. A Simple Generalisation of the Area Under the ROC Curve for Multiple Class Classification Problems. *Machine Learning*, 2001, 45(2):171–186.
- [186] R. P. Sheridan, S. B. Singh, E. M. Fluder, and S. K. Kearsley. Protocols for Bridging the Peptide to Nonpeptide Gap in Topological Similarity Searches. *J. Chem. Inf. Comput. Sci.*, 2001, 41(5):1395–1406.
- [187] B. Q. Wei, W. A. Baase, L. H. Weaver, B. W. Matthews, and B. K. Shoichet. A Model Binding Site for Testing Scoring Functions in Molecular Docking. *J. Mol. Biol.*, 2002, 322(2):339–355.
- [188] *Molecular Operating Environment (MOE) 2010.10*. Chemical Computing Group Inc., 2010.
- [189] J. Wang, P. Cieplak, and P. A. Kollman. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.*, 2000, 21(12):1049–1074.
- [190] R. Meier, M. Pippel, F. Brandt, W. Sippl, and C. Baldauf. ParaDockS: A Framework for Molecular Docking with Population-Based Metaheuristics. *J. Chem. Inf. Model.*, 2010, 50(5):879–889.
- [191] D. Moustakas, Lang, S. Pegg, E. Pettersen, I. Kuntz, N. Brooijmans, and R. Rizzo. Development and validation of a modular, extensible docking program: DOCK 5. *J. Comput. Aided Mol. Des.*, 2006, 20(10-11):601–619.
- [192] A. P. Graves, D. M. Shivakumar, S. E. Boyce, M. P. Jacobson, D. A. Case, and B. K. Shoichet. Rescoring docking hit lists for model cavity sites: predictions and experimental testing. *J. Mol. Biol.*, 2008, 377(3):914–934.
- [193] D. Case, T. Darden, I. T.E. Cheatham, C. Simmerling, J. Wang, R. Duke, R. Luo, R. Walker, W. Zhang, K. Merz, B. Roberts, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossvai, K. Wong, F. Paesani, J. Vanicek, J. Liu, X. Wu, S. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D. Roe, D. Mathews, M. Seetin,

- C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P. Kollman. *AMBER 11*. 2010.
- [194] T. Förster. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Ann. Phys.*, 1948, 437(1-2):55–75.
- [195] L. Stryer. Fluorescence Energy Transfer as a Spectroscopic Ruler. *Ann. Rev. Biochem.*, 1978, 47(1):819–846.
- [196] C. S. Lebakken, S. M. Riddle, U. Singh, W. J. Frazee, H. C. Eliason, Y. Gao, L. J. Reichling, B. D. Marks, and K. W. Vogel. Development and Applications of a Broad-Coverage, TR-FRET-Based Kinase Binding Assay Platform. *J. Biomol. Screen.*, 2009, 14(8):924–935.
- [197] M. Krug, G. Erlenkamp, W. Sippl, C. Schächtele, F. Totzke, and A. Hilgeroth. Discovery and selectivity-profiling of 4-benzylamino 1-aza-9-oxafluorene derivatives as lead structures for IGF-1R inhibitors. *Bioorg. Med. Chem. Lett.*, 2010, 20(23):6915–6919.
- [198] W.-S. Huang, X. Zhu, Y. Wang, M. Azam, D. Wen, R. Sundaramoorthi, R. M. Thomas, S. Liu, G. Banda, S. P. Lentini, S. Das, Q. Xu, J. Keats, F. Wang, S. Wardwell, Y. Ning, J. T. Snodgrass, M. I. Broudy, K. Russian, G. Q. Daley, J. Iulucci, D. C. Dalgarno, T. Clackson, T. K. Sawyer, and W. C. Shakespeare. 9-(Arenethenyl)purines as Dual Src/Abl Kinase Inhibitors Targeting the Inactive Conformation: Design, Synthesis, and Biological Evaluation. *J. Med. Chem.*, 2009, 52(15):4743–4756.
- [199] I. Aliagas-Martin, D. Burdick, L. Corson, J. Dotson, J. Drummond, C. Fields, O. W. Huang, T. Hunsaker, T. Kleinheinz, E. Krueger, J. Liang, J. Moffat, G. Phillips, R. Pulk, T. E. Rawson, M. Ultsch, L. Walker, C. Wiesmann, B. Zhang, B.-Y. Zhu, and A. G. Cochran. A Class of 2,4-Bisanilino-pyrimidine Aurora A Inhibitors with Unusually High Selectivity against Aurora B †. *J. Med. Chem.*, 2009, 52(10):3300–3307.
- [200] M. S. Coumar, J.-S. Leou, P. Shukla, J.-S. Wu, A. K. Dixit, W.-H. Lin, C.-Y. Chang, T.-W. Lien, U.-K. Tan, C.-H. Chen, J. T. A. Hsu, Y.-S. Chao, S.-Y. Wu, and H.-P. Hsieh. Structure-Based Drug Design of Novel Aurora Kinase A Inhibitors: Structural Basis for Potency and Specificity. *J. Med. Chem.*, 2009, 52(4):1050–1062.
- [201] S. Howard, V. Berdini, J. A. Boulstridge, M. G. Carr, D. M. Cross, J. Curry, L. A. Devine, T. R. Early, L. Fazal, A. L. Gill, M. Heathcote, S. Maman, J. E. Matthews, R. L. McMenamin, E. F. Navarro, M. A. O'Brien, M. O'Reilly, D. C. Rees, M. Reule, D. Tisi, G. Williams, M. Vinković, and P. G. Wyatt. Fragment-Based Discovery of the Pyrazol-4-yl Urea (AT9283), a Multitargeted Kinase Inhibitor with Potent Aurora Kinase Activity †. *J. Med. Chem.*, 2008, 52(2):379–388.
- [202] M. Menichincheri, A. Bargiotti, J. Berthelsen, J. A. Bertrand, R. Bossi, A. Ciavolella, A. Cirila, C. Cristiani, V. Croci, R. D'Alessio, M. Fasolini, F. Fiorentini, B. Forte, A. Isacchi, K. Martina, A. Molinari, A. Montagnoli, P. Orsini, F. Orzi, E. Pesenti, D. Pezzetta, A. Pillan, I. Poggesi, F. Roletto, A. Scolaro, M. Tatò, M. Tibolla, B. Valsasina, M. Varasi, D. Volpi, C. Santocanale, and E. Vanotti. First Cdc7 kinase inhibitors: pyrrolopyridinones as potent and orally active antitumor agents. 2. Lead discovery. *J. Med. Chem.*, 2009, 52(2):293–307.
- [203] M. G. Brasca, N. Amboldi, D. Ballinari, A. Cameron, E. Casale, G. Cervi, M. Colombo, F. Colotta, V. Croci, R. D'Alessio, F. Fiorentini, A. Isacchi, C. Mercurio, W. Moretti,

- A. Panzeri, W. Pastori, P. Pevarello, F. Quartieri, F. Roletto, G. Traquandi, P. Vianello, A. Vulpetti, and M. Ciomei. Identification of n,1,4,4-tetramethyl-8-[4-(4-methylpiperazin-1-yl)phenyl]amino-4,5-dihydro-1h-pyrazolo[4,3-h]quinazoline-3-carboxamide (pha-848125), a potent, orally available cyclin dependent kinase inhibitor. *J. Med. Chem.*, 2009, 52(16):5152–5163.
- [204] F. Popowycz, G. Fournet, C. Schneider, K. Bettayeb, Y. Ferandin, C. Lamigeon, O. M. Tirado, S. Mateo-Lozano, V. Notario, P. Colas, P. Bernard, L. Meijer, and B. Joseph. Pyrazolo[1,5-a]-1,3,5-triazine as a Purine Bioisostere: Access to Potent Cyclin-Dependent Kinase Inhibitor (R)-Roscovitine Analogue. *J. Med. Chem.*, 2009, 52(3):655–663.
- [205] H.-R. Tsou, X. Liu, G. Birnberg, J. Kaplan, M. Otteng, T. Tran, K. Kutterer, Z. Tang, R. Suayan, A. Zask, M. Ravi, A. Bretz, M. Grillo, J. P. McGinnis, S. K. Rabindran, S. Ayril-Kaloustian, and T. S. Mansour. Discovery of 4-(Benzylaminomethylene)isoquinoline-1,3-(2H,4H)-diones and 4-[(Pyridylmethyl)aminomethylene]isoquinoline-1,3-(2H,4H)-diones as Potent and Selective Inhibitors of the Cyclin-Dependent Kinase 4. *J. Med. Chem.*, 2009, 52(8):2289–2310.
- [206] A. Ermoli, A. Bargiotti, M. G. Brasca, A. Ciavolella, N. Colombo, G. Fachin, A. Isacchi, M. Menichincheri, A. Molinari, A. Montagnoli, A. Pillan, S. Rainoldi, F. R. Sirtori, F. Sola, S. Thieffine, M. Tibolla, B. Valsasina, D. Volpi, C. Santocanale, and E. Vanotti. Cell Division Cycle 7 Kinase Inhibitors: 1 H -Pyrrolo[2,3- b]pyridines, Synthesis and Structure-Activity Relationships. *J. Med. Chem.*, 2009, 52(14):4380–4390.
- [207] T. P. Matthews, S. Klair, S. Burns, K. Boxall, M. Cherry, M. Fisher, I. M. Westwood, M. I. Walton, T. McHardy, K.-M. J. Cheung, R. Van Montfort, D. Williams, G. W. Aherne, M. D. Garrett, J. Reader, and I. Collins. Identification of Inhibitors of Checkpoint Kinase 1 through Template Screening. *Journal of Medicinal Chemistry*, 2009, 52(15):4810–4819.
- [208] C. Peifer, M. Abadleh, J. Bischof, D. Hauser, V. Schattel, H. Hirner, U. Knippschild, and S. Laufer. 3,4-Diaryl-isoxazoles and -imidazoles as Potent Dual Inhibitors of p38 α Mitogen Activated Protein Kinase and Casein Kinase 1 δ †. *J. Med. Chem.*, 2009, 52(23):7618–7630.
- [209] S. K. De, J. L. Stebbins, L.-H. Chen, M. Riel-Mehan, T. Machleidt, R. Dahl, H. Yuan, A. Emdadi, E. Barile, V. Chen, R. Murphy, and M. Pellicchia. Design, Synthesis, and Structure-Activity Relationship of Substrate Competitive, Selective, and in Vivo Active Triazole and Thiadiazole Inhibitors of the c-Jun N-Terminal Kinase. *J. Med. Chem.*, 2009, 52(7):1943–1952.
- [210] I. N. Gaisina, F. Gallier, A. V. Ougolkov, K. H. Kim, T. Kurome, S. Guo, D. Holzle, D. N. Luchini, S. Y. Blond, D. D. Billadeau, and A. P. Kozikowski. From a Natural Product Lead to the Identification of Potent and Selective Benzofuran-3-yl-(indol-3-yl)maleimides as Glycogen Synthase Kinase 3 β Inhibitors That Suppress Proliferation and Survival of Pancreatic Cancer Cells. *J. Med. Chem.*, 2009, 52(7):1853–1863.
- [211] G. M. Schroeder, Y. An, Z.-W. Cai, X.-T. Chen, C. Clark, L. A. M. Cornelius, J. Dai, J. Gullo-Brown, A. Gupta, B. Henley, J. T. Hunt, R. Jeyaseelan, A. Kamath, K. Kim, J. Lippy, L. J. Lombardo, V. Manne, S. Oppenheimer, J. S. Sack, R. J. Schmidt, G. Shen, K. Stefanski, J. S. Tokarski, G. L. Trainor, B. S. Wautlet, D. Wei, D. K. Williams, Y. Zhang,

- Y. Zhang, J. Fargnoli, and R. M. Borzilleri. Discovery of N-(4-(2-Amino-3-chloropyridin-4-yloxy)-3-fluorophenyl)-4-ethoxy-1-(4-fluorophenyl)-2-oxo-1,2-dihydropyridine-3-carboxamide (BMS-777607), a Selective and Orally Efficacious Inhibitor of the Met Kinase Superfamily. *J. Med. Chem.*, 2009, 52(5): 1251–1254.
- [212] T. J. Cheeseright, M. Holm, F. Lehmann, S. Luik, M. Göttert, J. L. Melville, and S. Laufer. Novel Lead Structures for p38 MAP Kinase via FieldScreen Virtual Screening. *J. Med. Chem.*, 2009, 52(14):4200–4209.
- [213] S. C. Karcher and S. A. Laufer. Aza-Analogue Dibenzepinone Scaffolds as p38 Mitogen-Activated Protein Kinase Inhibitors: Design, Synthesis, and Biological Data of Inhibitors with Improved Physicochemical Properties. *J. Med. Chem.*, 2009, 52(6): 1778–1782.
- [214] K. Qian, L. Wang, C. L. Cywin, B. T. Farmer, E. Hickey, C. Homon, S. Jakes, M. A. Kashem, G. Lee, S. Leonard, J. Li, R. Magboo, W. Mao, E. Pack, C. Peng, A. Prokopowicz, M. Welzel, J. Wolak, and T. Morwick. Hit to Lead Account of the Discovery of a New Class of Inhibitors of Pim Kinases and Crystallographic Studies Revealing an Unusual Kinase Binding Mode†. *J. Med. Chem.*, 2009, 52(7):1814–1827.
- [215] Z. Xia, C. Knaak, J. Ma, Z. M. Beharry, C. McInnes, W. Wang, A. S. Kraft, and C. D. Smith. Synthesis and Evaluation of Novel Inhibitors of Pim-1 and Pim-2 Protein Kinases. *J. Med. Chem.*, 2009, 52(1):74–86.
- [216] P. A. Renhowe, S. Pecchi, C. M. Shafer, T. D. Machajewski, E. M. Jazan, C. Taylor, W. Antonios-McCrea, C. M. McBride, K. Frazier, M. Wiesmann, G. R. Lapointe, P. H. Feucht, R. L. Warne, C. C. Heise, D. Menezes, K. Aardalen, H. Ye, M. He, V. Le, J. Vora, J. M. Jansen, M. E. Wernette-Hammond, and A. L. Harris. Design, Structure-Activity Relationships and in Vivo Characterization of 4-Amino-3-benzimidazol-2-ylhydroquinolin-2-ones: A Novel Class of Receptor Tyrosine Kinase Inhibitors. *J. Med. Chem.*, 2008, 52(2):278–292.
- [217] J. B. Cross, D. C. Thompson, B. K. Rai, J. C. Baber, K. Y. Fan, Y. Hu, and C. Humblet. Comparison of Several Molecular Docking Programs: Pose Prediction and Virtual Screening Accuracy. *J. Chem. Inf. Model.*, 2009, 49(6):1455–1474.
- [218] F. N. Novikov, V. S. Stroylov, A. A. Zeifman, O. V. Stroganov, V. Kulkov, and G. G. Chilov. Lead Finder docking and virtual screening evaluation with Astex and DUD test sets. *J. Comp. Aid. Mol. Des.*, 2012, 26(6):725–735.
- [219] N. Mishra and A. Basu. Exploring different virtual screening strategies for acetylcholinesterase inhibitors. *BioMed research international*, 2013, 2013.
- [220] P. C. D. Hawkins, A. G. Skillman, and A. Nicholls. Comparison of Shape-Matching and Docking as Virtual Screening Tools. *J. Med. Chem.*, 2006, 50(1):74–82.
- [221] R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, and P. S. Shenkin. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.*, 2004, 47(7):1739–1749.

