

Strukturbasierte Leitstruktursuche ATP-kompetitiver Inhibitoren der Gyrase Untereinheit B

Dissertation zur Erlangung des Doktorgrades
an der Fakultät für Mathematik, Informatik und Naturwissenschaften
Fachbereich Chemie der Universität Hamburg

vorgelegt von

Moritz Niels Münsterberg

aus Reinbek

Hamburg, 2015

Tag der Disputation: 16.10.2015

Folgende Gutachter empfehlen die Annahme der Dissertation:

Name: Prof. Dr. Hans-Jürgen Duchstein

Name: Prof. Dr. Andrew Torda

Die vorliegende Arbeit wurde in der Zeit von Mai 2011 bis August 2015 am Institut für Pharmazie im Fachbereich Chemie der Universität Hamburg angefertigt. Sie entstand unter der Leitung

von

Herrn Professor Dr. Hans-Jürgen Duchstein,

dem ich an dieser Stelle für die freundliche Aufnahme in seinen Arbeitskreis danken möchte.

Mein besonderer Dank gilt außerdem

Herrn Dr. Thomas Lemcke

für die Überlassung des interessanten Themas, für seine engagierte Betreuung und Unterstützung sowie für zahlreiche wertvolle Anregungen während der Durchführung dieser Arbeit.

Danken möchte ich zudem

Herrn Professor Dr. Andrew Torda

für die Übernahme des Zweitgutachtens.

Inhaltsverzeichnis

1	Einleitung und Zielsetzung.....	1
1.1	Topoisomerasen	4
1.2	Strukturbasierte Leitstruktursuche	9
1.3	Kompetitive Inhibitoren der ATPase der GyrB-Untereinheit	14
1.4	Zielsetzung.....	23
2	Analyse der Zielstrukturen.....	27
2.1	Ausgangsstrukturen.....	27
2.2	Vorbereitung der Röntgenkristallstrukturen.....	40
2.2.1	Material und Methoden	40
2.2.2	Röntgenkristallstrukturen	42
2.3	Analyse der Bedeutung der Bindetaschen Wassermoleküle	47
2.3.1	Material und Methoden	48
2.3.2	Wassermoleküle in der ATP-Bindetasche der Gyrase	50
3	Bibliothekerstellung.....	61
3.1	Chemischer Raum.....	61
3.2	Material und Methoden	64
3.2.1	ZINC-Datenbank	64
3.2.2	Mona	65
3.3	Filterkriterien für die ZINC-Datenbank.....	66
3.4	Anwendung des Filters auf den „Drug-Like Clean Dataset“	73
4	Virtual Screening.....	79
4.1	Material und Methoden	80
4.1.1	cRAISE	80
4.1.2	Evaluierungsmethoden.....	82
4.1.3	Directory of Useful Decoys – Enhanced	85
4.2	Docking mit cRAISE	87

4.2.1	Erstellung der Pharmakophore	87
4.2.2	Erstes cRAISE-Docking	92
4.2.3	Zweites cRAISE-Docking.....	98
5	Postprocessing.....	103
5.1	Material und Methoden	106
5.1.1	Piecewise Linear Potential Score.....	106
5.1.2	Glide.....	106
5.1.3	GlideScore.....	107
5.1.4	DSX Score.....	108
5.1.5	Prime MM-GBSA-Score	108
5.1.6	Daylight-Fingerprint.....	109
5.1.7	Clustering.....	109
5.2	Planung.....	110
5.3	Rescoring	111
5.4	Leader-Clustering	117
5.5	Redocking mit Glide.....	120
5.6	Visuelle Auswahl.....	123
5.6.1	3G7E_1	124
5.6.2	3G7E_2	129
5.6.3	1EI1	131
5.6.4	1AJ6	137
5.6.5	Erste Auswahl	141
5.6.6	Zweite Auswahl	144
6	Assay-Entwicklung	151
6.1	Assays	151
6.2	Material und Methoden	157
6.2.1	Supercoiling Assay	157
6.2.2	Absorption des Phosphomolybdat-Farbstoff-Komplexes	157

6.2.3	Fluoreszenzauslöschung des Phosphomolybdat-Farbstoff-Komplexes	157
6.2.4	Regressionen für IC_{50} -Werte.....	159
6.3	Absorptionsmethode.....	160
6.4	Fluoreszenzauslöschungsmethode	169
7	Biologische Testung	183
7.1	Erste Auswahl	183
7.2	Zweite Auswahl	185
8	Evaluierung	191
9 a	Zusammenfassung	203
9 b	Summary.....	205
10	Ausblick.....	209
11	Experimenteller Teil.....	213
11.1	Arbeitstechniken	213
11.2	Arbeitsvorschriften.....	216
11.3	Computermethoden.....	220
12	Anhang.....	229
13	Literatur	239
14	Gefahrstoffverzeichnis.....	261

Tagungsbeiträge

Teile der Arbeit wurden bereits vorab in Form von Tagungsbeiträgen an folgender Stelle veröffentlicht.

- 2014 Münsterberg, M.; Fransson, I.; Lemcke T., *Hit Search for ATPase Inhibitors of the GyrB-Subunit*, EFMC-ISMC, XXIII International Symposium on Medicinal Chemistry, Lissabon.
- 2014 Münsterberg, M.; Fransson, I.; Lemcke T., *Virtual Screening and Biological Evaluation of New Inhibitors of the ATPase-Domain in GyrB*, DPhG: Jahrestagung 2014, Trends and Perspectives in Pharmaceutical Sciences, Frankfurt a. Main.
- 2012 Münsterberg, M; Lemcke T., *On the Importance of Water Molecules on Ligand Binding in the ATPase of GyrB*, DPhG: Jahrestagung 2012, Moleküle, Targets und Tabletten – Translationale Forschung für Arzneimittel der Zukunft, Greifswald.
- 2012 Münsterberg, M; Lemcke T., *Evaluation of Water Molecules in the ATPase Binding Site of GyrB*, 19th EuroQSAR, Knowledge Enabled Ligand Design, Wien.

Abkürzungsverzeichnis

4pl-Funktion	4-Parameter-Logarithmische Funktion (bekannt als Hill-Funktion)
5pl-Funktion	5-Parameter-Logarithmische Funktion
A_(calc)	(berechnete) Absorption
Abk.	Abkürzung
ADME	s. (L)ADME-(T)
ADP	Adenosindiphosphat
ANP	5'-Adenylyl- β - γ -imidodiphosphat
ATP	Adenosintriphosphat
B-Factor	Temperaturfaktor
BP	Basenpaar
BSA	bovines Serumalbumin
bspw.	beispielsweise
bzw.	beziehungsweise
(c)LogP	Logarithmus des (berechneten) Octanol-Wasser-Verteilungskoeffizienten
CSD	Cambridge Structural Database
DMSO	Dimethylsulfoxid
DNA	Desoxyribonukleinsäure (engl. <i>Deoxyribonucleic acid</i>)
DTT	Dithiothreitol
EDTA	Ethylendiamintetraessigsäure (engl. <i>Ethylenediaminetetraacetic acid</i>)
engl.	englisch
ESI	Elektrosprayionisation
GPU	Grafikprozessor (engl. <i>graphics processing unit</i>)
GUI	Grafische Benutzeroberfläche (engl. <i>graphical user interface</i>)
GyrA	A-Untereinheit der DNA Gyrase
GyrB	B-Untereinheit der DNA Gyrase
HPLC	Hochleistungsflüssigkeitschromatographie (engl. <i>high performance liquid chromatography</i>)
HSP	Hitzeschockproteine (engl. <i>Heat Shock Protein</i>)
HT(v)S	engl. <i>High Throughput (virtual) Screening</i> (s. Glossar)
IC₅₀	Konzentration bei der die Aktivität zu 50 % gehemmt wird
IUPAC	Internationale Union für reine und angewandte Chemie (engl. <i>International Union of Pure and Applied Chemistry</i>)
(L)ADME-(T)	Vorgänge im Rahmen der Pharmakokinetik eines Arzneistoffes: (Freisetzung) Absorption, Distribution, Metabolisierung, Exkretion-(Toxizität)
Mio.	Million
NAD⁺/NADH	oxidierte/reduzierte Form von Nicotinamidadenindinukleotid
PDB	Protein Data Base
pKa	negativ dekadischer Logarithmus der Säurekonstante
PSA	polare Oberfläche (engl. <i>polar surface area</i>)
RMSD	mittlere quadratische Abweichung (engl.: <i>root mean square deviation</i>); s. Glossar
ROC-Kurve	Grenzwertoptimierungskurve (engl. <i>Receiver Operating Characteristic</i>)
S.	Seite
s.	siehe
s. o.	siehe oben

s. u.	siehe unten
SAR	Struktur-Wirkungsbeziehung (engl. <i>structure-activity relationship</i>)
SAS	dem Lösungsmittel zugängliche Oberfläche (engl. <i>solvent accesible surface</i>)
SBVS	strukturbasiertes virtuelles Screening (engl. <i>structure-based virtual screening</i>)
SQL	Datenbanksprache (engl. <i>Structured Query Language</i>)
s_{rel}	relative Standardabweichung
TAE-Puffer	Tris-Acetat-EDTA-Puffer
Topo IV	DNA Topoisomerase IV
Tris	Tris-(hydroxymethyl)-aminomethan (freie Base)
Tris-HCl	Tris-(hydroxymethyl)-aminomethan-hydrochlorid
usw.	und so weiter
vdW	Van-der-Waals
WHO	Weltgesundheitsorganisation (engl. <i>World Health Organization</i>)
\bar{x}	Mittelwert
z. B.	zum Beispiel

Für Aminosäuren wurden folgende 3- und 1-Buchstabencodes verwendet:

Aminosäure	3-Buchstabencode	1-Buchstabencode
Alanin	Ala	A
Arginin	Arg	R
Asparagin	Asn	N
Asparaginsäure	Asp	D
Cystein	Cys	C
Glutamin	Gln	Q
Glutaminsäure	Glu	E
Glycin	Gly	G
Histidin	His	H
Isoleucin	Ile	I
Leucin	Leu	L
Lysin	Lys	K
Methionin	Met	M
Phenylalanin	Phe	F
Prolin	Pro	P
Serin	Ser	S
Threonin	Thr	T
Tryptophan	Trp	W
Tyrosin	Tyr	Y
Valin	Val	V

Glossar

In der medizinischen Chemie haben sich verschiedene, häufig englische, Fachausdrücke etabliert. Deren Bedeutungen oder Übersetzungen sind nicht selbsterklärend, so dass in diesem Glossar eine kurze Definition gegeben wird. Bei entsprechender Relevanz werden die Begriffe zusätzlich im Fließtext präzisiert.