- [222] M. D. Eldridge, C. W. Murray, T. R. Auton, G. V. Paolini, and R. P. Mee. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comp. Aid. Mol. Des.*, 1997, 11(5):425–445.
- [223] P. C. Hawkins, A. G. Skillman, G. L. Warren, B. A. Ellingson, and M. T. Stahl. 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):572–584.
- [224] J.-F. Truchon and C. I. Bayly. Evaluating Virtual Screening Methods: Good and Bad Metrics for the “Early Recognition” Problem. *J. Chem. Inf. Model.*, 2007, 47(2):488–508.
- [225] R. Teramoto and H. Fukunishi. Structure-Based Virtual Screening with Supervised Consensus Scoring: Evaluation of Pose Prediction and Enrichment Factors. *J. Chem. Inf. Model.*, 2008, 48(4):747–754.
- [226] A. Oda, K. Tsuchida, T. Takakura, N. Yamaotsu, and S. Hirono. Comparison of Consensus Scoring Strategies for Evaluating Computational Models of Protein-Ligand Complexes. *J. Chem. Inf. Model.*, 2005, 46(1):380–391.
- [227] J. Köhler, G. Erlenkamp, A. Eberlin, T. Rumpf, I. Slynko, E. Metzger, R. Schüle, W. Sippl, and M. Jung. Lestaurtinib Inhibits Histone Phosphorylation and Androgen-Dependent Gene Expression in Prostate Cancer Cells. *PLoS ONE*, 2012, 7(4):1–8.
- [228] B. Båtshake and J. Sundelin. The Mouse Genes for the EP1Prostanoid Receptor and the PKN Protein Kinase Overlap. *Biochemical and Biophysical Research Communications*, 1996, 227(1):70–76.
- [229] K. Arnold, L. Bordoli, J. Kopp, and T. Schwede. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*, 2006, 22(2):195–201.
- [230] F. Kiefer, K. Arnold, M. Künzli, L. Bordoli, and T. Schwede. The SWISS-MODEL Repository and associated resources. *Nucl. Acid. Res.*, 2009, 37(Database issue): 387–392.
- [231] M. C. Peitsch. Protein Modeling by E-mail. *Nature Biotechnology*, 1995, 13(7):658–660.
- [232] M. Takahashi, H. Mukai, M. Toshimori, M. Miyamoto, and Y. Ono. Proteolytic activation of PKN by caspase-3 or related protease during apoptosis. *PNAS*, 1998, 95(20): 11566–11571.
- [233] P. Tapley, F. Lamballe, and M. Barbacid. K252a is a selective inhibitor of the tyrosine protein kinase activity of the trk family of oncogenes and neurotrophin receptors. *Oncogene*, 1992, 7(2):371–381.
- [234] A. Morotti, S. Mila, P. Accornero, E. Tagliabue, and C. Ponzetto. K252a inhibits the oncogenic properties of Met, the HGF receptor. *Oncogene*, 2002, 21(32):4885–4893.
- [235] L. A. Kirby, J. T. Schott, B. L. Noble, D. C. Mendez, P. S. Caseley, S. C. Peterson, T. J. Routledge, and N. V. Patel. Glycogen Synthase Kinase 3 (GSK3) Inhibitor, SB-216763, Promotes Pluripotency in Mouse Embryonic Stem Cells. *PLoS ONE*, 2012, 7(6):1–13.
- [236] M. P. Coghlan, A. A. Culbert, D. A. Cross, S. L. Corcoran, J. W. Yates, N. J. Pearce, O. L. Rausch, G. J. Murphy, P. S. Carter, L. Roxbee Cox, D. Mills, M. J. Brown, D. Haigh,

- R. W. Ward, D. G. Smith, K. J. Murray, A. D. Reith, and J. C. Holder. Selective small molecule inhibitors of glycogen synthase kinase-3 modulate glycogen metabolism and gene transcription. *Chemistry & biology*, 2000, 7(10):793–803.
- [237] H. C. Kolb, M. G. Finn, and K. B. Sharpless. Click-Chemie: diverse chemische Funktionalität mit einer Handvoll guter Reaktionen. *Angew. Chem.*, 2001, 113(11): 2056–2075.
- [238] A. Rohe, F. Erdmann, C. Bäßler, K. Wichapong, W. Sippl, and M. Schmidt. In vitro and in silico studies on substrate recognition and acceptance of human PKMYT1, a Cdk1 inhibitory kinase. *Bioorg. Med. Chem. Lett.*, 2012, 22(2):1219–1223.
- [239] P. Kolb, D. Huang, F. Dey, and A. Caflisch. Discovery of Kinase Inhibitors by High-Throughput Docking and Scoring Based on a Transferable Linear Interaction Energy Model. *J. Med. Chem.*, 2008, 51(5):1179–1188.
- [240] M. A. Fabian, W. H. Biggs, D. K. Treiber, C. E. Atteridge, M. D. Azimioara, M. G. Benedetti, T. A. Carter, P. Ciceri, P. T. Edeen, M. Floyd, J. M. Ford, M. Galvin, J. L. Gerlach, R. M. Grotzfeld, S. Herrgard, D. E. Insko, M. A. Insko, A. G. Lai, J.-M. Lelias, S. A. Mehta, Z. V. Milanov, A. M. Velasco, L. M. Wodicka, H. K. Patel, P. P. Zarrinkar, and D. J. Lockhart. A small molecule-kinase interaction map for clinical kinase inhibitors. *Nat Biotech*, 2005, 23(3):329–336.
- [241] B. Zhao, M. J. Bower, P. J. McDevitt, H. Zhao, S. T. Davis, K. O. Johanson, S. M. Green, N. O. Concha, and B.-B. S. Zhou. Structural Basis for Chk1 Inhibition by UCN-01. *J. Biol. Chem.*, 2002, 277(48):46609–46615.
- [242] M. I. Davis, J. P. Hunt, S. Herrgard, P. Ciceri, L. M. Wodicka, G. Pallares, M. Hocker, D. K. Treiber, and P. P. Zarrinkar. Comprehensive analysis of kinase inhibitor selectivity. *Nat Biotech*, 2011, 29(11):1046–1051.
- [243] M. Bantscheff, D. Eberhard, Y. Abraham, S. Bastuck, M. Boesche, S. Hobson, T. Mathieson, J. Perrin, M. Raida, C. Rau, V. Reader, G. Sweetman, A. Bauer, T. Bouwmeester, C. Hopf, U. Kruse, G. Neubauer, N. Ramsden, J. Rick, B. Kuster, and G. Drewes. Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. *Nat Biotech*, 2007, 25(9):1035–1044.
- [244] N. P. Shah, C. Tran, F. Y. Lee, P. Chen, D. Norris, and C. L. Sawyers. Overriding Imatinib Resistance with a Novel ABL Kinase Inhibitor. *Science*, 2004, 305(5682):399–401.
- [245] J. S. Tokarski, J. A. Newitt, C. Ying, J. D. Cheng, M. Wittekind, S. E. Kiefer, K. Kish, F. Y. F. Lee, R. Borzilleri, L. J. Lombardo, D. Xie, Y. Zhang, and H. E. Klei. The Structure of Dasatinib (BMS-354825) Bound to Activated ABL Kinase Domain Elucidates Its Inhibitory Activity against Imatinib-Resistant ABL Mutants. *Cancer Research*, 2006, 66(11):5790–5797.
- [246] Y. Wang, J. Li, R. N. Booher, A. Kraker, T. Lawrence, W. R. Leopold, and Y. Sun. Radiosensitization of p53 Mutant Cells by PD0166285, a Novel G2 Checkpoint Abrogator. *Cancer Research*, 2001, 61(22):8211–8217.
- [247] J. PosthumaDeBoer, T. Würdinger, H. C. Graat, V. W. van Beusechem, M. N. Helder, B. J. van Royen, and G. J. Kaspers. WEE1 inhibition sensitizes osteosarcoma to radiotherapy. *BMC cancer*, 2011, 11:156–164.

- [248] R. L. Panek, G. H. Lu, S. R. Klutchko, B. L. Batley, T. K. Dahring, J. M. Hamby, H. Hallak, A. M. Doherty, and J. A. Keiser. In Vitro Pharmacological Characterization of PD 166285, a New Nanomolar Potent and Broadly Active Protein Tyrosine Kinase Inhibitor. *J. Pharm. Exp. Ther.*, 1997, 283(3):1433–1444.
- [249] N. Osherov and A. Levitzki. Epidermal-Growth-Factor-Dependent Activation of the Src-Family Kinases. *European Journal of Biochemistry*, 1994, 225(3):1047–1053.
- [250] Y. Han, C. G. Caday, A. Nanda, W. K. Cavenee, and H. J. S. Huang. Tyrphostin AG 1478 Preferentially Inhibits Human Glioma Cells Expressing Truncated Rather than Wild-Type Epidermal Growth Factor Receptors. *Cancer Research*, 1996, 56(17):3859–3861.
- [251] L. Shewchuk, A. Hassell, B. Wisely, W. Rocque, W. Holmes, J. Veal, and L. F. Kuyper. Binding Mode of the 4-Anilinoquinazoline Class of Protein Kinase Inhibitor: X-ray Crystallographic Studies of 4-Anilinoquinazolines Bound to Cyclin-Dependent Kinase 2 and p38 Kinase. *J. Med. Chem.*, 1999, 43(1):133–138.
- [252] M. Krug, K. Wichapong, G. Erlenkamp, W. Sippl, C. Schächtele, F. Totzke, and A. Hilgeroth. Discovery of 4-Benzylamino-Substituted α -Carbolines as a Novel Class of Receptor Tyrosine Kinase Inhibitors. *ChemMedChem*, 2011, 6(1):63–72.
- [253] K. Brachwitz, B. Voigt, L. Meijer, O. Lozach, C. Schächtele, J. Molnár, and A. Hilgeroth. Evaluation of the First Cytostatically Active 1-Aza-9-oxafluorenes as Novel Selective CDK1 Inhibitors with P-Glycoprotein Modulating Properties. *J. Med. Chem.*, 2003, 46(5):876–879.
- [254] B. Voigt, M. Krug, C. Schächtele, F. Totzke, and A. Hilgeroth. Probing Novel 1-Aza-9-oxafluorenes as Selective GSK-3 β Inhibitors. *ChemMedChem*, 2008, 3(1):120–126.
- [255] Z. Zhou, Y. Wang, and S. H. Bryant. Computational analysis of the cathepsin B inhibitors activities through LR-MMPBSA binding affinity calculation based on docked complex. *J. Comput. Chem.*, 2009, 30(14):2165–2175.
- [256] S. Wang, G. W. A. Milne, M. C. Nicklaus, V. E. Marquez, J. Lee, and P. M. Blumberg. Protein Kinase C. Modeling of the Binding Site and Prediction of Binding Constants. *J. Med. Chem.*, 1994, 37(9):1326–1338.
- [257] G. B. McGaughey, J. C. Culberson, B. P. Feuston, C. Kreatsoulas, V. Maiorov, and J. Shpungin. Scoring of KDR Kinase Inhibitors: Using Interaction Energy as a Guide for Ranking. *Molecular Diversity*, 2006, 10(3):341–347.
- [258] A. Kuznetsov and J. Järv. Allosteric cooperativity in inhibition of protein kinase α catalytic subunit. *Open Enzyme Inhib J*, 2008, 1(1):42–47.
- [259] T. Zhou, D. Huang, and A. Caflisch. Is Quantum Mechanics Necessary for Predicting Binding Free Energy? *J. Med. Chem.*, 2008, 51(14):4280–4288.
- [260] A. Del Rio, B. F. Baldi, and G. Rastelli. Activity Prediction and Structural Insights of Extracellular Signal-Regulated Kinase 2 Inhibitors with Molecular Dynamics Simulations. *Chemical Biology & Drug Design*, 2009, 74(6):630–635.
- [261] F. Uhlmann, C. Bouchoux, and S. López-Avilés. A quantitative model for cyclin-dependent kinase control of the cell cycle: revisited. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 2011, 366(1584):3572–3583.