Alignment	Ausrichtung von Elementen zueinander. Diese kann auf der Struktur (bei kleinen Molekülen oder Proteinen) oder auch auf Sequenz (Proteine) basieren.
Backbone	als Rückgrat eines Makromoleküls bezeichnete Kette von sich wiederholenden Einzelbausteinen; im engeren Sinn: bei Proteinen die Folge von Peptidbindungen
Bindungsmodus	eine wiederkehrende Pose mit definierten Interaktionsmuster für eine Substanz oder Substanzgruppe in einer Zielstruktur
Clash	engl. <i>Kollision</i> ; Überlappung zweier Strukturelemente; häufig von vdW-Radien
Clustering	Gruppieren einzelner Elemente eine Menge nach ähnlichen Eigenschaften
Docking	Generierung und Bewertung von Posen einer Substanz in einer Zielstruktur. Dies beinhaltet die Lösung a) des Docking-Problems: Generierung der wahren Pose und b) des Scoring-Problems: Identifizierung der wahren Pose aus falschen Posen.
GHKL-Domäne	konservierte ATPase, die bei funktional sehr unterschiedlichen Enzymen auftritt: <u>G</u> yrase, <u>H</u> eat Shock Protein 90, Histidink <u>i</u> nasen, Mut <u>L</u> (ein DNA-Mismatch-Reparaturprotein)
Hit-Struktur/Hit	eine in einem HT(v)S gefundene Substanz mit nachgewiesener Aktivität
HT(v)S	engl. <i>High-Throughput (virtual) Screening</i> , Testen einer großen Menge an Substanzen auf eine bestimmte biologische Aktivität. Bei einem HTS wird die Testreihe <i>in vitro/vivo</i> , bei einem HTvS mithilfe von Computermethoden (<i>in silicio</i>) durchgeführt.
Induced Fit	gegenseitige Anpassung von Zielstruktur und Ligand bei der Bildung eines Protein-Ligand-Komplexes
Leitstruktur	ein Repräsentant einer Substanzserie mit ausreichend Potential (z. B. Selektivität, Potenz), um für die Entwicklung eines Arzneistoffs verwendet zu werden
Lipinski's Rule of Five	s. Rule of Five
Minimierung	Geometrieoptimierung von Molekülen, Proteinen oder Protein-Ligand-Komplexen zum nächstliegenden lokalen energetischen Minimum

Moleküldynamik (MD)-Simulation	ermöglichen die Untersuchung von biologischen und chemischen Systemen auf atomarer Ebene über die Zeit, auch unter Berücksichtigung der Bewegung
Pose	die genaue räumliche Platzierung und Orientierung, welche ein Molekül in einer Zielstruktur einnimmt
Probe	engl. für Sonde; eine Sphäre, die eine bestimmte chemische oder physikalische Eigenschaften besitzt
Refinement	engl. für Verfeinerung, Anpassung/Verbesserung einer Struktur
Restraint	extern definierte energetische Barriere; häufig verwendet bei MD-Simulationen
RMSD	Wert [\AA] für die mittlere quadratische Abweichung der Distanz der (schweren) Atome zweier überlagerter Moleküle zueinander
Rule of Five (Ro5)	empirische „Faustregel“ von Anforderungen an eine Substanz, damit sie oral verfügbar ist: ≤ 10 Wasserstoffbrückenakzeptoren, ≤ 5 Wasserstoffbrückendonoren, ≤ 500 Da Molekülmasse und $\text{cLogP} \leq 5$
Scoring	Bewertung eines Systems; bei Docking-Experimenten: Bewertungen von Protein-Ligand-Komplexen zum Abschätzen der Freien Bindungsenthalpie. Sie werden dazu verwendet, das Scoring-Problem (s. Docking) zu lösen.
SMARTS	engl. <i>Simplified Molecular Input Line Entry Specification</i> , chemischer Strukturcode, um Moleküle in einer einfachen Zeichenfolge darzustellen
Zielstruktur	engl. <i>Target</i> , biologisches Molekül (z. B. Enzym, Rezeptor, DNA), das von einem Liganden moduliert (bspw. inhibiert) wird

1 Einleitung und Zielsetzung

Bakterielle Infektionen gehörten zu den größten Gefahren für die Menschheit. Schätzungen gehen davon aus, dass zwischen 1347 und 1351 ca. 30 bis 40 % der gesamten Bevölkerung der damals bekannten Welt alleine durch die Infektion mit *Yersinia Pestis*, bekannt als der Schwarze Tod, starben [1]. Dank höherer hygienischer Standards und der Entwicklung von Antibiotika konnten die Gefahren durch Bakterien in der jüngeren Geschichte erheblich gesenkt werden. Einen Meilenstein stellte dabei 1929 die Entdeckung des Penicillins durch FLEMING dar [2]. Bis zur industriellen Produktion und Anwendung von Penicillin dauerte es noch über zehn Jahre [3]. In der Folge wurde das so genannte Goldene Zeitalter der Antibiotikaentwicklung, 1940er bis 1960er, eingeläutet [4]. In dieser Zeit wurde ein Großteil der noch heute verwendeten Antibiotika entwickelt (s. Abbildung 1.1). Aber schon 1945 warnte FLEMING während seiner *Nobel Lecture* für den Nobelpreis [5]:

„It is not difficult to make microbes resistant to penicillin in the laboratory by exposing them to concentrations not sufficient to kill them, and the same thing has occasionally happened in the body.“

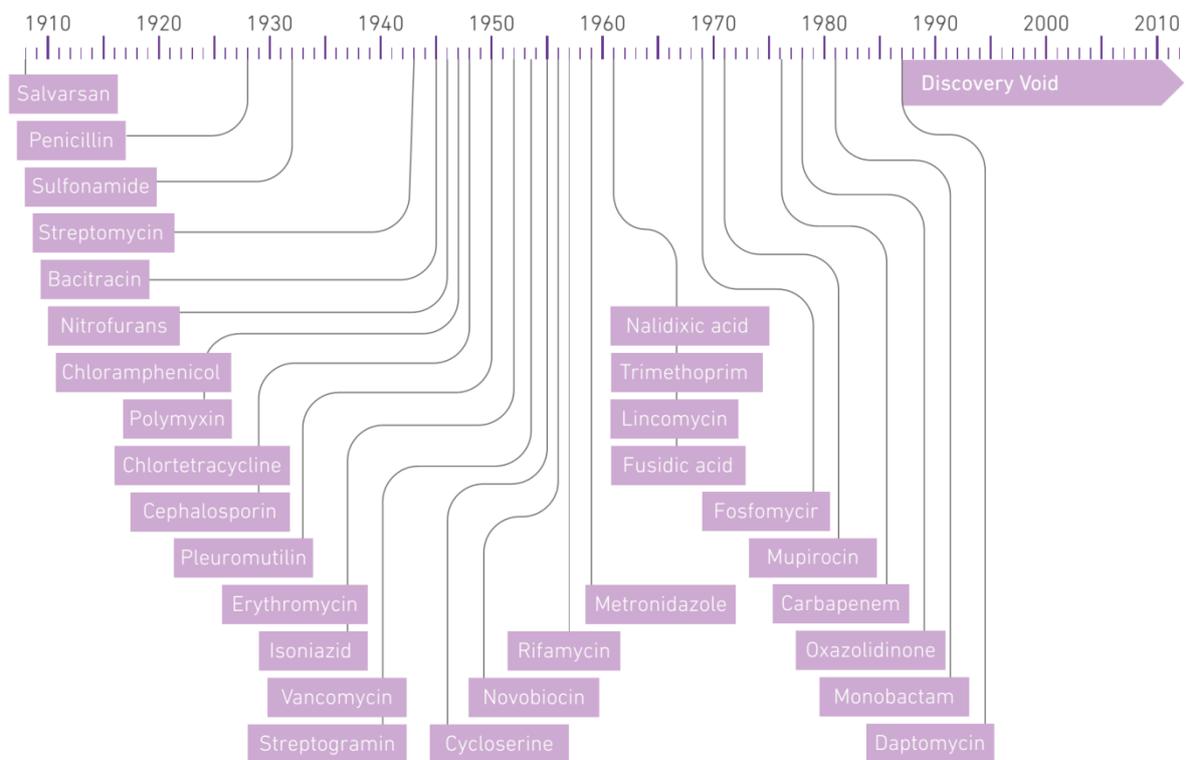


Abbildung 1.1 – Zeitstrahl über die Entwicklung neuer Vertreter von Antibiotikaklassen. Die Jahreszahlen geben das Jahr der Erstveröffentlichung oder der Patentanmeldung des ersten Vertreters der jeweiligen Klasse an. Seit 1987 wurden keine neuen Antibiotikaklassen mehr entdeckt („Discovery Void“). Abbildung aus dem WHO-Report: Antimicrobial Resistance – Global Report on Surveillance 2014 [6].