- [262] R. M. Stephenson. Mutual solubility of water and pyridine derivatives. *J. Chem. Eng. Data*, 1993, 38(3):428–431.
- [263] R. M. Dickhut, A. W. Andren, and D. E. Armstrong. Naphthalene solubility in selected organic solvent/water mixtures. *J. Chem. Eng. Data*, 1989, 34(4):438–443.
- [264] A. Suenaga, N. Okimoto, Y. Hirano, and K. Fukui. An Efficient Computational Method for Calculating Ligand Binding Affinities. *PLoS ONE*, 2012, 7(8):1–9.

ANHANG

5.1. Screening

Hier sind alle errechneten Werte für die AUC, RIE und EF für jede verwendete Kristallstruktur und jede einzelne Scoring-Funktion aufgeführt. Aus diesen Tabellen wurden die Durchschnittswerte für die einzelnen Kinasen und Scoring-Funktionen im Ergebnisteil ermittelt.

5.1.1 AUC

Tabelle 5.1: Brasca AUC

Name	DFG-Motiv	normalisierter Score					ohne Wasser			GLIDEScore-Rescore			CONSENSUS-SCORE
		GOLD	GLIDE	PMF	PScore	ROCS	CHEMScore-Rescore	GOLD	PMF	PScore	GOLD	PMF	
2FVD	in	0,10	0,17	0,98	0,97		0,83	0,83	0,88	0,59	0,33	0,38	0,78
1P2A	in	0,99	0,17	0,99	0,99	0,99	0,82	0,85	0,85	0,49	0,33	0,31	0,77
1KE6	out	0,90	0,52	0,97	0,93	0,99	0,90	0,78	0,62	0,43	0,31	0,24	0,75
1DI8	out	0,10	0,11	0,98	0,97	0,99	0,89	0,90	0,77	0,46	0,32	0,22	0,782
mit Wasser													
2FVD	in	0,17	0,19	0,97	0,70		0,23	0,43	0,26	0,43	0,47	0,44	0,61
1KE6	out	0,76	0,67	0,96	0,88		0,23	0,74	0,64	0,51	0,46	0,41	0,82
1DI8	put	0,98	0,66	0,98	0,72		0,38	0,88	0,71	0,41	0,31	0,31	0,84

Tabelle 5.2: Peifer AUC

Name	DFG-Motiv	normalisierter Score					ohne Wasser			GLIDEScore-Rescore			CONSENSUS-SCORE
		GOLD	GLIDE	PMF	PScore	ROCS	CHEMScore-Rescore	GOLD	PMF	PScore	GOLD	PMF	
1EH4	in	0,88	0,87	0,75	0,76		0,81	0,71	0,60	0,67	0,30	0,38	0,33
2IZR	in	0,74	0,52	0,72	0,66		0,67	0,63	0,63	0,60	0,40	0,44	0,42
3GC9	in	0,66	0,55	0,47	0,66	0,50	0,50	0,57	0,57	0,48	0,30	0,30	0,58
3FLQ	out	0,73	0,52	0,74	0,74	0,50	0,61	0,61	0,70	0,52	0,46	0,49	0,63
2GTM	out	0,53	0,44	0,39	0,39	0,50	0,64	0,38	0,45	0,46	0,37	0,26	0,45
2BAQ	out	0,47	0,34	0,44	0,38	0,50	0,58	0,49	0,60	0,57	0,30	0,39	0,39
mit Wasser													
1EH4	in	0,73	0,70	0,68	0,70		0,55	0,46	0,47	0,51	0,47	0,37	0,71
2IZR	in	0,32	0,52	0,30	0,36		0,49	0,42	0,35	0,57	0,40	0,33	0,34
3GC9	in	0,74	0,49	0,57	0,57		0,58	0,67	0,60	0,54	0,43	0,42	0,59
2GTM	out	0,38	0,43	0,38	0,45		0,47	0,45	0,39	0,50	0,43	0,33	0,40

Tabelle 5.3: Matthews AUC

Name	DFG-Motiv	normalisierter Score				ROCS	ohne Wasser CHEMScore-Rescore			GLIDEScore-Rescore			CONSENSUS-SCORE
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2WMV	in	0,58	0,88	0,65	0,37	0,45	0,61	0,58	0,74	0,74	0,63	0,76	0,40
2XEZ	in	0,68	0,92	0,58	0,67		0,76	0,57	0,57	0,67	0,62	0,70	0,51

Tabelle 5.4: Huang AUC

Name	DFG-Motiv	normalisierter Score				ROCS	ohne Wasser CHEMScore-Rescore			GLIDEScore-Rescore			CONSENSUS-SCORE
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2GQG	in	0,83	0,76	0,67	0,82	0,81	0,61	0,58	0,50	0,64	0,42	0,45	0,72
3KF4	in	0,68	0,71	0,65	0,89		0,62	0,60	0,54	0,49	0,46	0,45	0,70
1IEP	out	0,81	0,75	0,69	0,82		0,71	0,60	0,59	0,48	0,44	0,42	0,76
2HIW	out	0,82	0,70	0,86	0,83		0,57	0,56	0,50	0,48	0,39	0,41	0,75
mit Wasser													
2GQG	in	0,83	0,57	0,67	0,85		0,68	0,94	0,96	0,51	0,42	0,45	0,61
3KF4	in	0,58	0,40	0,68	0,57		0,70	0,96	0,96	0,45	0,48	0,49	0,42
2HIW	out	0,55	0,43	0,83	0,72		0,63	0,96	0,96	0,51	0,41	0,41	0,69

Tabelle 5.5: Cheeseright AUC

Name	DFG-Motiv	normalisierter Score				ROCS	ohne Wasser CHEMScore-Rescore			GLIDEScore-Rescore			CONSENSUS-SCORE
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
3GC9	in	0,68	0,83	0,18	0,26	0,52	0,73	0,64	0,71	0,58	0,62	0,61	0,56
3FLQ	in	0,60	0,72	0,17	0,28		0,72	0,72	0,64	0,59	0,65	0,60	0,49
2BAQ	out	0,68	0,66	0,31	0,47		0,58	0,71	0,69	0,56	0,60	0,49	0,73
2GTM	out	0,56	0,83	0,31	0,27		0,59	0,71	0,62	0,54	0,59	0,61	0,41
mit Wasser													
3GC9	in	0,63	0,83	0,13	0,20		0,71	0,63	0,70	0,59	0,39	0,41	0,49
2GTM	out	0,62	0,83	0,53	0,59		0,73	0,85	0,83	0,57	0,33	0,40	0,55

Tabelle 5.6: Ermoli AUC

Name	DFG-Motiv	normalisierter Score				ROCS	ohne Wasser CHEMScore-Rescore			GLIDEScore-Rescore			CONSENSUS-SCORE
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
1UA2	in	0,49	0,96	0,62	0,49	0,89	0,58	0,61	0,74	0,55	0,68	0,79	0,60

Tabelle 5.7: Aliagas-Martin AUC

Name	DFG-Motiv	normalisierter Score				ROCS	ohne Wasser CHEMScore-Rescore			GLIDEScore-Rescore			CONSENSUS-SCORE
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2W1I	in	0,87	0,68	0,57	0,94	0,81	0,88	0,87	0,58	0,50	0,30	0,24	0,78
2F4J	in	0,83	0,62	0,50	0,88		0,83	0,85	0,66	0,52	0,31	0,28	0,75
1MUO	out	0,77	0,83	0,40	0,71		0,87	0,84	0,62	0,49	0,29	0,27	0,65
2J5O	out	0,60	0,63	0,44	0,83		0,57	0,84	0,61	0,17	0,29	0,29	0,641
mit Wasser													
2F4J	in	0,84	0,69	0,54	0,66		0,92	0,83	0,65	0,50	0,32	0,28	0,68

Tabelle 5.8: Tsou AUC

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE	
		normalisierter Score					ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
		GOLD	GLIDE	PMF	PScore	GOLD		PMF	PScore	GOLD	PMF	PScore		
2F2C	in	0,22	0,89	0,33	0,47	0,58	0,73	0,61	0,66	0,55	0,52	0,69	0,48	
1XO2	in	0,32	0,76	0,35	0,67		0,51	0,59	0,49	0,47	0,58	0,60	0,52	
1B18	out	0,19	0,55	0,38	0,46		0,42	0,59	0,52	0,47	0,55	0,54	0,75	
1G3N	out	0,20	0,44	0,36	0,46		0,66	0,52	0,54	0,44	0,50	0,50	0,37	

Tabelle 5.9: Qian AUC

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE	
		normalisierter Score					ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
		GOLD	GLIDE	PMF	PScore	GOLD		PMF	PScore	GOLD	PMF	PScore		
2J2I	in	0,87	0,61	0,90	0,75	0,77	0,92	0,79	0,84	0,83	0,34	0,41	0,71	
3BWF	in	0,92	0,69	0,90	0,75		0,85	0,85	0,89	0,77	0,29	0,40	0,75	
mit Wasser														
2J2I	in	0,91	0,63	0,91	0,88		0,93	0,82	0,94	0,68	0,41	0,78	0,71	
3BWF	in	0,89	0,51	0,91	0,87		0,91	0,89	0,90	0,83	0,65	0,48	0,70	

Tabelle 5.10: Gaisina AUC

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE	
		normalisierter Score					ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
		GOLD	GLIDE	PMF	PScore	GOLD		PMF	PScore	GOLD	PMF	PScore		
3F88	in	0,77	0,71	0,31	0,88	0,95	0,78	0,69	0,68	0,38	0,53	0,36	0,58	
3L1S	in	0,73	0,70	0,28	0,79		0,69	0,76	0,65	0,40	0,53	0,54	0,59	
mit Wasser														
3F88	in	0,84	0,80	0,21	0,80		0,88	0,68	0,67	0,51	0,47	0,47	0,55	
3L1S	in	0,86	0,80	0,23	0,94		0,86	0,67	0,67	0,49	0,44	0,46	0,60	

Tabelle 5.11: Surya AUC

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE	
		normalisierter Score					ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
		GOLD	GLIDE	PMF	PScore	GOLD		PMF	PScore	GOLD	PMF	PScore		
2GMX	in	0,56	0,65	0,23	0,40	0,62	0,33	0,37	0,28	0,48	0,49	0,56	0,49	
3G90	in	0,49	0,61	0,24	0,45		0,23	0,37	0,33	0,59	0,47	0,48	0,52	

Tabelle 5.12: Karcher AUC

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE	
		normalisierter Score					ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
		GOLD	GLIDE	PMF	PScore	GOLD		PMF	PScore	GOLD	PMF	PScore		
3GC9	in	0,82	0,83	0,60	0,88	0,78	0,97	0,93	0,95	0,39	0,71	0,55	0,87	
3FLQ	in	0,45	0,74	0,48	0,89		0,89	0,95	0,98	0,77	0,60	0,74	0,77	
2BAQ	out	0,71	0,94	0,83	0,91		0,95	0,97	0,96	0,65	0,70	0,83	0,86	
2GTM	out	0,65	0,97	0,80	0,96		0,91	0,96	0,95	0,66	0,83	0,72	0,85	
mit Wasser														
3GC9	in	0,54	0,90	0,53	0,79		0,87	0,94	0,95	0,49	0,70	0,58	0,62	
2GTM	out	0,85	0,82	0,74	0,93		0,89	0,92	0,95	0,61	0,49	0,39	0,80	

Tabelle 5.13: Schroeder AUC

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		normalisierter Score				ROCS	CHEMScore-Resscore			GLIDEScore-Resscore			
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2WGJ	in	0,74	0,24	0,52	0,87		0,20	0,52	0,31	0,41	0,33	0,29	0,70
2WD1	in	0,76	0,30	0,50	0,86	1,0	0,23	0,61	0,22	0,42	0,31	0,40	0,72
3F82	out	0,86	0,95	0,37	0,89		0,42	0,67	0,45	0,28	0,18	0,30	0,78
3CTH	out	0,85	0,89	0,36	0,90		0,38	0,51	0,48	0,34	0,47	0,35	0,78
mit Wasser													
2WGJ	in	0,87	0,37	0,72	0,99		0,29	0,67	0,63	0,37	0,34	0,55	0,73

Tabelle 5.14: Coumar AUC

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		normalisierter Score				ROCS	CHEMScore-Resscore			GLIDEScore-Resscore			
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2W11		0,51	0,55	0,97	0,70	0,42	0,59	0,59	0,60	0,41	0,52	0,35	0,63
2F4J	in	0,48	0,50	0,95	0,74		0,62	0,58	0,55	0,62	0,54	0,35	0,62
1MUO	out	0,59	0,55	0,91	0,68		0,66	0,49	0,49	0,61	0,51	0,39	0,63
2J5O	out	0,62	0,59	0,92	0,68		0,67	0,45	0,52	0,63	0,60	0,46	0,64
mit Wasser													
2F4J	in	0,44	0,65	0,86	0,60		0,58	0,60	0,47	0,61	0,37	0,43	0,646

Tabelle 5.15: Popowycz AUC

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		normalisierter Score				ROCS	CHEMScore-Resscore			GLIDEScore-Resscore			
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
CDK2													
1E9H	in	0,68	0,48	0,99	0,86		0,50	0,41	0,56	0,58	0,24	0,46	0,82
1FQ1	in	0,44	0,38	0,98	0,38		0,29	0,46	0,34	0,56	0,41	0,48	0,62
CDK4													
1XO2	in	0,91	0,97	0,91	0,64	0,75	0,72	0,46	0,23	0,76	0,58	0,39	0,83
1JOW	in	0,58	0,62	0,99	0,59		0,71	0,41	0,36	0,80	0,24	0,45	0,77
1B17	out	0,63	0,80	1,00	0,59		0,58	0,47	0,41	0,72	0,42	0,46	0,82
1G3N	out	0,79	0,78	0,99	0,63		0,63	0,63	0,55	0,70	0,42	0,36	0,77
CDK5													
1UNL	in	0,83	0,69	0,98	0,87		0,77	0,45	0,58	0,82	0,53	0,62	0,80
1UNG	in	0,76	0,75	0,99	0,75		0,39	0,57	0,55	0,80	0,48	0,79	0,81
CDK7													
1UA2	in	0,73	0,90	1,00	0,57		0,28	0,37	0,42	0,47	0,55	0,47	0,82
CDK9													
3BLR	in	0,93	0,89	0,99	0,78		0,74	0,52	0,67	0,74	0,67	0,87	0,84
3MIA	in	0,67	0,15	0,97	0,43		0,57	0,39	0,31	0,76	0,53	0,68	0,76
mit Wasser													
1UNL	in	0,87	0,71	0,99	0,79		0,56	0,52	0,74	0,36	0,34	0,88	0,76

Tabelle 5.16: Menichincheri AUC

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		normalisierter Score				ROCS	CHEMScore-Resscore			GLIDEScore-Resscore			
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
3DU8	in	0,41	0,67	0,36	0,18	1,0	0,78	0,73	0,82	0,42	0,63	0,44	0,39
mit Wasser													
3DU8	in	0,47	0,70	0,31	0,13		0,79	0,70	0,78	0,42	0,49	0,54	0,40

Tabelle 5.17: Renhowe AUC

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		normalisierter Score				ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
3B8R	in	0,41	0,87	0,82	0,73	0,58	0,73	0,91	0,91	0,67	0,44	0,62	0,69
2P2H	in	0,34	0,80	0,66	0,53		0,79	0,89	0,86	0,60	0,40	0,55	0,59
3EWH	out	0,29	0,69	0,79	0,61		0,55	0,88	0,84	0,60	0,71	0,74	0,59
3CPC	out	0,48	0,70	0,66	0,77		0,76	0,83	0,82	0,70	0,36	0,49	0,62
mit Wasser													
2P2H	in	0,32	0,69	0,64	0,53		0,61	0,84	0,74	0,55	0,33	0,54	0,55
3EWH	out	0,38	0,68	0,76	0,68		0,64	0,85	0,92	0,55	0,28	0,58	0,61

Tabelle 5.18: Howard AUC

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		normalisierter Score				ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2W11	in	0,76	0,76	0,78	0,64	1,0	0,70	0,61	0,60	0,72	0,51	0,37	0,78
2F4J	in	0,74	0,78	0,78	0,68		0,39	0,70	0,61	0,65	0,52	0,38	0,79
1MUO	out	0,83	0,71	0,79	0,69		0,40	0,62	0,55	0,80	0,64	0,60	0,80
2J5O	out	0,89	0,74	0,76	0,72		0,83	0,65	0,60	0,71	0,23	0,53	0,80
mit Wasser													
2F4J	in	0,74	0,76	0,69	0,30		0,44	0,61	0,66	0,70	0,49	0,49	0,63

Tabelle 5.19: Xia AUC

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		normalisierter Score				ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2J2I	in	0,29	0,84	0,17	0,14	0,71	0,56	0,05	0,05	0,45	0,47	0,39	0,11
3BWF	in	0,30	0,88	0,19	0,19		0,58	0,07	0,07	0,50	0,46	0,45	0,22
mit Wasser													
2J2I	in	0,17	0,84	0,16	0,17		0,59	0,74	0,62	0,38	0,45	0,42	0,18
3BWF	in	0,30	0,83	0,15	0,16		0,62	0,78	0,63	0,48	0,42	0,41	0,19

5.1.2 Enrichmentfaktor 5

Tabelle 5.20: Brasca EF5

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		normalisierter Score				ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2FVD	in	10,00	2,07	10,00	10,00	10,00	10,00	6,22	5,53	4,15	3,46	4,15	8,30
1P2A	in	10,00	2,07	10,00	10,00		5,53	6,91	4,15	1,38	3,46	3,46	10,00
1KE6	out	6,22	5,53	10,00	10,00		0,69	3,46	2,05	1,38	2,07	2,77	8,30
1DI8	out	10,00	1,38	10,00	10,00		10,00	7,61	6,22	3,46	2,76	2,05	10,00
mit Wasser													
2FVD	in	0,00	1,38	10,00	3,46		2,05	0,69	3,46	0,00	4,15	2,77	6,22
1KE6	out	1,38	6,22	10,00	10,00		2,05	4,15	2,77	2,07	4,15	4,15	6,22
1DI8	put	10,00	6,91	10,00	8,30		3,46	8,30	4,15	0,69	3,46	5,53	8,30

Tabelle 5.21: Peifer EF5

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		normalisierter Score				ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
1EH4	in	10,00	8,50	6,96	7,73		4,64	4,64	3,09	3,86	1,55	2,32	4,64
2IZR	in	7,73	5,44	6,96	10,00		7,00	10,00	10,00	4,64	3,86	3,86	4,64
3GC9	in	0,77	5,44	3,09	7,73	6,18	3,09	2,32	2,32	0,00	1,55	2,32	5,41
3FLQ	out	3,09	6,18	6,18	10,00		1,55	1,55	3,86	1,55	3,86	4,64	5,41
2GTM	out	0,00	4,64	0,00	1,55		1,55	0,00	0,00	0,00	2,32	2,32	5,41
2BAQ	out	0,77	3,86	0,00	1,55		0,77	1,55	1,55	0,00	2,32	2,32	4,64
mit Wasser													
1EH4	in	2,32	2,32	5,41	6,18		3,86	0,77	1,55	0,00	2,32	2,32	7,00
2IZR	in	0,00	0,00	1,55	0,77		3,86	0,00	1,55	3,86	4,64	3,86	4,64
3GC9	in	5,41	3,86	1,55	0,00		1,55	3,86	1,55	2,32	2,32	2,32	4,64
2GTM	out	0,00	3,86	5,41	0,00		3,09	1,55	0,77	0,00	2,32	2,32	4,64

Tabelle 5.22: Matthews EF5

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		normalisierter Score				ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2WMV	in	2,50	7,50	6,25	2,50	3,75	3,75	2,50	5,00	6,65	6,25	7,50	5,00
2XEZ	in	3,75	8,75	6,25	1,25	3,75	5,00	2,50	5,00	3,75	6,25	7,50	5,00

Tabelle 5.23: Huang EF5

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		normalisierter Score				ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2GQG	in	4,63	2,70	2,31	6,17		0,00	2,31	0,00	1,16	0,77	2,31	4,63
3KF4	in	0,39	1,54	0,77	5,78	5,78	0,00	0,39	1,54	0,00	3,47	1,54	3,48
1IEP	out	5,78	3,47	3,08	5,40		2,70	2,31	0,77	3,47	1,54	1,54	5,39
2HIW	out	3,47	3,08	6,94	6,17		1,54	1,54	1,54	0,77	0,77	2,31	5,39
mit Wasser													
2GQG	in	3,47	0,39	1,93	6,17		0,00	10,00	10,00	0,39	0,39	0,77	3,48
3KF4	in	0,00	0,39	1,16	2,70		1,54	10,00	10,00	0,77	0,39	1,16	2,70
2HIW	out	0,39	0,00	4,63	1,16		0,77	10,00	10,00	0,39	0,77	0,39	2,70

Tabelle 5.24: Cheeseright EF5

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		normalisierter Score				ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
3GC9	in	3,28	2,30	0,00	0,00		2,30	1,97	0,99	3,611	1,97	1,64	1,31
3FLQ	in	1,31	2,30	0,00	0,38	4,56	3,28	3,61	2,63	2,626	0,99	1,97	1,31
2BAQ	out	0,99	1,64	0,66	0,38		1,64	3,28	3,61	1,641	0,38	0,38	1,64
2GTM	out	0,99	2,30	0,66	0,00		1,97	2,30	1,97	2,626	0,99	0,66	1,64
mit Wasser													
3GC9	in	1,97	2,30	0,00	0,00		2,30	1,31	2,63	1,64	0,99	1,64	1,31
2GTM	out	1,64	2,63	1,36	3,28		3,28	7,80	6,89	1,64	1,99	0,98	1,64

Tabelle 5.25: Ermoli EF5

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		GOLD	normalisierter Score			ROCS	CHEMScore-Resscore			GLIDEScore-Resscore			
			GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
1UA2	in	0,912	0,912	2,735	1,367	3,190	0,912	2,735	3,190	0,912	2,735	1,367	0,912

Tabelle 5.26: Aliagas-Martin EF5

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		GOLD	normalisierter Score			ROCS	CHEMScore-Resscore			GLIDEScore-Resscore			
			GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2W11	in	10,00	6,67	6,67	3,33	6,67	6,67	6,67	6,67	6,67	3,333	6,67	6,67
2F4J	in	3,33	3,33	3,33	3,33		6,67	6,67	6,67	3,33	3,333	10,00	3,33
1MUO	out	3,33	6,67	6,67	6,67		6,67	6,67	3,33	0,00	3,333	6,67	6,67
2J5O	out	0,00	3,33	0,00	6,67		0,00	10,00	0,00	0,00	6,667	3,33	3,33
mit Wasser													
2F4J	in	3,33	6,67	3,33	3,33		10,00	10,00	6,67	0,00	3,33	3,33	3,33

Tabelle 5.27: Tsou EF5

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		GOLD	normalisierter Score			ROCS	CHEMScore-Resscore			GLIDEScore-Resscore			
			GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2F2C	in	0,00	0,00	1,30	1,56	1,82	4,30	1,82	1,56	2,07	1,56	1,04	1,56
1XO2	in	0,78	0,26	0,52	1,30		0,78	1,82	0,52	2,07	1,30	1,30	1,30
1BI8	out	0,00	0,52	0,53	0,00		0,26	1,82	1,56	2,33	1,04	0,52	2,33
1G3N	out	0,26	0,26	0,78	0,26		1,30	1,30	0,52	1,30	1,04	1,30	2,07

Tabelle 5.28: Qian EF5

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		GOLD	normalisierter Score			ROCS	CHEMScore-Resscore			GLIDEScore-Resscore			
			GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2J2I	in	5,00	2,62	6,43	3,52	7,14	10,00	7,14	7,14	7,86	2,62	5,00	7,86
3BWF	in	2,62	6,43	6,43	2,62		2,62	7,86	6,43	2,62	2,62	7,86	7,86
mit Wasser													
2J2I	in	9,29	7,14	6,43	7,14		10,00	7,86	9,29	4,29	7,14	7,86	9,29
3BWF	in	7,14	9,29	5,71	5,71		7,14	10,00	10,00	2,62	10,00	4,29	9,29

Tabelle 5.29: Gaisina EF5

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		GOLD	normalisierter Score			ROCS	CHEMScore-Resscore			GLIDEScore-Resscore			
			GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
3F88	in	2,50	3,50	0,50	5,50	4,00	5,00	2,50	2,00	1,00	1,50	1,00	2,50
3L1S	in	3,00	3,00	0,00	3,00		2,50	2,50	1,50	1,50	1,00	0,50	2,50
mit Wasser													
3F88	in	5,50	4,00	0,00	4,00		4,00	3,00	3,00	1,50	2,00	1,50	3,50
3L1S	in	6,50	3,50	0,00	9,50		3,00	3,50	3,00	1,50	2,00	2,50	4,50

Tabelle 5.30: Surya EF5

Name	DFG-Motiv	normalisierter Score						ohne Wasser			GLIDEScore-Rescore			CONSENSUS-SCORE
		GOLD	GLIDE	PMF	PScore	ROCS	CHEMScore-Rescore			GOLD	PMF	PScore		
							GOLD	PMF	PScore					
2GMX	in	2,00	2,50	0,00	0,50	5,00	0,50	1,50	1,00	0,50	2,00	2,50	2,00	
3G90	in	2,00	2,50	0,00	1,50		0,00	0,50	0,00	3,00	1,50	1,50	2,00	

Tabelle 5.31: Karcher EF5

Name	DFG-Motiv	normalisierter Score						ohne Wasser			GLIDEScore-Rescore			CONSENSUS-SCORE
		GOLD	GLIDE	PMF	PScore	ROCS	CHEMScore-Rescore			GOLD	PMF	PScore		
							GOLD	PMF	PScore					
3GC9	in	3,00	4,00	1,00	9,00		3,00	8,00	10,00	2,00	2,00	10,00	7,00	
3FLQ	in	0,00	5,00	0,00	4,00	5,00	9,00	8,00	10,00	4,00	6,00	10,00	5,00	
2BAQ	out	0,00	5,00	3,00	7,00		10,00	10,00	10,00	3,00	4,00	10,00	5,00	
2GTM	out	0,00	7,00	7,00	9,00		10,00	10,00	10,00	7,00	4,00	10,00	5,00	
mit Wasser														
3GC9	in	0,00	4,00	0,00	2,00		9,00	8,00	9,00	3,00	2,00	9,00	5,00	
2GTM	out	3,00	7,00	4,00	10,00		8,00	10,00	10,00	3,00	4,00	9,00	6,00	

Tabelle 5.32: Schroeder EF5

Name	DFG-Motiv	normalisierter Score						ohne Wasser			GLIDEScore-Rescore			CONSENSUS-SCORE
		GOLD	GLIDE	PMF	PScore	ROCS	CHEMScore-Rescore			GOLD	PMF	PScore		
							GOLD	PMF	PScore					
2WGJ	in	2,00	0,00	2,00	4,00		0,00	0,00	0,00	2,00	2,00	0,00	4,00	
2WD1	in	2,00	0,00	0,00	8,00	10,00	0,00	0,00	0,00	2,00	0,00	2,00	8,00	
3F82	out	0,00	0,00	0,00	6,00		0,00	2,00	0,00	0,00	2,00	2,00	4,00	
3CTH	out	4,00	0,00	0,00	8,00		0,00	0,00	0,00	2,00	0,00	2,00	2,00	
mit Wasser														
2WGJ	in	4,00	2,00	3,00	10,00		0,00	4,00	4,00	2,00	2,00	4,00	6,00	