Damit erkannte FLEMMING schon zu Beginn der Antibiotika-Ära die Gefahr der Resistenzentwicklung von Bakterien gegenüber Antibiotika. Unter einer Resistenz wird die Widerstandsfähigkeit von Bakterien gegenüber einem eigentlich wirksamen Antibiotikum verstanden, die zu einer verlängerten oder sogar unwirksamen Therapie führt. Es wird zwischen einer einfachen Resistenz gegenüber einer Klasse, *multidrug*-Resistenz gegenüber mindestens drei Klassen, *extensively drug*-Resistenz gegenüber nahezu allen Klassen und *pandrug*-Resistenz gegenüber allen bekannten Antibiotikaklassen unterschieden [7]. Die Resistenzmechanismen sind gut untersucht und basieren auf sehr unterschiedlichen Mechanismen, wie verminderte Penetration in oder erhöhte Ausschleusung aus der Zelle, enzymatische Inaktivierung oder Modifikation des Antibiotikums, Überproduktion oder Modifizierung der Zielstruktur und Umgehung des gehemmten Stoffwechselwegs [4]. Die Resistenzentwicklung von Bakterien gegenüber Antibiotika stellt eine große Gefahr dar: Die Entwicklung in den letzten 14 Jahren in England zeigt eine deutliche Zunahme von resistenten, immer schwieriger zu therapierenden bakteriellen Infektionen [8, 9]. Der gleiche Prozess wird weltweit beobachtet. Inzwischen wurden klinische Isolate von Bakterien gefunden, die eine *pandrug*-Resistenz besitzen; dies sind die ersten Infektionen, die nicht mehr therapierbar sind [10, 11]. Die Ursachen für die rasante Resistenzentwicklung in den letzten Jahren sind multifaktoriell. Sowohl die zu häufige Anwendung von Antibiotika bei unzureichender Indikation beim Menschen als auch die Anwendung in der Masttierzucht spielen eine wichtige Rolle [12]. Es handelt sich dabei aber nur um beschleunigende Faktoren für die eigentliche Problematik: das Fehlen neuer Antibiotika, die nicht von bekannten Resistenzen bedroht sind. Seit 1987 wurde kein Antibiotikum mehr eingeführt, das einen neuen Wirkmechanismus besaß oder eine neue Substanzklasse darstellte (s. auch die Entwicklungslücke in Abbildung 1.1). Die Entwicklung von Resistenzen ist ein evolutionärer Prozess, der höchstens verlangsamt, aber niemals aufgehalten werden kann [10]. Daher besteht grundsätzlich die Notwendigkeit, neue Antibiotika zu entwickeln. Derzeit befinden sich zwar einige neue Substanzen in der *Pipeline*, jedoch sind nur wenige interessante Vertreter in der Anfangsphase ihrer Entwicklung, während die fortgeschrittenen Substanzen gerade bei den problematischen Bakterien wirkungslos sind [13]. Die Gründe für die fehlende Weiterentwicklung wurden in der Literatur sehr ausführlich diskutiert und bewertet [4, 12, 14–16]:

Zum einen besteht die Problematik in dem Erfolg der Antibiotika selbst, die weitestgehend gut verträglich und, bis auf die Resistenzentwicklung, wirksam sind. Die Anforderungen an neue Antibiotika sind sehr hoch, gleichzeitig sind die so genannten „tief hängenden Früchte“ der Antibiotikaklassen schon entdeckt worden. Neue Klassen erfordern daher einen deutlich höheren Aufwand. Dies wird durch vielfältige Probleme, wie z. B. schnelle Resistenzentwicklung und hohe pharmakokinetische Hürden (bspw. Zellwand der Bakterien), erschwert. Die gute Verträglichkeit vorhandener Antibiotika führen zu hohen regulatorischen Hürden bei der Zulassung neuer Wirkstoffe. Beispielsweise beurteilte SPELLBERG die Anforderungen für klinische Studien von Antibiotika der *Food and Drug Administrati-*

on (FDA) in den USA sehr kritisch: „(...) driven to an irrational and dangerous extreme (...)“ [17]. In der Wissenschaft wurde zu lange Zeit mit den gleichen Methoden ähnliche Leitstrukturen entwickelt. Es ist notwendig, mit komplexeren, zeit- und kostenaufwendigeren Methoden neue chemische Entitäten zu untersuchen. Neben den Anforderungen an die Substanzen selbst sollte auch die Zielstruktur für die Therapie sorgfältig ausgesucht und evaluiert werden. Dieser Kritikpunkt bezieht sich unter anderem auf die Entschlüsselung des Genoms der Bakterien in den 90ern. Die Erwartungen daran, neue Zielstrukturen für die Therapie zu entdecken, konnten nicht erfüllt werden: Es konnte in den letzten Jahrzehnten keine einzige neue Zielstruktur etabliert werden.

Neben allen diesen Gründen ist einer der wichtigsten Einflussfaktoren in der Entwicklung die Wirtschaftlichkeit der Antibiotika. Wie bisher dargelegt, gibt es viele Hürden für die Entwicklung eines neuen Wirkstoffes. Gleichzeitig sind die Gewinnerwartungen für ein Antibiotikum gering. Zum einen werden Antibiotika nicht chronisch gegeben. Eine Infektion und Therapie dauert in der Regel ein bis zwei Wochen. Zum anderen werden neue Antibiotika mit überlegener Wirkung möglichst restriktiv angewendet, um eine Resistenzentwicklung zu verlangsamen. Diese Beobachtungen lassen sich gut belegen: Wie bei Bakterien ist auch bei Viren die Entwicklung von Arzneistoffen aufwendig und schwierig. Bei Viren ist die Mutationsrate und Resistenzentwicklung sogar noch höher. Dennoch wurden zwischen 1993 und 2007 in den USA genauso viele Antibiotika für alle bakteriellen Infektionen kombiniert zugelassen wie HIV-Therapeutika – eine einzige Infektionskrankheit, die meistens im jungen Alter erworben wird und ein Leben lang therapiert werden muss [10]. Bei der Firma Merck generierte das Antidiabetikum mit dem höchsten Umsatz Sitagliptin im Vergleich zum antibiotischen Pendant Ertapenem von 2008 bis 2012 den 7.5fachen Umsatz [18]. Der Anreiz für die Industrie sich mit der Entwicklung neuer Antibiotika zu beschäftigen ist folglich gering, weil es aufwendig und wirtschaftlich nur wenig lukrativ ist. Inzwischen werden verschiedene Lösungsansätze verfolgt, um dieses Dilemma zu lösen: bspw. Patentverlängerungen, Förderung der Zusammenarbeit innerhalb der Industrie und mit Academia, neue Finanzierungsoptionen für Arzneimittel oder sogar öffentliche Finanzierung der Forschung [10, 15, 16, 19].

Ein Fokus bei der Erforschung von neuen Substanzen für die Antibiotikaentwicklung sollte folglich darauf liegen, bei einer gut validierten Zielstruktur neue chemische Entitäten zu entwickeln, die nicht von den Resistenzmechanismen der bekannten Antibiotika betroffen sind.

1.1 Topoisomerasen

Die Länge der Desoxyribonukleinsäure (DNA) im Zellkern einer einzigen eukaryotischen Zelle kann unverdrillt 2 m betragen [20]. Daher muss die DNA zur Aufbewahrung sehr kompakt gelagert werden. Andererseits muss sie zur Erfüllung ihrer Funktion abgelesen werden können. Für diese aufwendige Regulation der Topologie der DNA sind unter anderem DNA Topoisomerasen (im Folgenden Topoisomerasen genannt) verantwortlich. Alle bekannten Zellen enthalten Topoisomerasen, da sie beispielsweise essentiell für die Transkription und Replikation sind [21]. Topoisomerasen werden in zwei Hauptgruppen kategorisiert. Topoisomerasen Typ I verändern die Topologie durch einen Einzelstrangbruch, bei Topoisomerasen Typ II wird dies durch einen Doppelstrangbruch erreicht. Für beide Typen gibt es noch verschiedene Subtypen von Topoisomerasen, die sich nochmals erheblich in der Sequenz und dem Reaktionsmechanismus unterscheiden können, auf die in dieser Arbeit nicht alle im Einzelnen eingegangen werden soll [22, 23]. Typ II Topoisomerasen werden in zwei Subtypen eingeteilt Typ IIa stellt dabei die größere Gruppe dar und enthält unter anderem die DNA Gyrase (im Folgenden Gyrase genannt) sowie die DNA Topoisomerase IV (Topo IV).

Topoisomerasen haben viele Funktionen, die, wie oben beschrieben, für das Überleben der Zelle notwendig sind. Sie können Knoten in der DNA entfernen oder verknüpfte ringförmige Plasmide voneinander trennen (so genannte *Decatenation*). Weiterhin sind sie in der Lage sowohl negative wie positive Supercoils in die DNA einzuführen als auch supercoiled oder unter Spannung stehende DNA wieder zu relaxieren (s. Abbildung 1.2). Diese teilweise gegensätzlichen Funktionen können sich noch innerhalb der Subtypen der verschiedenen Topoisomerasen unterscheiden, z. B. führt die Gyrase negative Supercoils ein und die Topo IV relaxiert die DNA wieder – beide gehören zu den Topoisomerasen Typ IIa (s. Abbildung 1.2). Verschiedene Topoisomerasen zeigen zusätzlich noch eine Präferenz für bestimmte Reaktionstypen. Bspw. weist die Topo IV besonders hohe Aktivität bei der *Decatenation* auf und ist für eine schnelle Zellteilung wichtig [23]. Damit stellt die Regulation der DNA durch Topoisomerasen ein kompliziertes Gleichgewicht her, das auch für den Zellzyklus wichtig ist.