Tabelle 5.33: Coumar EF5

Name	DFG-Motiv	normalisierter Score						ohne Wasser			GLIDEScore-Rescore			CONSENSUS-SCORE
		GOLD	GLIDE	PMF	PScore	ROCS	CHEMScore-Rescore			GOLD	PMF	PScore		
							GOLD	PMF	PScore					
2W11	in	0,00	1,88	9,38	8,69	2,50	0,00	2,500	1,25	1,25	1,88	9,38	6,88	
2F4J	in	0,00	7,50	10,00	6,88		2,50	0,625	0,63	2,50	1,88	6,88	6,88	
1MUO	out	0,63	1,88	7,50	1,88		1,25	1,25	0,625	0,00	1,25	8,69	3,75	
2J5O	out	0,00	6,88	9,38	3,75		2,50	0,63	1,250	2,50	6,88	3,75	3,75	
mit Wasser														
2F4J	in	0,00	7,50	7,50	2,50		0,63	0,00	0,00	0,63	2,50	2,50	3,75	

Tabelle 5.34: Popowycz EF5

Name	DFG-Motiv	normalisierter Score				ROCS	ohne Wasser CHEMScore-Rescore			GLIDEScore-Rescore			CONSENSUS-SCORE
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
CDK2													
1E9H	in	0,00	3,33	10,00	6,67		0,00	0,00	0,00	3,33	3,33	6,67	6,67
1FQ1	in	0,00	3,33	10,00	0,00		0,00	0,00	0,00	3,33	6,67	6,67	6,67
CDK4													
1XO2	in	6,67	3,33	3,33	6,67	3,33	3,33	0,00	0,00	6,67	6,67	6,67	3,33
1JOW	in	0,00	6,67	10,00	0,00		3,33	0,00	0,00	3,33	0,00	0,00	3,33
1B17	out	0,00	6,67	10,00	3,33		0,00	3,33	0,00	3,33	3,33	3,33	3,33
1G3N	out	3,33	0,00	10,00	3,33		0,00	3,33	0,00	6,67	3,33	3,33	6,67
CDK5													
1UNL	in	6,67	6,67	6,67	6,67		6,67	0,00	3,33	3,33	6,67	6,67	10,00
1UNG	in	3,33	3,33	10,00	3,33		3,33	0,00	0,00	6,67	3,33	3,33	6,67
CDK7													
1UA2	in	0,00	0,00	10,00	0,00		0,00	0,00	3,33	6,67	3,33	10,00	3,33
CDK9													
3BLR	in	10,00	6,67	10,00	6,67		3,33	0,00	0,00	6,67	3,33	3,33	10,00
3MIA	in	0,00	3,33	6,67	0,00		0,00	0,00	0,00	6,67	10,00	3,33	3,33
mit Wasser													
1UNL	in	10,00	6,67	10,00	6,67		6,67	0,00	3,33	3,33	3,33	6,67	6,67

Tabelle 5.35: Menichincheri EF5

Name	DFG-Motiv	normalisierter Score				ROCS	ohne Wasser CHEMScore-Rescore			GLIDEScore-Rescore			CONSENSUS-SCORE
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
3DU8	in	0,83	0,83	0,83	0,83	10,00	5,47	4,17	6,25	5,00	2,09	1,25	0,83
mit Wasser													
3DU8	in	0,00	0,00	1,25	0,42		5,47	2,04	4,17	5,42	2,08	6,25	1,25

Tabelle 5.36: Renhowe EF5

Name	DFG-Motiv	normalisierter Score				ROCS	ohne Wasser CHEMScore-Rescore			GLIDEScore-Rescore			CONSENSUS-SCORE
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
3B8R	in	0,00	3,56	7,57	3,12		2,23	6,23	5,34	1,34	0,89	5,34	5,79
2P2H	in	0,00	4,01	2,23	1,78		3,12	4,90	2,67	3,12	2,23	4,01	3,56
3EWH	out	0,00	4,45	5,79	1,34	5,34	0,00	4,90	1,78	1,78	1,78	1,78	4,01
3CPC	out	0,45	1,78	3,56	3,12		2,67	6,23	0,00	4,01	0,00	1,34	3,56
mit Wasser													
2P2H	in	0,00	3,12	3,56	0,45		1,78	6,23	4,45	1,78	5,79	4,01	4,01
3EWH	out	0,00	3,12	7,12	4,01		0,89	5,79	4,45	1,78	6,23	4,90	4,90

Tabelle 5.37: Howard EF5

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		normalisierter Score				ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2W11	in	2,86	8,57	6,43	6,43		1,43	0,00	1,43	2,86	6,43	7,14	6,43
2F4J	in	1,43	7,14	7,14	7,14	7,14	1,43	0,00	0,00	2,86	8,57	5,71	7,14
1MUO	out	5,71	8,57	8,57	7,14		0,00	0,00	0,00	1,43	2,88	5,71	9,26
2J5O	out	10,00	8,57	8,57	8,57		8,57	0,00	2,86	8,75	2,88	8,57	9,26
mit Wasser													
2F4J	in	1,43	7,14	2,88	0,00		2,86	2,86	1,43	0,00	7,14	5,71	4,28

Tabelle 5.38: Xia EF5

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		normalisierter Score				ROCS	CHEMScore-Resscore			GLIDEScore-Resscore			
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2J2I	in	1,10	0,38	0,38	0,76	1,52	0,00	0,00	0,00	0,38	1,52	0,76	1,83
3BWF	in	0,76	0,76	0,76	1,14		0,00	0,38	0,38	0,38	1,52	0,76	1,52
mit Wasser													
2J2I	in	1,16	0,38	0,76	0,76		1,83	1,53	1,83	0,76	1,14	1,83	1,14
3BWF	in	0,00	0,00	0,38	0,38		1,83	2,66	1,14	0,76	1,14	1,14	1,52

5.1.3 RIE

Tabelle 5.39: Brasca RIE

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		normalisierter Score				ROCS	CHEMScore-Resscore			GLIDEScore-Resscore			
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2FVD	in	8,29	5,23	0,00	0,00	6,59	8,05	5,01	5,05	0,21	4,00	4,31	6,75
1P2A	in	8,08	4,97	0,00	0,00		5,11	5,53	4,08	0,49	3,21	5,55	6,69
1KE6	out	4,82	2,71	0,00	0,00		1,83	3,42	1,97	0,57	4,88	1,75	6,02
1DI8	out	8,33	6,45	0,00	0,00		7,41	6,01	4,91	1,01	5,12	3,49	6,76
mit Wasser													
2FVD	in	0,29	1,11	0,00	0,15		0,55	0,77	0,04	0,67	0,75	0,82	3,30
1KE6	out	2,06	2,95	0,00	0,00		2,00	3,75	2,21	0,41	1,00	0,87	3,36
1DI8	put	7,75	1,50	0,00	0,27		3,48	6,29	3,48	0,77	0,83	1,01	2,00

Tabelle 5.40: Peifer RIE

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		normalisierter Score				ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
1EH4	in	6,72	4,60	0,98	2,79		3,92	3,92	3,03	0,56	0,46	1,57	3,30
2IZR	in	5,50	4,12	1,37	3,12		1,87	6,51	6,51	1,24	0,84	1,87	2,78
3GC9	in	1,76	3,45	1,02	2,64	4,39	1,94	2,17	2,23	0,93	0,43	2,00	3,33
3FLQ	out	2,29	2,93	0,84	2,64		1,98	1,54	3,40	0,71	0,55	1,24	5,34
2GTM	out	0,38	3,81	0,88	2,85		1,77	0,03	0,26	1,83	0,67	1,50	2,95
2BAQ	out	1,17	1,15	0,44	2,11		0,99	1,77	1,38	0,37	0,32	2,46	3,00
mit Wasser													
1EH4	in	2,33	1,30	1,05	2,97		1,00	0,90	0,86	1,35	0,40	0,76	2,30
2IZR	in	0,05	1,56	2,80	2,64		1,21	0,30	0,05	0,92	0,46	0,80	2,34
3GC9	in	4,46	1,64	1,54	2,33		1,56	3,39	1,77	1,53	0,86	0,95	3,38
2GTM	out	0,09	0,77	1,92	2,96		0,48	1,79	0,14	0,88	0,64	0,85	1,98

Tabelle 5.41: Matthews RIE

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		normalisierter Score				ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2WMV	in	2,52	2,87	2,04	5,82	4,83	2,84	2,12	4,04	0,38	1,77	1,56	2,32
2XEZ	in	3,48	3,00	3,02	6,76		4,49	2,14	3,52	0,57	1,89	1,67	1,07

Tabelle 5.42: Huang RIE

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		normalisierter Score				ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2GQG	in	3,83	3,50	0,01	0,05		0,87	1,88	0,75	1,07	1,30	1,49	1,25
3KF4	in	0,83	2,96	0,01	0,00	3,80	0,41	0,94	1,73	0,64	0,90	1,13	0,04
1IEP	out	5,03	3,12	0,01	0,00		2,41	1,47	1,73	0,80	0,83	1,02	2,28
2HIW	out	3,02	1,89	0,01	0,00		1,50	2,11	1,56	1,43	1,50	0,99	5,01
mit Wasser													
2GQG	in	3,42	3,51	0,02	0,00		0,94	7,83	8,00	1,37	1,29	1,29	1,24
3KF4	in	0,26	2,23	0,03	0,15		0,37	7,98	7,99	0,75	0,94	0,88	0,31
2HIW	out	0,42	1,82	0,00	0,02		1,45	7,99	8,01	1,54	1,44	1,00	3,45

Tabelle 5.43: Cheeseright RIE

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		normalisierter Score				ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
		GOLD	GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
3GC9	in	3,06	2,93	4,39	2,87		2,26	1,75	1,41	1,47	2,39	1,87	1,51
3FLQ	in	1,07	1,97	4,60	2,54	2,99	2,94	3,26	2,65	1,05	1,05	1,00	0,57
2BAQ	out	1,25	2,00	4,01	0,98		1,36	2,81	1,98	2,44	2,20	1,01	3,77
2GTM	out	1,22	1,11	2,08	3,08		1,50	2,30	1,98	1,49	0,94	0,85	0,64
mit Wasser													
3GC9	in	1,93	1,40	5,43	3,36		1,83	1,53	2,35	1,76	2,39	2,00	0,84
2GTM	out	1,62	1,23	1,73	2,10		1,68	5,93	5,60	1,60	1,39	1,01	3,10

Tabelle 5.44: Ermoli RIE

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		GOLD	normalisierter Score			ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
			GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
1UA2	in	0,66	1,03	0,09	1,84	1,48	1,08	0,39	2,46	0,83	0,17	2,01	3,17

Tabelle 5.45: Aliagas-Martin RIE

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		GOLD	normalisierter Score			ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
			GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2W11	in	7,04	5,44	6,15	3,07	2,88	5,96	7,03	4,06	2,91	2,79	1,87	2,86
2F4J	in	3,77	3,98	5,99	1,70		5,86	5,95	2,52	1,72	1,50	1,48	3,34
1MUO	out	3,74	2,96	5,91	0,20		6,35	6,15	3,11	1,32	1,50	2,05	3,22
2J5O	out	0,13	1,04	1,41	1,00		0,12	1,04	1,46	3,39	0,32	0,97	2,05
mit Wasser													
2F4J	in	2,44	2,67	3,85	0,93		7,21	6,20	4,84	3,88	4,10	3,99	3,27

Tabelle 5.46: Tsou RIE

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		GOLD	normalisierter Score			ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
			GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2F2C	in	0,08	0,13	1,05	1,07	1,45	3,83	1,43	2,05	1,04	1,09	1,04	2,15
1XO2	in	0,75	2,98	1,46	0,02		0,92	1,67	0,70	0,65	1,10	0,75	3,12
1BI8	out	0,00	0,36	1,05	0,52		0,34	2,00	1,39	1,09	1,91	1,31	2,77
1G3N	out	0,16	0,56	1,06	0,67		1,35	1,29	0,50	1,17	0,93	1,19	0,31

Tabelle 5.47: Qian RIE

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		GOLD	normalisierter Score			ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
			GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2J2I	in	4,58	0,61	0,46	0,76	3,95	7,60	5,45	5,81	0,33	0,34	0,41	4,14
3BWF	in	1,31	0,69	0,62	0,78		1,31	6,00	5,40	0,29	0,26	0,40	1,31
mit Wasser													
2J2I	in	7,44	0,63	0,40	0,76		7,54	6,39	7,76	1,07	1,41	1,78	4,83
3BWF	in	5,09	0,50	0,66	0,87		5,11	7,00	7,68	1,32	1,65	1,48	4,26

Tabelle 5.48: Gaisina RIE

Name	DFG-Motiv	ohne Wasser											CONSENSUS-SCORE
		GOLD	normalisierter Score			ROCS	CHEMScore-Rescore			GLIDEScore-Rescore			
			GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
3F88	in	2,41	2,52	1,26	0,46	3,08	4,34	2,17	1,81	2,24	0,53	0,36	2,84
3L1S	in	2,64	3,76	1,43	0,45		1,69	2,53	1,49	1,84	0,53	0,54	3,00
mit Wasser													
3F88	in	4,19	3,20	2,34	0,31		4,52	2,78	2,79	0,06	0,70	0,23	1,69
3L1S	in	4,87	2,74	2,56	0,26		5,87	4,86	2,74	0,77	0,84	0,23	3,30

Tabelle 5.49: Surya RIE

Name	DFG-Motiv	GOLD	normalisierter Score			ROCS	ohne Wasser CHEMScore-Rescore			GLIDEScore-Rescore			CONSENSUS-SCORE
			GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2GMX	in	1,67	0,46	4,68	2,39	1,69	0,49	1,20	0,94	0,83	1,70	0,78	0,08
3G90	in	1,13	0,52	4,96	1,96		0,21	0,70	0,06	0,72	2,81	0,94	0,43