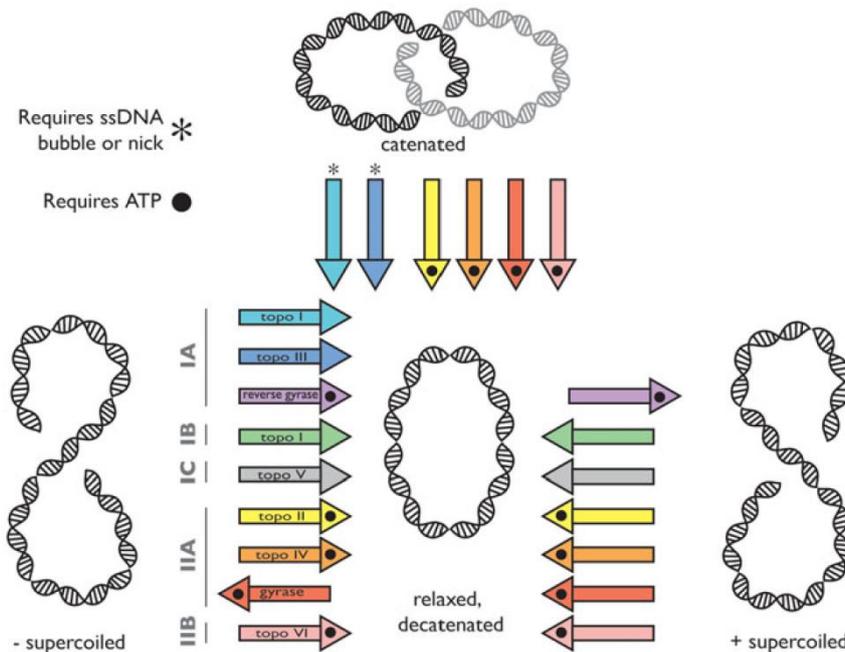


Abbildung 1.2 – Aufgaben der Topoisomerase. Verschiedene Topoisomerasen und ihre Funktion sind entsprechend des Farbcodes (Mitte links) dargestellt. ATPabhängige Enzyme sind mit einem Punkt markiert. Bei einem Sternchen können die Topoisomerasen die entsprechende Reaktion nur bei einer Einzelstrang-DNA katalysieren. Abbildung mit Erlaubnis von SCHOEFFLER *et al.* aus [23].

– Gyrase

Bei der Gyrase handelt es sich, wie oben beschrieben, um eine Topoisomerase Typ IIa. Die Gyrase kommt in allen Prokaryoten und bei einigen Eukaryoten vor. In menschlichen Zellen gibt es keine Gyrase. Unter Hydrolyse von Adenosintriphosphat (ATP) führt sie negative Supercoils in die DNA ein, relaxiert positive Supercoils, kann *Decatenation* durchführen und Knoten entfernen. Bei hoher Spannung der DNA kann sie auch ohne ATP-Hydrolyse negative Supercoils relaxieren [24, 25]. *In vivo* steht vor allem die Einführung von negativen Supercoils im Vordergrund, da die Gyrase die einzig bekannte Topoisomerase dafür ist. Ohne diese Funktion ist ein Bakterium nicht überlebensfähig.

Für die Funktionsweise der Gyrase ist das Verständnis der Struktur essentiell (s. Abbildung 1.3). Die prokaryotischen Gyrase ist ein Tetramer, das aus zwei Homodimeren besteht: zwei Gyrase Untereinheiten A (GyrA) und zwei Gyrase Untereinheiten B (GyrB). In Abbildung 1.3 ist der Aufbau und die Struktur der Gyrase dargestellt. Während der Katalyse wird ein DNA-Segment durch den Doppelstrangbruch eines zweiten DNA-Segments geführt. Das erstere DNA-Segment wird T-Segment, *Transported Segment* (engl. für „transportiertes Segment“), das gespaltene DNA-Segment G-Segment, *Gate-Segment* (engl. für „Tor-Segment“), genannt. Während der Reaktion durchschreitet das T-Segment drei Tore (engl. *Gate*).

Das erste Tor, das als *N-Gate* (N-terminale Ende der GyrB) bezeichnet wird, wird von der GHKL-Domäne und dem *Transducer* von zwei GyrB-Untereinheiten gebildet. Bei der GHKL-Domäne handelt

es sich um eine konservierte ATPase, die bei funktional sehr unterschiedlichen Enzymen auftritt: Gyrase, Heat Shock Protein 90, Histidinkinasen, MutL (ein DNA-Mismatch-Reparaturprotein) [26]. An der GHKL schließt sich die übertragende Domäne (*Transducer*) an (in Abbildung 1.3 beide orange).

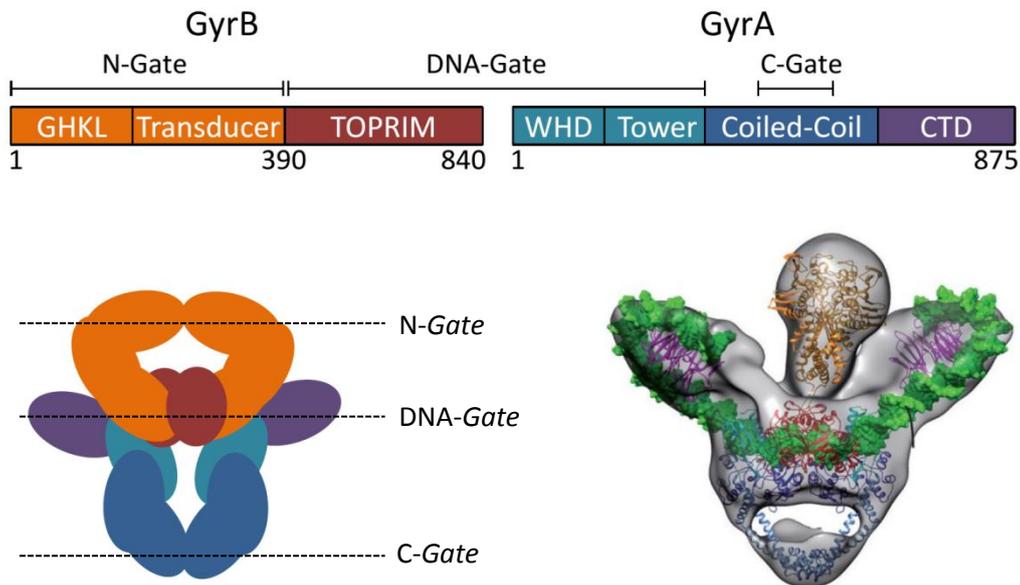


Abbildung 1.3 – Struktur der Gyrase. **Oben:** Einteilung der Domänen des Enzyms. *N-Gate*: N-terminales Tor, *DNA-Gate*: Tor durch die DNA, *C-Gate*: C-terminales Tor; für einzelne Domänen s. Text. **Unten links:** Schematische Darstellung der Struktur der Gyrase. Die Domänen sind in den Farben des oberen Schemas dargestellt. (nach [27]). **Unten rechts:** Modell für die Gesamtstruktur der Gyrase. PAPILLON *et al.* konnten mittels Kryo-Elektronenmikroskopie die grau dargestellte Oberfläche bestimmen und modellierten nachträglich die Abschnitte mit verschiedenen Röntgenkristallstrukturen in die Oberfläche hinein (G-Segment ist grün dargestellt, weitere Farben identisch zu den oben gezeigten Domänen) [28].

Die TOPRIM-Domäne, eine konservierte Domäne der Topoisomerasen und Primasen, folgt dem Transducer (in Abbildung 1.3 rot). Über diese Domäne interagiert die GyrB-Untereinheit während der Katalyse mit dem G-Segment und stabilisiert den Doppelstrangbruch der DNA. Weiterhin findet über die Domäne auch die Dimerisierung mit dem N-terminalen Ende der GyrA-Untereinheit statt.

Das N-terminale Ende der GyrA interagiert auch mit dem G-Segment. Während der Spaltung ist das G-Segment kovalent an einem konservierten Tyrosin in der WHD-Domäne (*Winged Helix Domäne*) gebunden (in Abbildung 1.3 türkis). Zusammen mit dem G-Segment formen die TOPRIM-, die WHD- und die sich anschließende *Tower*-Domäne (engl. für Turm) das so genannte *DNA-Gate*. Schließlich bildet die *Coiled-Coil*-Domäne (engl. für „Doppelwendel“), dies sind sich gegenseitig umwindende α -Helices, das *C-Gate* (in Abbildung 1.3 blau). Über das *C-Gate* wird das T-Segment nach der Katalyse wieder freigegeben. Bei der CTD (C-terminalen Domäne) handelt es sich um β -Faltblätter, die sich windradartig anordnen (β -*Pinwheel*; in Abbildung 1.3 magenta). Um dieses schlingt sich während der Katalyse das G-Segment (in Abbildung 1.3 unten rechts grün).

Der vollständige Reaktionsmechanismus der Gyrase und der Topoisomerasen Typ II ist noch immer nicht aufgeklärt. In den letzten Jahren konnten jedoch durch verschiedene Methoden wie Röntgen-

kristallographie [29, 30], Förster-Resonanzenergietransfer (FRET) [27, 31, 32], Kryoelektronmikroskopie [28] oder Röntgenkleinwinkelstreuung (SAXS) [33] immer genauere Einblicke in den Mechanismus gewonnen werden.

In Abbildung 1.4 ist die Katalyse schematisch dargestellt. Zunächst bindet das G-Segment an die Gyrase (1, s. Abbildung 1.4). Das G-Segment windet sich um die CTD (in der Abbildung nicht dargestellt) und das T-Segment wird über das G-Segment geführt (2). ATP bindet und die GyrB-Untereinheiten dimerisieren am N-terminalen Ende; das N-Gate schließt sich. Gleichzeitig öffnet sich das DNA-Gate, dabei wird auch das G-Segment gespalten. Anschließend wird das T-Segment durch das G-Gate geführt (3). Das DNA-Gate schließt sich, möglicherweise unter ATP-Hydrolyse, und das C-Gate wird geöffnet (4). Nach Freisetzung des ADPs unter Öffnung des N-Gates und Schließen des C-Gates befindet sich der DNA-Enzym-

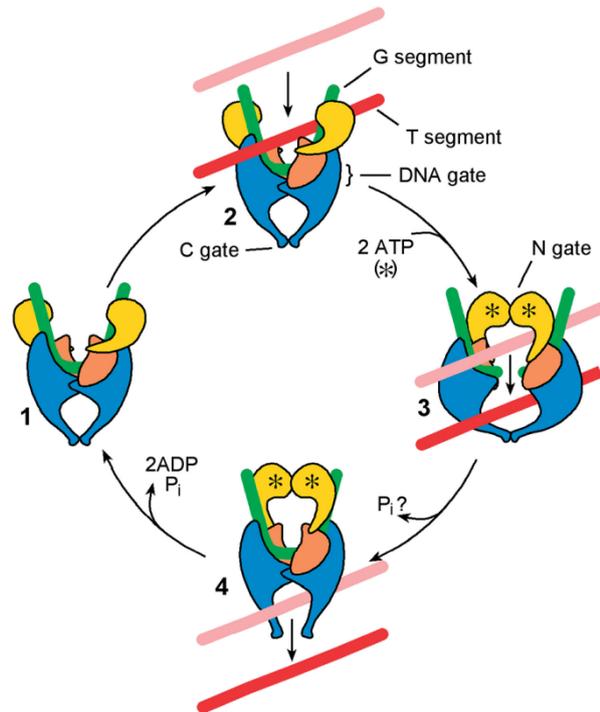


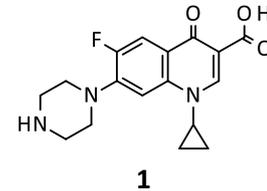
Abbildung 1.4 – Schematische Darstellung des Mechanismus der Katalyse. Die GHKL- und Transducer-Domäne sind gelb, die TOPRIM-Domäne orange und die GyrA-Untereinheit blau dargestellt. Erklärung s. Text. Abbildung mit Erlaubnis von BATES *et al.* aus [34].

Komplex wieder im Ausgangszustand (1). Das G-Segment kann freigesetzt werden oder der Zyklus läuft erneut ab [27]. Essentiell für diese Reaktion ist unter anderem die Bindung von ATP in Schritt 3, die zum Schließen des N-Gates und zur Öffnung des G-Gates führt [34, 35]. Wird die Bindung von ATP gehemmt wird der ganze Mechanismus unterbrochen und die Topologie der DNA kann nicht mehr durch die Gyrase reguliert werden. Dies führt letztendlich zum Tod der Zelle [36]. Bisher konnte aber noch nicht abschließend geklärt werden, warum Topoisomerasen Typ II überhaupt die Energiezufuhr in Form von ATP benötigen. Bis auf die Einführung von negativen Supercoils in die DNA durch die Gyrase sind alle anderen katalysierten Reaktionen energetisch günstig und benötigen keine ATP-Hydrolyse. Die meisten Topoisomerasen Typ I benötigen für ähnliche Reaktion kein ATP (s. Abbildung 1.2 auf S. 5). Mögliche Erklärungen lieferten BATES *et al.* [34], die den Hauptgrund in der Kontrolle der Protein-Protein-Interaktion und damit der Vermeidung von unkontrollierten Doppelstrangbrüchen sehen.

Die Gyrase besitzt also die einzigartige und für Bakterien essentielle Eigenschaft negative Supercoils in die DNA einzuführen. Gleichzeitig gibt es in menschlichen Zellen kein vergleichbares Enzym, das diese Funktion übernimmt oder strukturell mit der Gyrase verwandt ist [36]. Damit stellt die Gyrase

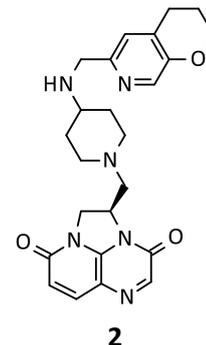
im Allgemeinen und die ATP-Bindestelle im Speziellen eine wichtige Zielstruktur für Antibiotika dar. Im Folgenden werden kurz die nicht ATP-kompetitiven Gyrase-Hemmstoffe vorgestellt, die eine klinische Relevanz besitzen. Die ATP-kompetitiven Inhibitoren werden ausführlich in Abschnitt 1.3 auf S. 14 vorgestellt.

Die bekanntesten Gyraseinhibitoren der GyrA-Untereinheit sind die Fluorchinolone (s. Abbildung 1.5). Sie stabilisieren den DNA-Gyrase-Komplex. Sie binden in der Region des Doppelstrangbruchs des G-Segments und führen zur Einfrierung des Komplexes. Letztendlich wird die Aktivität der Gyrase gehemmt, die DNA segmentiert und das Bakterium stirbt [37]. Aufgrund des großen Erfolgs von Fluorchinolonen wurden Strukturvariationen schon sehr intensiv erforscht – sowohl um Resistenzmechanismen zu umgehen als auch um das Spektrum zu erweitern. Möglicherweise sind für diese Antibiotikum-Gruppe die Grenzen der Optimierung erreicht [36]. Insbesondere die Resistenzentwicklung nimmt, wie bei allen Antibiotika-Klassen (s. auch Resistenzentwicklung auf S. 2), dramatisch zu [38].



1
Abbildung 1.5 – Ciprofloxacin. Ein Vertreter der Fluorchinolone.

2010 konnten die Mitarbeiter von GlaxoSmithKline eine neue Bindungsstelle für Gyrasinhibitoren entdecken, die sich in der Nähe der Bindungsstelle der Fluorchinolone befindet, aber nicht mit dieser überschneidet [39]. Der Wirkungsmechanismus unterscheidet sich auch deutlich zu den Fluorchinolonen, auch wenn die Bindung ebenfalls an dem DNA-Gyrase-Komplex stattfindet. Vermutlich führt diese neue Klasse von Gyraseinhibitoren zu Einzelstrangbrüchen in der DNA. Diese Gruppe wurde *Novel Bacterial Topoisomerase Inhibitors* (NBTI) genannt. Sie stellt ein gutes Beispiel dafür dar, dass noch längst nicht alle



2
Abbildung 1.6 – GSK2140944. Klinischer Kandidat eines NBTIs von GlaxoSmithKline.

Wirkmechanismen für die Inhibition der Gyrase bekannt sind und betont die Wichtigkeit der mechanistischen Aufklärung der katalytischen Reaktion. Diese Entdeckung ermöglichte die Entwicklung zahlreicher weiterer potentieller Kandidaten für Antibiotika [40–42]. Mit GSK2140944 gibt es von GlaxoSmithKline seit 2015 auch einen Kandidaten in der klinischen Prüfung [43].

Diese zwei Klassen von Gyrasinhibitoren stellen nur einen kleinen Auszug von den bekannten Wirkmechanismen dar. Neben diesen Klassen gibt es noch zahlreiche niedermolekulare (Natur-)Stoffe mit ähnlichen oder anderen Wirkmechanismen und auch bisher unerwähnte hochmolekulare Inhibitoren (z. B. Proteine), die die Gyrase inhibieren [36]. Allerdings wurde aus keinem dieser Strukturen bisher ein klinischer Kandidat oder Arzneistoff entwickelt.

1.2 Strukturbasierte Leitstruktursuche

– Leitstruktur

Im Rahmen dieser Arbeit soll eine strukturbasierte Leitstruktursuche durchgeführt werden. Der Begriff der „Leitstruktur“ (engl. *lead* oder *lead compound*) wird in der Literatur häufig verwendet, ohne dass dieser definiert wird. KLEBE beschreibt eine Leitstruktur als eine Substanz, die bereits eine biologische Wirkung besitzt, aber noch kein Wirkstoff ist, da wichtige Eigenschaften fehlen, z. B. Selektivität oder Wirkstärke. Essentiell hingegen ist für die Leitstruktur, dass sie noch weiter chemisch variiert und optimiert werden kann [44]. Dies sind deutlich weniger Anforderungen an eine Leitstruktur, als nach der Meinung von OPREA *et al.* von AstraZeneca erfüllt werden müssen: einfache chemische Struktur, die für kombinatorische oder medizinisch chemische Optimierungen gut zugänglich ist, Mitglied in einer gut untersuchten und nachvollziehbaren Struktur-Wirkungsbeziehungen (SAR), günstige Patentlage und gute ADME-Eigenschaftenⁱ [45]. Der Begriff der Leitstruktur kann also abhängig von der Perspektive und der Zielsetzung unterschiedlich definiert werden. Eine allgemeingültige Formulierung für eine Leitstruktur lautet nach VALLER *et al.* folgendermaßen [46]:

„A representative of a compound series with sufficient potential (as measured by potency, selectivity, pharmacokinetics, physicochemical properties, absence of toxicity and novelty) to progress to a full drug development programme.“

Eine Leitstruktur ist folglich ein Repräsentant einer Substanzserie mit ausreichend Potential, um für die Entwicklung eines Arzneistoffs verwendet zu werden. Dabei kann das Potential je nach Perspektive unterschiedliche Anforderungen beinhalten: Synthetische Variationen und SAR nehmen dabei eine Schlüsselrolle ein. Pharmakokinetik, Selektivität, Toxizität, physikochemische Parameter oder Neuheit sind Beispiele für Eigenschaften, die fallbezogen bedeutend sein können. Im Rahmen dieser Arbeit wird Folgendes von einer Leitstruktur erwartet:

1. Bei der Leitstruktur soll es sich um eine neue chemische Entität handeln, die noch nicht als Inhibitor der ATPase der Gyrase beschrieben wurde.
2. Die Leitstruktur soll nachvollziehbare SAR-Eigenschaften besitzen.
3. Die Leitstruktur soll synthetisch gut erschließbar und modifizierbar sein.
4. Die Leitstruktur soll zu einem oral verfügbaren Wirkstoff weiterentwickelt werden können.

ⁱ ADME-Eigenschaften beschreiben die Pharmakokinetik eines Wirkstoffes *in vivo*: Absorption, Distribution, Metabolisierung, Exkretion.