Tabelle 5.50: Karcher RIE

Name	DFG-Motiv	GOLD	normalisierter Score			ROCS	ohne Wasser CHEMScore-Rescore			GLIDEScore-Rescore			CONSENSUS-SCORE
			GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
3GC9	in	1,43	1,72	0,18	0,73		1,44	6,64	6,90	1,38	0,36	0,82	1,44
3FLQ	in	0,20	1,34	1,71	0,28	1,60	5,70	5,69	8,43	0,05	1,97	0,64	2,25
2BAQ	out	0,81	3,39	0,35	0,01		7,88	8,15	8,31	1,76	0,44	0,12	5,45
2GTM	out	0,77	1,84	0,00	0,00		7,17	8,02	7,79	1,34	0,22	0,14	4,70
mit Wasser													
3GC9	in	0,07	1,36	0,83	0,73		6,32	6,47	6,88	1,80	0,46	0,99	1,04
2GTM	out	1,92	2,00	0,23	1,00		5,78	7,66	7,69	1,32	0,30	0,48	2,11

Tabelle 5.51: Schroeder RIE

Name	DFG-Motiv	GOLD	normalisierter Score			ROCS	ohne Wasser CHEMScore-Rescore			GLIDEScore-Rescore			CONSENSUS-SCORE
			GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2WGJ	in	1,67	3,92	0,62	0,00		0,00	0,11	0,01	2,15	0,94	0,01	3,32
2WD1	in	2,23	2,77	0,29	0,00	4,11	0,00	0,28	0,00	2,14	0,43	0,11	0,01
3F82	out	1,53	2,84	0,33	0,00		0,02	1,75	0,06	1,19	0,46	0,27	5,11
3CTH	out	3,87	3,95	0,58	0,00		0,00	0,11	0,03	1,93	1,04	0,23	5,04
mit Wasser													
2WGJ	in	3,72	4,87	0,00	0,00		0,00	1,61	0,19	2,69	1,11	1,05	3,13

Tabelle 5.52: Coumar RIE

Name	DFG-Motiv	GOLD	normalisierter Score			ROCS	ohne Wasser CHEMScore-Rescore			GLIDEScore-Rescore			CONSENSUS-SCORE
			GLIDE	PMF	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2W11	in	0,23	0,17	0,00	0,79	2,87	1,10	0,65	0,47	0,13	0,38	0,77	0,00
2F4J	in	0,26	0,31	0,00	0,79		2,39	0,71	0,63	0,83	0,45	0,61	0,63
1MUO	out	0,85	0,61	0,00	0,98		1,62	1,17	0,35	0,25	0,57	0,68	0,06
2J5O	out	0,37	0,38	0,00	2,19		1,90	0,57	0,79	0,62	0,33	0,90	0,08
mit Wasser													
2F4J	in	0,15	0,27	0,00	0,89		0,67	0,69	0,48	0,72	0,38	0,55	1,02

Tabelle 5.53: Popowycz RIE

Name	DFG-Motiv	normalisierter Score					ROCS	ohne Wasser CHEMScore-Rescore			GLIDEScore-Rescore			CONSENSUS-SCORE
		GOLD	GLIDE	PMF	PScore	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
CDK2														
1E9H	in	0,10	5,74	0,00	0,00			0,05	0,28	0,05	0,28	0,00	0,03	6,85
1FQ1	in	0,01	3,09	0,00	1,29			0,00	0,13	0,00	0,14	0,01	1,00	4,63
CDK4														
1XO2	in	4,83	2,09	0,00	0,01	2,48		2,25	0,20	0,00	3,08	0,03	0,18	6,72
1JOW	in	0,55	1,10	0,00	0,01			1,58	0,05	0,00	0,67	0,05	0,58	4,54
1B17	out	0,10	0,18	0,00	0,09			0,30	1,48	0,43	0,33	0,11	0,44	6,91
1G3N	out	2,47	1,00	0,00	1,04			0,38	1,67	0,39	2,77	0,02	2,02	5,97
CDK5														
1UNL	in	3,17	2,19	0,00	0,00			4,36	0,06	1,09	4,11	0,21	0,18	5,90
1UNG	in	1,41	1,89	0,00	0,01			0,03	0,05	0,55	2,21	0,46	0,33	4,83
CDK7														
1UA2	in	0,94	4,30	0,00	0,75			0,00	0,10	1,46	2,73	0,38	1,86	5,70
CDK9														
3BLR	in	6,70	6,15	0,00	0,00			1,86	0,05	0,97	3,50	0,46	1,00	7,30
3MIA	in	0,29	0,84	0,00	0,75			0,41	0,01	0,01	2,51	0,38	2,14	0,90
mit Wasser														
1UNL	in	6,18	4,99	0,00	0,54			2,97	0,11	2,76	3,97	0,57	1,11	5,83

Tabelle 5.54: Menichincheri RIE

Name	DFG-Motiv	normalisierter Score					ROCS	ohne Wasser CHEMScore-Rescore			GLIDEScore-Rescore			CONSENSUS-SCORE
		GOLD	GLIDE	PMF	PScore	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
3DU8	in	0,62	4,87	2,45	6,09	3,36		4,47	3,91	5,71	1,20	4,87	6,43	5,46
mit Wasser														
3DU8	in	0,50	2,99	3,28	6,92			3,98	1,98	4,10	2,20	2,64	3,91	3,01

Tabelle 5.55: Renhowe RIE

Name	DFG-Motiv	normalisierter Score					ROCS	ohne Wasser CHEMScore-Rescore			GLIDEScore-Rescore			CONSENSUS-SCORE
		GOLD	GLIDE	PMF	PScore	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
3B8R	in	0,08	0,87	0,13	0,44			2,06	5,02	4,53	0,16	0,48	0,59	0,47
2P2H	in	0,05	0,79	2,17	2,98	1,42		2,87	4,13	2,40	0,88	2,17	3,00	1,02
3EWH	out	0,02	0,65	0,33	1,10			0,37	4,60	2,61	0,32	0,97	2,17	0,14
3CPC	out	0,58	0,71	2,02	0,12			2,32	5,15	4,73	1,00	0,22	1,03	0,28
mit Wasser														
2P2H	in	0,01	0,44	2,52	0,30			1,83	5,33	3,56	0,92	0,86	0,67	0,99
3EWH	out	0,12	0,57	0,61	0,64			0,99	5,21	6,40	0,40	0,71	0,87	1,16

Tabelle 5.56: Howard RIE

Name	DFG-Motiv	normalisierter Score					ROCS	ohne Wasser CHEMScore-Rescore			GLIDEScore-Rescore			CONSENSUS-SCORE
		GOLD	GLIDE	PMF	PScore	PScore		GOLD	PMF	PScore	GOLD	PMF	PScore	
2W1I	in	1,59	1,37	1,38	2,01			1,68	0,33	0,63	0,72	0,85	1,88	2,58
2F4J	in	2,06	2,00	1,08	2,60	3,01		1,48	0,47	0,57	1,18	0,94	1,34	1,54
1MUO	out	5,73	2,86	1,91	2,84			0,05	0,17	0,12	1,50	1,30	0,16	1,98
2J5O	out	6,44	3,00	2,66	3,55			5,26	0,93	1,88	1,03	0,62	2,04	2,58
mit Wasser														
2F4J	in	2,25	2,01	1,44	2,71			2,67	1,22	1,23	1,11	0,99	1,94	0,02

Tabelle 5.57: Xia RIE

Name	DFG-Motiv	normalisierter Score						ohne Wasser			GLIDEScore-Rescore			CONSENSUS-SCORE
		GOLD	GLIDE	PMF	PScore	Rocs	CHEMScore-Rescore			GOLD	PMF	PScore		
							GOLD	PMF	PScore					
2J2I	in	1,04	1,94	6,60	6,76	2,000	2,003	0,008	0,008	2,701	2,632	1,554	3,005	
3BWF	in	0,66	0,95	6,26	6,24		1,96	0,38	0,38	1,91	1,89	2,12	2,85	
mit Wasser														
2J2I	in	0,95	1,88	6,56	6,30		2,02	1,69	1,96	0,99	1,00	1,23	2,30	
3BWF	in	0,34	0,76	6,66	6,69		1,94	2,84	1,75	1,04	0,40	2,06	2,89	

5.1.4 RIE für die jeweilige Methode

Tabelle 5.58: RIE für GOLDScore für die jeweilige Kinase

	Gesamt		ohne Wasser		mit Wasser	
	Ø	SD	Ø	SD	Ø	SD
CDK2	5,660	3,336	7,380	1,713	3,367	3,896
CK18	3,649	3,031	6,110	0,865	1,188	1,609
P38	1,690	1,585	1,399	0,819	2,274	3,087
CHK1	2,997	0,677	–	–	–	–
ABL1	3,402	1,887	3,178	1,831	1,366	1,784
P38	1,692	0,742	1,651	0,946	1,775	0,215
CDK7	0,657	–	–	–	–	–
Aurora-2	3,424	2,507	3,670	2,824	2,438	–
CDK4	0,247	0,343	–	–	–	–
PIM	4,605	2,524	2,945	2,315	6,264	1,656
GSK	3,528	1,194	2,525	0,159	4,531	0,477
C-Jun	1,402	0,383	–	–	–	–
P38	0,867	0,709	0,802	0,505	0,996	1,304
MET	2,605	1,120	2,326	1,074	3,721	–
Aurora	0,371	0,278	0,425	0,288	0,153	–
CDK2	0,057	0,064	–	–	–	–
CDK4	1,986	2,154	–	–	–	–
CDK5	3,586	2,413	2,290	1,247	6,18	–
CDK7	0,943	–	–	–	–	–
CDK9	3,498	4,533	–	–	–	–
CDC7	0,560	0,090	0,623	–	0,496	–
VGFR	0,143	0,217	0,181	0,265	0,067	0,081
Aurora	3,614	2,281	3,956	2,482	2,249	–
PIM	0,749	0,318	0,854	0,268	0,644	0,433

Tabelle 5.59: RIE für GLIDEScore für die jeweilige Kinase

	Gesamt		ohne Wasser		mit Wasser	
	Ø	SD	Ø	SD	Ø	SD
CDK2	3,562	2,020	4,843	1,559	1,854	0,969
CK18	2,893	1,708	4,360	0,337	1,426	0,184
P38	2,291	1,273	2,834	1,182	1,206	0,615
CHK1	2,937	0,091	–	–	–	–
ABL1	2,719	0,728	2,868	0,669	2,520	0,882
P38	1,774	0,678	2,004	0,742	1,314	0,118
CDK7	1,032	–	–	–	–	–
Aurora-2	3,217	1,631	3,353	1,850	2,671	–
CDK4	1,001	1,329	–	–	–	–
PIM	0,607	0,077	0,650	0,052	0,564	0,086
GSK	3,055	0,551	3,142	0,880	2,968	0,327
C-Jun	0,487	0,040	–	–	–	–
P38	1,944	0,757	2,075	0,905	1,682	0,452
MET	3,672	0,878	3,372	0,654	4,872	–
Aurora	0,350	0,166	0,369	0,185	0,273	–
CDK2	4,416	1,874	–	–	–	–
CDK4	1,094	0,782	–	–	–	–
CDK5	3,026	1,709	2,043	0,214	4,991	–
CDK7	4,301	–	–	–	–	–
CDK9	3,494	3,755	–	–	–	–
CDC7	3,933	1,329	4,872	–	2,993	–
VGFR	0,670	0,155	0,754	0,094	0,502	0,092
Aurora	2,248	0,678	2,307	0,767	2,012	–
PIM	1,384	0,613	1,448	0,701	1,321	0,787

Tabelle 5.60: RIE für PMF für die jeweilige Kinase

	Gesamt		ohne Wasser		mit Wasser	
	Ø	SD	Ø	SD	Ø	SD
CDK2	0,001	0,000	0,001	0,000	0,001	0,000
CK18	1,550	0,852	1,176	0,279	1,924	0,852
P38	1,106	0,535	0,794	0,252	1,731	0,271
CHK1	2,531	0,688	–	–	–	–
ABL1	0,013	0,008	0,011	0,007	0,015	0,013
P38	3,707	1,476	3,771	1,153	3,579	2,618
CDK7	0,090	–	–	–	–	–
Aurora-2	4,662	2,048	4,865	2,305	3,849	–
CDK4	1,152	0,203	–	–	–	–
PIM	0,533	0,127	0,537	0,112	0,529	0,190
GSK	1,896	0,649	1,343	0,124	2,449	0,155
C-Jun	4,820	0,193	–	–	–	–
P38	0,551	0,632	0,559	0,777	0,534	0,431
MET	0,365	0,251	0,455	0,171	0,002	–
Aurora	0,001	0,000	0,001	0,000	0,001	–
CDK2	0,001	0,000	–	–	–	–
CDK4	0,001	0,000	–	–	–	–
CDK5	0,001	0,000	0,001	0,000	0,001	–
CDK7	0,001	–	–	–	–	–
CDK9	0,001	0,000	–	–	–	–
CDC7	2,864	0,593	2,445	–	3,283	–
VGFR	1,296	1,053	1,163	1,082	1,562	1,351
Aurora	1,695	0,617	1,759	0,693	1,440	–
PIM	6,519	0,178	6,428	0,241	6,610	0,065

Tabelle 5.61: RIE für PSCORE für die jeweilige Kinase

	Gesamt		ohne Wasser		mit Wasser	
	Ø	SD	Ø	SD	Ø	SD
CDK2	0,061	0,107	0,001	0,000	0,141	0,133
CK18	2,879	0,211	2,954	0,238	2,803	0,232
P38	2,589	0,318	2,560	0,314	2,646	0,445
CHK1	6,290	0,660	–	–	–	–
ABL1	0,032	0,056	0,031	0,062	0,058	0,083
P38	2,488	0,859	2,367	0,950	2,730	0,895
CDK7	1,843	–	–	–	–	–
Aurora-2	1,382	1,083	1,495	1,216	0,931	–
CDK4	0,596	0,432	–	–	–	–
PIM	0,791	0,050	0,771	0,014	0,811	0,076
GSK	0,369	0,102	0,456	0,012	0,283	0,040
C-Jun	2,175	0,310	–	–	–	–
P38	0,459	0,419	0,257	0,341	0,864	0,191
MET	0,001	0,000	0,001	0,000	0,001	–
Aurora	1,128	0,596	1,187	0,672	0,894	–
CDK2	0,644	0,909	–	–	–	–
CDK4	0,288	0,506	–	–	–	–
CDK5	0,180	0,307	0,288	0,506	0,535	–
CDK7	0,749	–	–	–	–	–
CDK9	0,375	0,529	–	–	–	–
CDC7	6,501	0,585	6,087	–	6,915	–
VGFR	0,927	1,059	1,157	1,280	0,467	0,243
Aurora	2,741	0,551	2,749	0,636	2,711	–
PIM	6,497	0,265	6,497	0,366	6,498	0,277

Tabelle 5.62: RIE für RocS für die jeweilige Kinase.