– Leitstruktursuche

Die Suche nach einer Leitstruktur kann auf verschiedenen Wegen erfolgen. In der Vergangenheit wurden Leitstrukturen durch Erfahrungen, Beobachtungen oder auch Zufallsentdeckungen entwickelt. Dabei spielten vor allem Naturstoffe eine große Rolle. Inhaltsstoffe von Pflanzen (z. B. Morphin, Papaverin, Cocain), von Mikroorganismen produzierte Substanzen (z. B. Penicillin, Tetracyclin, Ergotamin) oder auch Substanzen tierischen Ursprungs (z. B. Teprotid (Schlangengift und Leitstruktur der ACE-Hemmer), Hirudin) waren und sind wichtige Quellen für Leitstrukturen. Weitere Ausgangsstrukturen stellen endogene Substanzen (z. B. Adrenalin, Insulin) dar, die im menschlichen Körper eine wichtige physiologische Rolle besitzen. Dabei gab es sowohl sehr systematische Ansätze, wie beim Teprotid, als auch reine Zufallsbeobachtungen, wie beim Penicillin, die zur Entdeckung von neuen Leitstrukturen und Arzneistoffen führten [44]. In den 90er Jahren des letzten Jahrtausends etablierten sich durch zunehmende Automatisierbarkeit und Miniaturisierung biologischer Assays *High-Throughput-Screenings* (HTS). Bei diesen werden in automatisierten Prozessen bis zu mehrere 100,000 Substanzen gegen eine Zielstruktur getestet (so genanntes Screening) [47, 48]. Kann die beim Screening gefundene Aktivität validiert werden, ist eine Hit-Struktur (oder einfach Hit) gefunden. Ein Hit besitzt damit zunächst nur die nachgewiesene Aktivität gegenüber der getesteten Zielstruktur. In weiteren Schritten, z. B. chemische Variationen des Hits, Anpassung der physikochemischen Eigenschaften, kann daraus eine Leitstruktur mit den gewünschten Eigenschaften (s. o.) generiert werden [44, 46, 49]. Die hohen Erwartungen durch die systematische und gründliche Suche, neue und hochpotente Leitstrukturen zu finden, konnten jedoch nicht erfüllt werden: Trotz intensiver Bemühungen konnte die Anzahl an neuen chemischen Entitäten bei zugelassenen Arzneimitteln zwischen 1990 und 2000 nicht erhöht werden [49]. Einen weiteren Meilenstein stellte die Strukturklärung der Zielstrukturen durch Röntgenkristallstrukturen oder auch Kernspinresonanz (NMR) dar [50]. Dies ermöglichte die strukturbasierte Leitstruktursuche.

– Strukturbasierte Leitstruktursuche

Bei der strukturbasierten Leitstruktursuche werden die strukturellen Informationen der Zielstruktur (engl. *Target*) verwendet, um eine Leitstruktur zu entwickeln. Die zusätzlichen Informationen und Erkenntnisse aus den Strukturinformationen führen dazu, dass, anders als bei klassischen HTS-Ansätzen, nicht zufallsbasiert nach neuen Hits gesucht wird, sondern durch computerbasierte Methoden [51, 52] gerichtet Substanzen zur biologischen Testung ausgewählt und optimiert werden. Die Durchführung wurde in der Literatur bereits ausführlich diskutiert [50, 53–57]. Im Folgenden soll der darauf basierende Ablauf, der in dieser Arbeit verwendet wurde, dargestellt werden.

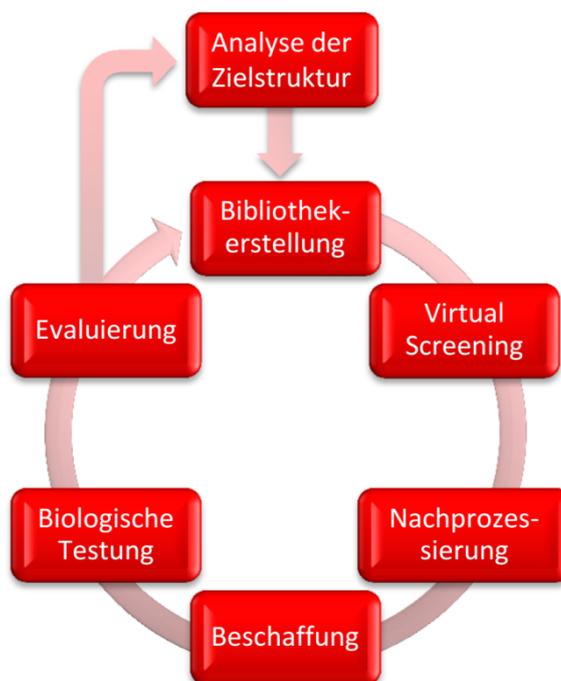


Abbildung 1.7 – Der zyklische Ablauf einer strukturbasierten Leitstruktursuche. Zunächst muss eine validierte Zielstruktur vorliegen und deren strukturellen Eigenschaften analysiert werden. Der Zielsetzung entsprechend wird eine Substanzbibliothek erzeugt und ein virtuelles Screening (*Virtual Screening*) durchgeführt, um potentielle Hits anzureichern. In weiteren Nachprozessierungsschritten werden die gescreenten Substanzen weiter gefiltert und priorisiert. Die schließlich beschafften Substanzen können abschließend getestet werden. Mithilfe der gewonnenen experimentellen Ergebnisse kann der gesamte Prozess evaluiert und ein neuer Zyklus gestartet werden.

In Abbildung 1.7 ist der schematische Zyklus der strukturbasierten Leitstruktursuche dargestellt. Zu Beginn dieser muss eine genaue Analyse der Zielstruktur stattfinden. Dies beinhaltet auch die Validierung der Zielstruktur. Nach einer erfolgreichen Validierung kann mit der Analyse der dreidimensionalen Proteinstrukturen begonnen werden. In der „*Protein Data Bank*“ [58] befinden sich öffentlich zugängliche Röntgenkristallstrukturen. In dieser sind z. B. auch Röntgenkristallstrukturen der ATPase der Gyrase enthalten. Diese Strukturen können allerdings nicht ohne weitere Bearbeitung weiterverwendet werden. Zum einen sind Wasserstoffatome zu klein, als dass sie regelmäßig in der Röntgenkristallographie bestimmt werden können. Daher müssen diese nachträglich hinzugefügt werden. Dies ist für die meisten Aminosäuren und Liganden unproblematisch. Aufwendiger wird dies, wenn verschiedene Tautomere möglich sind (z. B. bei Histidin). Zusätzlich können saure oder basische Funktionen (z. B. bei Asparaginsäure oder Lysin) in Abhängigkeit ihrer Umgebung geladen oder ungeladen vorliegen. Des Weiteren ist die Elektronendichte, die bei der Röntgenkristallographie bestimmt wird, von Hydroxyl- und Amino-Gruppen sehr ähnlich. Daher muss auch die Orientierung der Amide (bspw. bei Asparagin oder Glutamin) untersucht werden. Die gleichen Entscheidungen müssen auch jeweils bei dem kokristallisierten Liganden getroffen werden. Schließlich können auch noch Atome, Seitenketten oder ganze Aminosäuren experimentell nicht bestimmt worden sein. Diese müssen, wenn sie einen Einfluss auf die Bindung von Liganden besitzen, ebenfalls modelliert werden. Erst mit der komplettierten Struktur kann in der Folge eine genaue Analyse der Bindetasche erfolgen.

Bei dieser werden die Interaktionen von einem oder mehreren Liganden mit dem Protein untersucht. Dadurch können Schlüsselinteraktionen, die ein neuentwickelter Ligand besetzen soll, festgelegt werden oder auch Bereiche in der Bindetasche bestimmt werden, die zusätzlich oder nicht mehr belegt werden sollen. Ebenfalls von Bedeutung sind Wassermoleküle, die an der Bindung von Liganden beteiligt sind. Aufgrund gegensätzlicher energetischer Einflüsse in der freien Bindungsenthalpie kann die Verdrängung oder die Konservierung von Wasser vorteilhaft sein (s. auch Abschnitt 2.3 auf S. 47). Obwohl viele dieser Arbeitsschritte zum standardmäßigen Vorgehen bei der Proteinvorbereitung gehören, sind sie und das Verständnis der Protein-Ligand-Interaktion von elementarer Bedeutung für den gesamten Ablauf.

Im nächsten Schritt wird die Substanzbibliothek erzeugt („Erzeugung der Bibliothek“; s. Abbildung 1.7). Die Zusammensetzung dieser Bibliothek ist vom gesamten Prozess abhängig. Für die Durchführung muss die Bibliothek von allen verwendeten Computermethoden prozessierbar sein, z. B. muss die Anzahl der Substanzen bearbeitbar sein. Selbstverständlich sollten auch nur Substanzen gescreent werden, die für eine Testung auch beschafft (s. u.) und untersucht werden können, z. B. besitzen Säurehalogenide eine zu hohe Reaktivität, um in einem enzymatischen Assay untersucht werden zu können. Von elementarer Bedeutung sind aber auch die Anforderungen an die Leitstruktur. Während bei einem Hit zunächst primär die biologische Aktivität im Vordergrund steht, soll aus diesem schließlich die Leitstruktur entwickelt werden (s. Leitstruktursuche auf S. 10). Dieser Prozess kann vereinfacht werden, indem schon bei der Erstellung der Substanzbibliothek die Anforderung an die Leitstruktur berücksichtigt und Substanzen mit für die Optimierung ungünstigen Eigenschaften entfernt werden.

Beim virtuellen Screening (s. Abbildung 1.7 Virtual Screening) wird schließlich der Screening-Prozess, welcher beim HTS *in vitro* stattfindet, *in silicio* durchgeführt. Bei diesem Vorgang wird versucht, vorherzusagen, ob und wie gut eine Substanz in der Zielstruktur bindet. Dies wird im Allgemeinen mithilfe eines Dockings (s. auch Abschnitt 4 auf S. 79) durchgeführt. Dabei können auch sehr große Substanzbibliotheken mit mehreren Millionen Substanzen gescreent werden. Dieses Vorgehen wird dann in Analogie zum HTS als HTvS (*High-Throughput virtual Screening*) bezeichnet. In diesem Schritt werden also die Informationen aus der Analyse der Zielstruktur mit der Substanzbibliothek vereinigt.

Bei der Nachprozessierung (engl. *Postprocessing*; s. Abbildung 1.7) werden die aus dem virtuellen Screening gewonnenen Ergebnisse weiter evaluiert. Dabei handelt es sich nicht um klar definierbare Schritte, vielmehr können nahezu alle Methoden des *Molecular Modellings* verwendet werden, die eine Priorisierung von Protein-Ligand-Komplexen erlauben (z. B. Scoring-Funktionen, Moleküldynamik-Simulationen). Es ist sogar möglich, ein erneutes Docking, wie beim Schritt zuvor, durchzuführen (s. o.). Im Allgemeinen besteht das sukzessive Vorgehen darin, die zuvor sortierten Substanzen zu filtern und mit aufwendigeren sowie genaueren Methoden neu zu bewerten und zu sortieren, um die

Substanzen erneut zu filtern usw. Die Anzahl der zu durchlaufenden Postprocessing-Zyklen, hängt sehr stark vom Gesamtprozess ab. Die Güte des virtuellen Screening, die Größe der Substanzbibliothek und auch die Anzahl der zu testenden Substanzen sind die maßgeblich beeinflussenden Parameter.

Steht die Auswahl der Substanzen fest, müssen diese im nächsten Schritt noch beschafft werden (s. Abbildung 1.7). Dies kann bspw. durch Synthese oder den Erwerb der Substanzen geschehen. Der geplante Beschaffungsweg ist von hoher Bedeutung für den Gesamtprozess und muss schon bei der Erstellung der Substanzbibliothek berücksichtigt werden.

Durch die biologische Testung (s. Abbildung 1.7) wird schließlich ermittelt, ob es sich bei den ausgewählten Substanzen um Hits handelt. Dabei werden in der Regel zunächst alle erworbenen Substanzen bei einem Konzentrationsniveau getestet. Zeigt eine Substanz in dem ersten Assay ihre vorhergesagte Aktivität, ist die Substanz jedoch nicht automatisch ein validierter Hit. Abhängig von der Assay-Methode kann die Testsubstanz den Assay beeinflussen, z. B. durch Eigenfluoreszenz oder Inhibition weiterer beteiligter Enzyme (s. auch Abschnitt 6 auf S. 151). Bei HTS-Kampagnen ist die hohe falsch-positiv Rate bei dem initialen Assay schon länger bekannt [59]. Über diese wurde auch bereits beim virtuellen Screening berichtet [60]. Daher sollte jeder initiale Hit zumindest durch einen orthogonalen Assay validiert werden. Orthogonal bedeutet in diesem Zusammenhang, dass der Assay auf einem anderen Messprinzip basiert und daher möglichst nicht durch die gleiche Eigenschaft der falsch-positiven Substanz beeinflusst wird. Bspw. kann bei einem Assay für einen Ionenkanal eine fluoreszenzbasierte Methode und zur Validierung eine auf dem elektrochemischen Membranpotential basierende Methode orthogonal dazu verwendet werden.

Zum Abschluss des Zyklus findet die Evaluierung (s. Abbildung 1.7) statt. Bei dieser sollten einerseits die kritischen Schritte der Prozess selbst und andererseits die Ergebnisse, also die gefundenen Hits, bewertet werden. Für das Verständnis des Prozesses ist es wichtig, herauszufiltern, ob die getroffenen Annahmen während des Durchlaufs korrekt waren. Das können Entscheidungen von Protonierungszuständen im Protein, Bindungshypothesen, Eigenschaften der Substanzbibliothek, die Wahl des Screening Tools usw. sein. Diese Analyse ist wichtig, um herauszufinden, wo die „Fallstricke“ im Prozess liegen. Da der Prozess zyklisch abläuft, werden diese Informationen benötigt, um den erneuten Durchlauf zu verbessern. Die Hit-Strukturen stellen bei diesem Zyklus die Ausgangsergebnisse für den erneuten Durchlauf dar. Daher sollte eine genaue Analyse bspw. den möglichen Bindungsmodus und die Erweiterungsmöglichkeiten untersuchen.

Nach der Evaluierung findet ein erneuter Durchgang des Zyklus statt. Dabei wird prinzipiell der gleiche schematische Prozess durchlaufen, die Methoden können aber sehr verschieden sein: z. B. wird keine große Substanzbibliothek erstellt, sondern es findet eine Erstellung von Bibliotheken analoger

(oder strukturell ähnlicher) Substanzen um die Hitstrukturen statt. Es werden also die Ergebnisse der Evaluierung benötigt, um den Gesamtprozess fortzuführen. Dieser Kreisprozess kann beliebig häufig iterativ durchlaufen werden, bis eine Leitstruktur mit den gewünschten Eigenschaften gefunden wird.

1.3 Kompetitive Inhibitoren der ATPase der GyrB-Untereinheit

– Naturstoffe

Zu den ersten erforschten Inhibitoren der ATPase-Untereinheit der Gyrase gehörten die Naturstoffe (s. Abbildung 1.8). Novobiocin (**3**) ist ein antibakteriell wirkendes Amincoumarin, welches von bestimmten *Streptomyces*-Arten produziert wird.

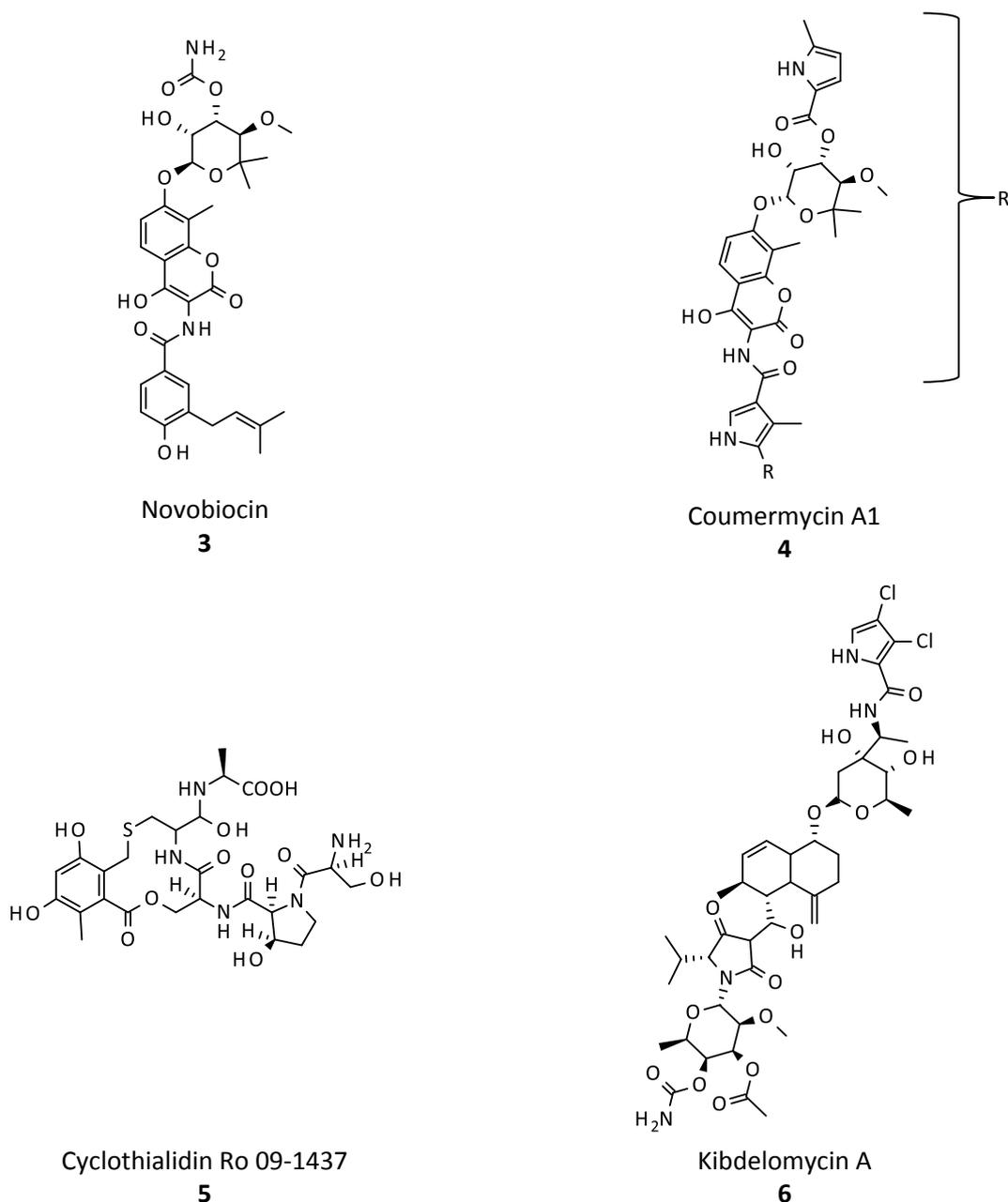


Abbildung 1.8 – Vertreter bekannter Naturstoffe von kompetitiven Inhibitoren der ATPase der GyrB-Untereinheit.

Obwohl Novobiocin (**3**) schon in den 1950ern als Antibiotikum klinisch eingesetzt wurde, konnte erst 1996 mithilfe von Röntgenkristallstrukturanalyse der Wirkungsmechanismus, die kompetitive Inhibition der ATP-Bindungsstelle, abschließend geklärt werden [61]. Aufgrund seiner schlechten Aktivität gegenüber gram-negativen Bakterien, Zytotoxizität gegenüber menschlichen Zellen und schlechter pharmakokinetischer Eigenschaften war Novobiocin (**3**) in der klinischen Anwendung jedoch nicht erfolgreich [36] und ist in Deutschland nicht zugelassen. Dennoch stellt Novobiocin (**3**) damit den *Proof of Concept* dafür dar, dass ein ATP-kompetitiver Mechanismus der Gyrasehemmung klinisch angewendet werden kann. Zur Optimierung der pharmakologischen und pharmkokinetischen Eigenschaften von Novobiocin (**3**) werden auch weiterhin neue Derivate dieses Aminocumarins erzeugt und getestet, bspw. durch Mutasynthese [62–64] oder chemische Modifikationen [65].