	Gesamt	
	Ø	SD
CDK2	6,587	–
CK18	4,387	–
P38	4,387	–
CHK1	4,829	–
ABL1	3,800	–
P38	2,994	–
CDK7	1,484	–
Aurora-2	2,884	–
CDK4	1,448	–
PIM	3,952	–
GSK	3,075	–
C-Jun	1,687	–
P38	1,598	–
MET	4,112	–
Aurora	2,865	–
CDK2	2,481	–
CDK4	2,481	–
CDK5	2,481	–
CDK7	2,481	–
CDK9	2,481	–
CDC7	3,360	–
VGFR	1,416	–
Aurora	3,013	–
PIM	2,000	–

Tabelle 5.63: RIE für CONSENSUS-SCORE für die jeweilige Kinase.

	Gesamt		ohne Wasser		mit Wasser	
	Ø	SD	Ø	SD	Ø	SD
CDK2	4,980	2,027	6,553	0,360	2,884	2,884
CK18	2,680	0,467	3,041	0,363	2,319	0,035
P38	3,328	1,105	3,652	1,135	2,679	0,989
CHK1	1,696	0,879	1,696	0,879	–	–
ABL1	1,938	1,783	2,141	1,811	1,668	1,615
P38	1,736	1,373	1,620	1,494	1,969	1,602
CDK7	3,167	–	3,167	–	–	–
Aurora-2	2,945	0,535	2,865	0,581	3,268	–
CDK4	2,087	1,252	2,087	1,252	–	–
PIM	3,632	1,577	2,726	2,005	4,538	0,393
GSK	2,708	0,706	2,925	0,114	2,491	1,138
C-Jun	0,254	0,247	0,254	0,247	–	–
P38	2,829	1,809	3,457	1,918	1,571	0,758
MET	3,322	2,074	3,369	2,391	3,131	–
Aurora	0,357	0,450	0,192	0,296	1,019	–
CDK2	5,739	1,570	5,739	1,570	–	–
CDK4	6,033	1,074	6,033	1,074	–	–
CDK5	5,521	0,597	5,364	0,752	5,834	–
CDK7	5,703	–	5,703	–	–	–
CDK9	4,100	4,530	4,100	4,530	–	–
CDC7	4,234	1,729	5,456	–	3,011	–
VGFR	0,675	0,434	0,475	0,386	1,076	0,117
Aurora	1,737	1,056	2,167	0,506	0,019	–
PIM	2,760	0,315	2,929	0,108	2,592	0,415

5.2. Moleküldynamik-Simulation

5.2.1 Vorbereitung der Simulation

Die Liganden wurden, wie folgt, mit dem Program *Antechamber* parametrisiert:

```
antechamber -i ligand.mol2 -fi mol2 -o ligand_p.mol2 -fo mol2 \\  
-c bcc -j 4 -at gaff
```

```
parmchk -i ligand_p.mol2 -o ligand.frcmod -f mol2
```

Mittels LEaP wurden die Komplexe aus Ligand und Protein erzeugt, und so für die anschließende Simulation vorbereitet:

```
source leaprc.ff99SB  
source leaprc.gaff  
lig = loadmol2 ligand_p.mol2  
loadamberparams ligand.frcmod  
prot = loadpdb protein.pdb  
com = combine {lig prot}  
saveamberparm lig lig.top lig.crd  
saveamberparm prot rec.top rec.crd  
saveamberparm com com.top com.crd  
addions com Na+ 0  
solvateoct com TIP3PBOX 12  
saveamberparm com w_com.top w_com.crd  
quit
```

Sofern nicht anders erwähnt wurden alle MD nach dem gleichen Schema durchgeführt. Für \$MAX wurde in Abhängigkeit vom verwendeten Protein die Anzahl der jeweilige Summe aus Aminosäureresten und Ligand eingetragen. Der eigentliche Simulationsschritt wurde dabei sooft wiederholt, bis die gewünschte Simulationszeit (10 ns) erreicht war.

Dazu wurden folgende Eingabeskripte verwendet:

1. Minimierung

```
Minimization  
&cntrl  
imin=1,maxcyc=1000,ncyc=500,  
cut=8.0,ntb=1,  
ntc=2,ntf=2,  
ntpr=100,  
ntr=1, restraint_wt=2.0, restraintmask=':1-$MAX',  
/  
/
```

2. Aufheizen

```
heat
  &cntrl
    imin=0,irest=0,ntx=1,
    nstlim=25000,dt=0.002,
    ntc=2,ntf=2,
    cut=8.0, ntb=1,
    ntp=500, ntwx=500,
    ntt=3, gamma_ln=2.0,
    tempi=0.0, temp0=300.0,
    ntr=1,
    restraint_wt=2.0, restraintmask=':1-$MAX',
    nmropt=1
  /
&wt TYPE='TEMPO', istep1=0, istep2=25000,
value1=0.1, value2=300.0, /
&wt TYPE='END' /
```

3. Dichtenausgleich

```
density
  &cntrl
    imin=0,irest=1,ntx=5,
    nstlim=25000,dt=0.002,
    ntc=2,ntf=2,
    cut=8.0, ntb=2, ntp=1, taup=1.0,
    ntp=500, ntwx=500,
    ntt=3, gamma_ln=2.0,
    temp0=300.0,
    ntr=1, restraintmask=':1-$MAX',
    restraint_wt=2.0,
  /
```

4. Equilibrierung

```
Equi
  &cntrl
    imin=0,irest=1,ntx=5,
    nstlim=250000,dt=0.002,
    ntc=2,ntf=2,
    cut=8.0, ntb=2, ntp=1, taup=2.0,
    ntp=1000, ntwx=1000,
    ntt=3, gamma_ln=2.0,
    temp0=300.0,
  /
```

5. Simulation

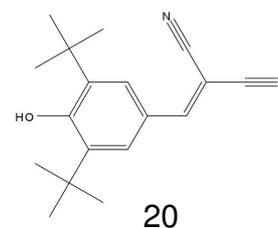
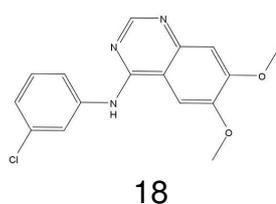
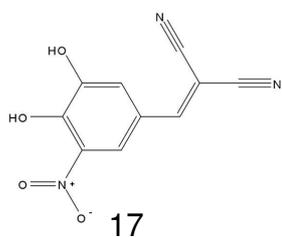
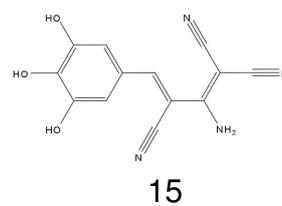
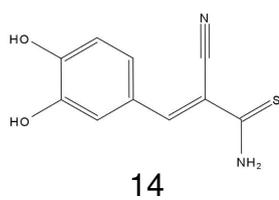
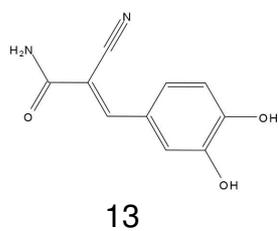
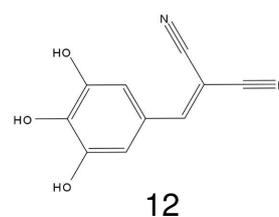
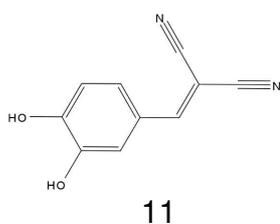
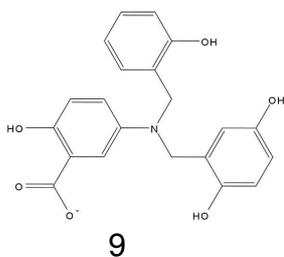
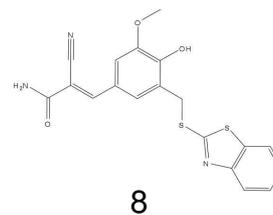
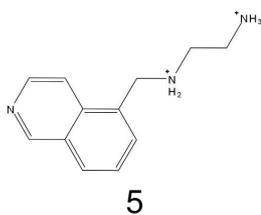
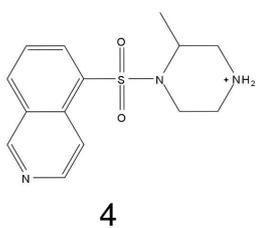
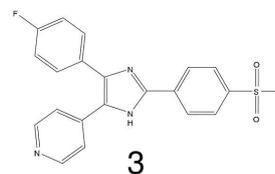
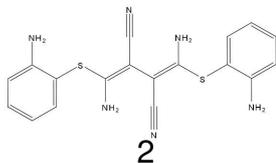
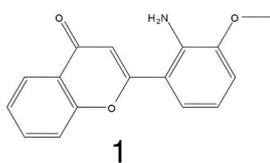
```
MD run
  &cntrl
    imin = 0,  irest=1,  ntx=5,
    ntt=3,  temp0=300,  gamma_ln=5.0,
    ntp=1,  taup=2.0,
    ntb=2,  ntc=2,  ntf=2,
    nstlim=500000,
    ntwe=100,  ntwx=1000,  ntp=200,
  /
```

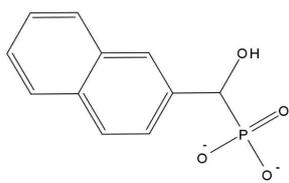
5.3. IGFR, EGFR, VEGFR

5.3.1 Durchführung des Affinitäts Assay

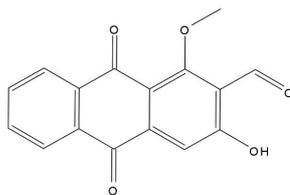
Diese wurden in 96-Well FlashPlates™ mit 50 µL Reaktionsvolumen durchgeführt. Dieses Volumen bestand aus 20 µL Puffer, 5 µL ATP in Wasser gelöst, 5 µL der Probensubstanz, 10 % DMSO sowie 10 µL Substrat und 10 µL Enzymlösung. Der Puffer enthielt HEPES/NaOH (60 nM, pH 7,5), MgCl₂ (3 mM), Na₃VO₄ (3 µM), DTT (1,2 mM), PEG 20 000 (50 µL⁻¹), [γ -³³P]-ATP (15 µM, 7,5x10⁴ cpm pro Well). Die Inkubationszeit betrug 80 Minuten bei 30 °C und die Reaktion wurde mit H₃PO₄ (50 µL, 2 % (v/v)) gestoppt, und die Wells wurden darauf mit H₂O (2x200 µL) oder NaCl (2x200 µL, 0,9 % (w/v)) ausgewaschen.

5.4. PRK1

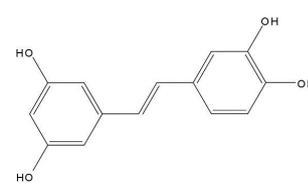




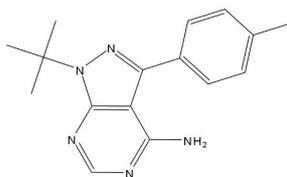
21



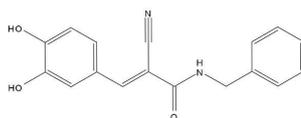
22



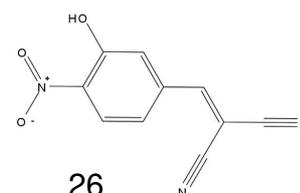
23



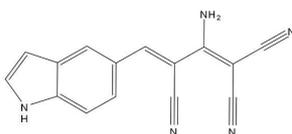
24



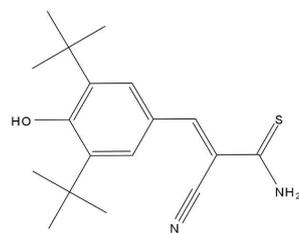
25



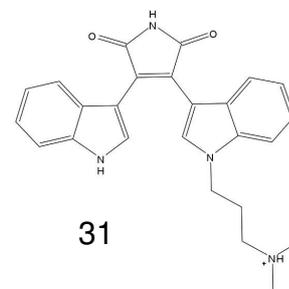
26



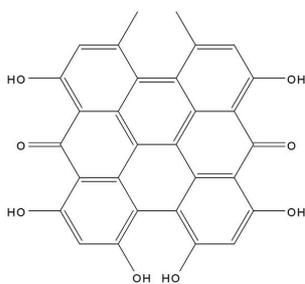
27



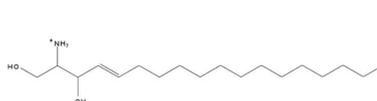
28



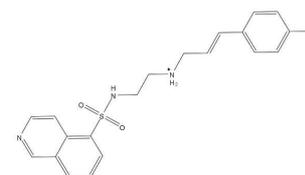
31



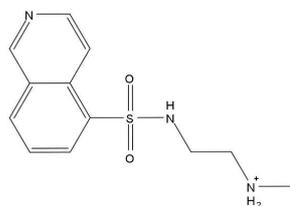
32



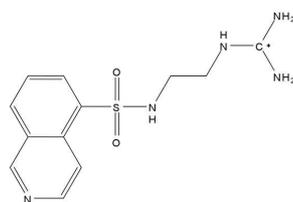
34



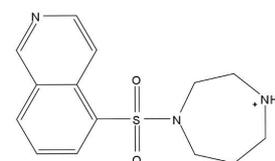
35



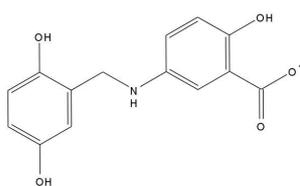
36



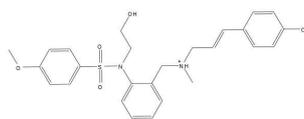
37



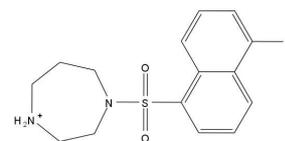
38



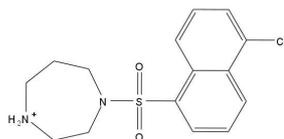
39



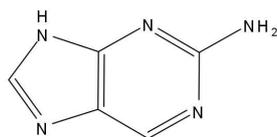
41



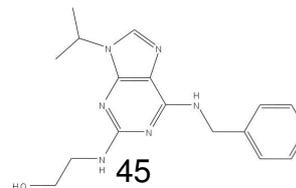
42



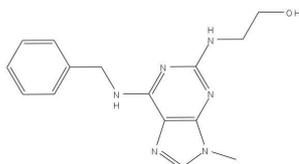
43



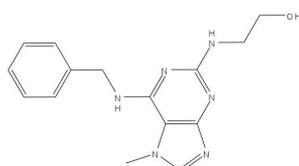
44



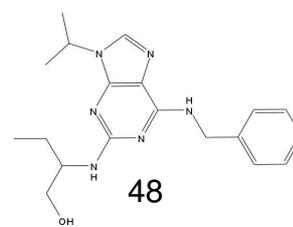
45



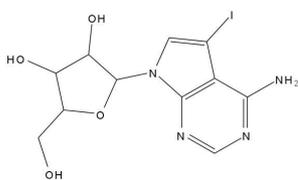
46



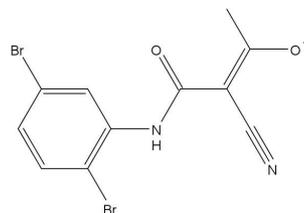
47



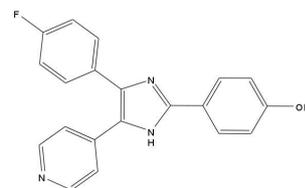
48



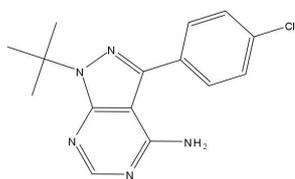
49



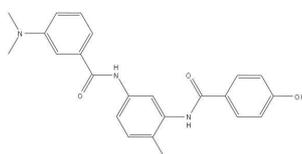
50



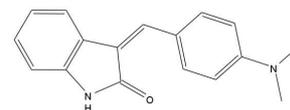
51



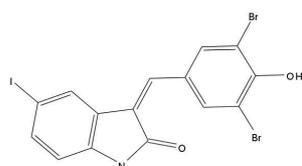
52



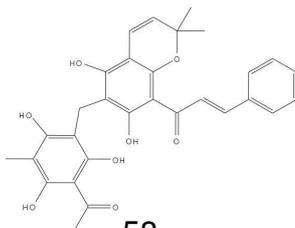
53



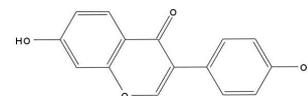
54



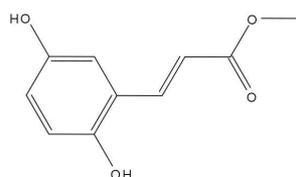
56



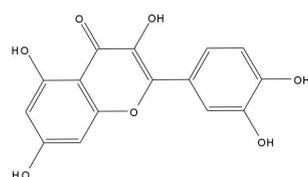
58



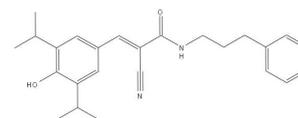
60



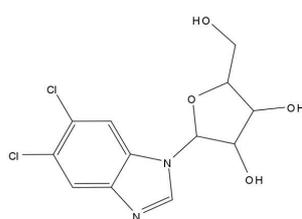
61



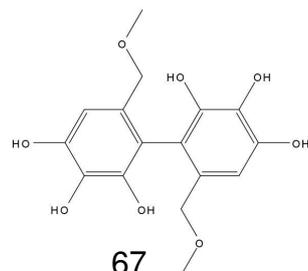
62



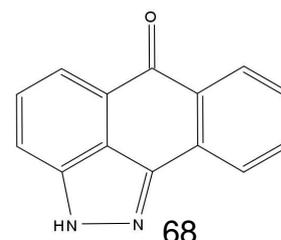
63



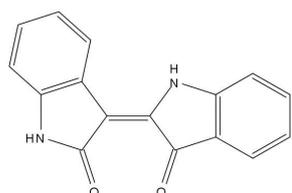
66



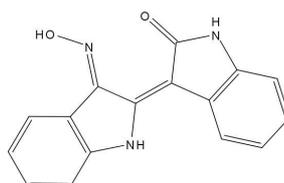
67



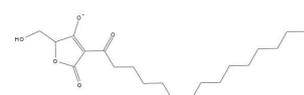
68



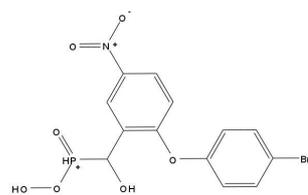
69



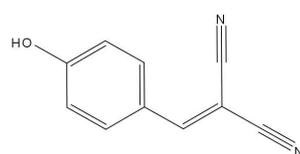
70



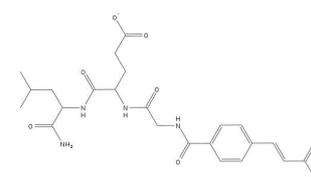
76



77



83



84

ABKÜRZUNGEN

AS Aminosäure

ATP Adenosintri-phosphat

cAMP zyklisches Adenosinmonophosphat

FDA Food and Drug Administration

RNA Ribonukleinsäure

EGFR epidermal growth factor receptor

Da Dalton

SF Scoring-Funktion

VS Virtuelles Screening

EF Enrichment Factor

RTK Rezeptor-Tyrosin-Kinasen

MD Moleküldynamik-Simulation

PB/SA Poisson-Boltzmann/Surface Area

GB/SA Generalized Born/Surface Area

MM-PB/SA Molecular Mechanics Poisson-Boltzmann/Surface Area

MM-GB/SA Molecular Mechanics generalized Born/Surface Area

LR-MMPBSA Linear Response Molecular Mechanics Poisson-Boltzmann/Surface Area

Danksagung

Die vorliegende Arbeit wurde in der Zeit von 2007 bis 2011 unter Leitung von Prof. Dr. Wolfgang Sippl im Institutsbereich für Pharmazeutische Chemie der Martin-Luther-Universität Halle-Wittenberg angefertigt. Für die Unterstützung seinerseits, die fruchtbaren Diskussionen und die Möglichkeit diese Arbeit anzufertigen zu können möchte ich ihm danken.

Dr. Andreas Hilgeroth und Dr. Martin Schmidt für die gute Zusammenarbeit bei den hier gestellten Fragen.

Auch allen Mitgliedern der Arbeitsgruppe möchte ich danken. Besonders hervorheben möchte ich dabei Dr. René Meier für die Wartung unseres Clusters, und für die vielen Dinge, die er mir im Bereich Netzwerke beigebracht hat. Dr. Ralf Heinke für die gute Zeit im Büro und für die Orientierung in Halle. Dr. Dina Rooba, Martin Pippel und Michael Scharfe für die Mühen, die sie mit dem Korrekturlesen meiner Arbeit hatten. Und allen anderen ehemaligen, und aktuellen Mitgliedern der Arbeitsgruppe für das schöne Arbeitsklima, vielen Diskussionen – nicht nur Fachbezogen – und die tolle Zeit in Halle.

Auch der Mittagessensgruppe muss gedankt werden. Ohne sie hätte ich meine Gattin Monika nie kennengelernt. Ihr gilt ein besonderer Dank für ihr Verständnis, ihre Liebe, und ihre bedingungslose Unterstützung, die mir sehr geholfen hat.

Allen weiteren Kollegen, vor allem denen aus dem Praktikum Qualitative Analytik sei ebenfalls gedankt. Dr. Jens Sproß sei dabei besonders erwähnt.

Und zu guter letzt meinen Eltern. Ihnen schulde ich mehr als Dank für all Ihre Unterstützung und Hilfe in jeder Lebenslage.

Curriculum Vitae

Persönliche Angaben

Name, Vorname	Erlenkamp, Daniel German
Geburtsdatum	09. Januar 1979
Geburtsort	Mönchengladbach
Familienstand	verheiratet
Staatsangehörigkeit	deutsch

Bildungsgang

07/2007–09/2011	Promotionsstudium am Institut für Pharmazie der MLU Halle-Wittenberg, Abteilung Medizinische Chemie
04/2006–12/2006	Diplomand am Institut für Physikalische Biologie der Heinrich-Heine-Universität Düsseldorf in der Arbeitsgruppe von Prof. Dr. Gerhard Steger
10/1999–12/2006	Studium der Biologie an der Heinrich-Heine-Universität Düsseldorf

Tätigkeiten

seit 02/2014	Wissenschaftlicher Mitarbeiter am Institut für Pharmazeutische und Medizinische Chemie an der Heinrich-Heine-Universität Düsseldorf
10/2011–12/2012	Wissenschaftlicher Mitarbeiter am Institut für Biochemie Universität Leipzig

Publikationen

Köhler, J., Erenkamp, G., Eberlin, A., Rumpf, T., Slynko, I., Metzger, E., Schüle, R., Sippl, W., Jung, M., *Lestaurtinib Inhibits Histone Phosphorylation and Androgen-Dependent Gene Expression in Prostate Cancer Cells*, PLoS ONE, 2012;7(4):e34973

Krug, M., Wichapong, K., Erenkamp, G., Sippl, W., Schächtele, C., Totzke, F., Hilgeroth, A., *Discovery of 4-Bezylamino Substituted α -Carbolines as a novel Class of Receptor Tyrosine Kinase Inhibitors*, ChemMedChem, 6, 63-72, 2011

Krug, M., Erenkamp, G., Sippl, W., Schächtele, C., Totzke, F., Hilgeroth, A., *Discovery and selectivity profiling of 4-bezylamino 1-aza-9-oxafluorene derivatives as lead structures for IGF-1P inhibitors*, Bioorg Med Chem Letters, 20, 6915-9, 2010

Wollman, J., Baumert, C., Erenkamp, G., Sippl, W., Hilgeroth, A., *Novel Insight in Inhibitor Binding of Highly Symmetric HIV-1 Protease*, ChemBioChem, 9, 874-78, 2008

Posterbeiträge

Slynko, I., Erenkamp, G., Köhler, J., Rumpf, T., Metzger, E., Schühle, R., Jung, M., Sippl, W., *Identification of PRK1 kinase inhibitors by virtual screening and binding free energy calculations*, 19th Summerschool Medicinal Chemistry Regensburg

Rumpf, T., Erenkamp, G., Metzger, E., Eberlin, A., Sippl, W., Schühle, R., Jung, M., *The clinical candidate Lestaurtinib inhibits the epigenetic kinase PRK1*, 1st Max-Planck Epigenetic Meeting Freiburg

Rumpf, T., Erenkamp, G., Metzger, E., Eberlin, A., Sippl, W., Schüle, R., Jung, M., *Lestaurtinib inhibits the kinase PRK1 in vivo*, DPhG-Jahrestagung Braunschweig

Rumpf, T., Erenkamp, G., Metzger, E., Eberlin, A., Sippl, W., Schühle, R., Jung, M., *The clinical candidate Lestaurtinib inhibits the epigenetic kinase PRK1*, EMBO Conference Series, Chemical Biology Heidelberg

Kumawat, N., Kannan, r., Erenkamp, G., Krug, M., Hilgeroth, A., Sippl, W., *Molecular docking and molecular dynamics simulations of GSK3-kinase inhibitors*, 5th Summerschool Medicinal Chemistry Regensburg

Erenkamp, G., Sippl, W., *Evaluating the performance of structure-and ligand-based screening methods for kinase inhibitors*, 5th Summerschool Medicinal Chemistry Regensburg

Erenkamp, G., Marx, C. K., Pietsch, M., Sippl, W., *Predicting thermostability of random mutants of microbial transglutaminase*, DPhG-Jahrestagung Jena

Erenkamp, G., Heinke, R., Meier, R., Sippl, W., *Evaluation of structure- and pharmacophore-based virtual screening methods*, 3rd German Conference on Cheminformatics Goslar

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides Statt, dass ich diese Dissertationsschrift selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen verwendet habe, und die daraus wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Die Dissertation wurde ausschließlich der Mathematisch-Naturwissenschaftlichen-Fakultät der Martin- Luther-Universität Halle-Wittenberg vorgelegt, und an keiner anderen universitären Einrichtung im In- und Ausland zur Erlangung des Doktorgrades eingereicht

Halle, 17. Juli 2014

German Erenkamp