Zusätzlich zu den Aminocumarinen gibt es weitere Naturstoffen, die die ATPase ebenfalls kompetitiv hemmen, aber noch keine klinische Anwendung gefunden haben. Neben strukturell verwandten Substanzen, z. B. Coumermycine (**4**), gibt es auch kleine Peptide wie Cyclothialidine (**5**). Cyclothialidine (**5**) zeigen eine gute, selektive Hemmung *in vitro*, weisen aber nahezu keine Wirkung bei lebenden Bakterien, also *in vivo*, auf. Weiterhin sind sie aufgrund ihrer strukturellen Eigenschaften pharmakokinetisch sehr ungünstig: Sie würden im menschlichen Körper schlecht resorbiert und schnell metabolisiert werden [36, 66]. 2011 entdeckte die Firma Merck einen neuen Naturstoff, Kibdelomycin (**6**), und konnte die Bindung in der ATPase-Tasche der Gyrase mit einer Röntgenkristallstruktur nachweisen. Diese Struktur ist ebenfalls ein kompetitiver Inhibitor von ATP, weist aber einen einzigartigen und bisher nicht bekannten Bindungsmodus auf [67, 68].

Zusammenfassend stellen die Naturstoffe die Grundlage für die Entwicklung von Inhibitoren der ATPase der GyrB-Untereinheit dar. Zum einen konnten mit ihrer Hilfe der Wirkmechanismus und damit die potentielle Zielstruktur für neue Wirkstoffe bestimmt werden. Zum anderen validierte der Einsatz von Novobiocin (**3**) in der Therapie die Anwendbarkeit dieses Wirkprinzips. Weiterhin führte die strukturelle Aufklärung des Bindungsortes, insbesondere durch Röntgenkristallstrukturen der verschiedenen Protein-Naturstoff-Komplexe, zur Entwicklung zahlreicher neuer synthetischer Substanzen (s. u.). Dabei ist davon auszugehen, dass noch längst nicht alle Naturstoffe entdeckt worden sind, die diesen oder einen ähnlichen Wirkmechanismus besitzen. Kibdelomycin (**6**) ist ein gutes Beispiel dafür, wie mit neuen Naturstoffen weitere Erkenntnisse über die Zielstruktur gewonnen werden können.

– Synthetische Inhibitoren

Die systematische Entwicklung von synthetischen Inhibitoren der ATPase-Untereinheit der Gyrase wurde vor allem durch die Röntgenkristallstrukturen von Naturstoffen (bspw. Novobiocin-GyrB-Komplex von 1997 [69]) und physiologischen ATP-Derivaten (ANP-GyrB-Komplex von 2000 [35]) ermöglicht. Im Folgenden wird ein kurzer Überblick über die verschiedenen gefundenen Substanzklassen gegeben. Dabei liegt der Fokus auf der verwendeten Methodik und den Strukturklassen. Für eine detailliertere Darstellung wird auf die zwei Reviews von TOMAŠIČ *et al.* [70] und MAYER *et al.* [66] verwiesen.ⁱⁱ

Allgemein lassen sich die bisherigen systematischen Entwicklungen in fragmentbasierte und HT(v)S-Ansätze einteilen. Bei den fragmentbasierten Ansätzen werden kleine Fragmente gesucht, die zu meist nur eine schwache und unselektive Bindung in der Zielstruktur aufweisen. Im weiteren Verlauf werden die Fragmente Schrittweise zu einer Leitstruktur (s. auch Leitstruktur auf S. 9) aufgebaut [71].

ⁱⁱDiese Arbeit wurde im Frühjahr 2011 begonnen. Um den damaligen Kenntnisstand zu verdeutlichen, sind für die verschiedenen Entwicklungen die jeweiligen Jahreszahlen angegeben. Erkenntnisse aus neueren Veröffentlichungen wurden im Verlauf der Arbeit, soweit wie möglich, berücksichtigt.

Der Vorteil bei diesem Verfahren besteht darin, dass die Substanz noch weitreichend an die Bindeta-sche angepasst werden kann. Während der Optimierung kann dabei noch Rücksicht auf die physiko-chemischen Eigenschaften der Substanz genommen werden. Die Substanz kann insgesamt flexibler optimiert werden, als dies mit Hits bei HTS möglich ist. Die Nachteile liegen in einem deutlich größe-ren Aufwand. Die Fragmente müssen zunächst eine gute Löslichkeit im Assay aufweisen, da sie in höheren Konzentrationen getestet werden. Weiterhin ist die Optimierung sehr aufwendig: Es müssen Schrittweise meist eine Vielzahl von Variationen synthetisiert werden, um die Auswirkungen der Va-riationen zu verstehen und die besten Optimierungen durchzuführen. Zusätzlich ist ein genaues Ver-ständnis des Bindungsmodus notwendig, um die Strukturen systematisch zu verbessern. Dies setzt experimentell aufwendige Methoden, meist Röntgenkristallstrukturanalysen, voraus [71]. Die Suche nach neuen Fragmenten kann ebenfalls mit computergestützten Methoden durchgeführt werden. Dafür können die zu testenden Fragmente, wie bei der strukturbasierten Leitstruktursuche (s. S. 10), mithilfe eines fragmentbasierten virtuellen Screening bestimmt werden [72].

– Synthetische Inhibitoren: Fragmentbasierte Ansätze

Die Firma Hoffman-La Roche führte im Jahr 2000 ein fragmentbasiertes virtuelles Screening durch, nachdem ein fragmentbasiertes experimentelles Screening nicht erfolgreich war [73]. Dafür wurde ein virtuelles Screening von Molekülen mit einem Molekulargewicht von unter 300 Da durchgeführt und die Hits anschließend getestet und validiert. Dabei wurde eine Vielzahl interessanter Fragmente identifiziert. Bemerkenswert ist, dass viele der dort gefundenen Fragmente, bspw. Carbamate, Pyr-rolpyrimidine oder Pyrazole, später bei anderen Leitstruktursuchen wieder auftreten. Aus einem der Fragmente, dem 3-Phenyl-1H-indazol (**7**), wurde schließlich ein potentes Indazolderivat (**8**) entwickelt (s. Abbildung 1.9) [73]. Die Firma Dainippon optimierte diese Indazolderivate 2004 mithilfe intensiver SAR-Studien zu Substanzen (bspw. **9**), die auch *in vivo* gute Aktivitäten zeigten [74].

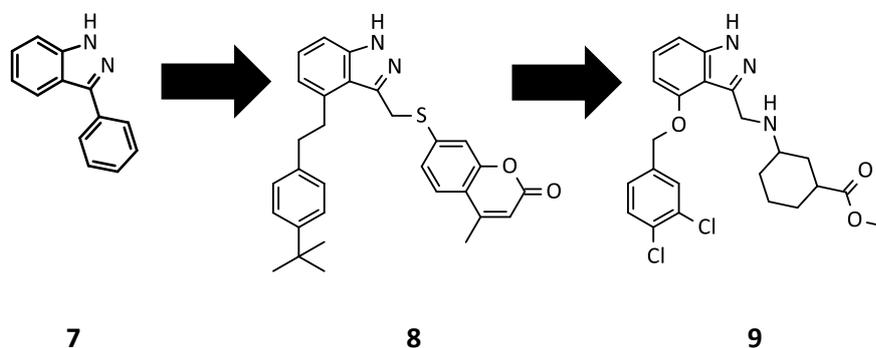


Abbildung 1.9 – Entwicklung verschiedener Indazolderivate aus dem Fragment **7**, welches bei einem fragmentbasierten Screening entdeckt wurde [73].

Von der Firma Hoffman-La Roche wiederum wurde 2007 ein zweites Fragment (**10**, s. Abbildung 1.10) weiterentwickelt. Dafür wurden die Erfahrungen aus den SAR-Eigenschaften und Kristallstrukturen der Cyclothialidine und Aminocumarine genutzt [75].

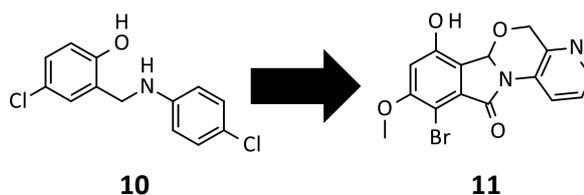


Abbildung 1.10 – Weiterentwicklung des Phenolderivat **10**, welches bei einem fragmentbasierten Screening entdeckt wurde [73, 75].

2012 führte die Firma AstraZeneca einen ähnlichen Ansatz durch: Zunächst wurde mit einem fragmentbasierten virtuellen Screening nach neuen Strukturen gesucht. Anschließend wurde das gefundene Azaindolderivat (**12**, s. Abbildung 1.11) schließlich weiter optimiert. Dabei wurde mithilfe von Röntgenkristallstrukturen und SAR-Analysen eine neue potente Stoffklasse (**13**) entwickelt [76].

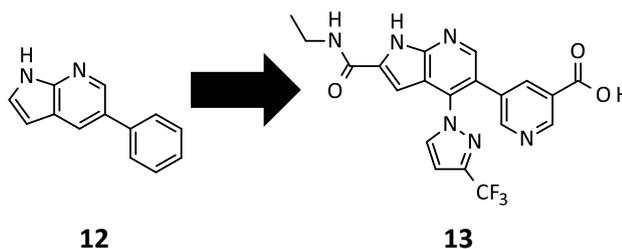


Abbildung 1.11 – Hitstruktur (**12**) in einem fragmentbasierten virtuellen Screening und die optimierte Substanz (**13**).

Die Firma AstraZeneca führte 2012 auch ein rein experimentelles fragmentbasiertes Screening durch und entdeckte dabei unter anderem Pyrrolester (**14**), die wiederum mithilfe von Kristallstrukturen und SAR-Analysen zu Pyrrolamiden weiterentwickelt wurden (s. Abbildung 1.12). Die Substanz **15** konnte in der Folge auch erfolgreich in einem Maus-Modell einer Infektion mit *Streptococcus pneumoniae* angewendet werden [77, 78]. Diese Substanz konnte von AstraZeneca 2014 erfolgreich zu einem Pyrrolamid (**16**) weiterentwickelt werden, das auch in einem Maus-Modell einer *Mycobacterium tuberculosis* (*M. tuberculosis*)-Infektion Aktivität zeigte [79, 80]. Dies ist ein Beispiel dafür, dass die Gyrase eine gute Zielstruktur für die Entwicklung von Antibiotika für sehr unterschiedliche Spezies darstellt.

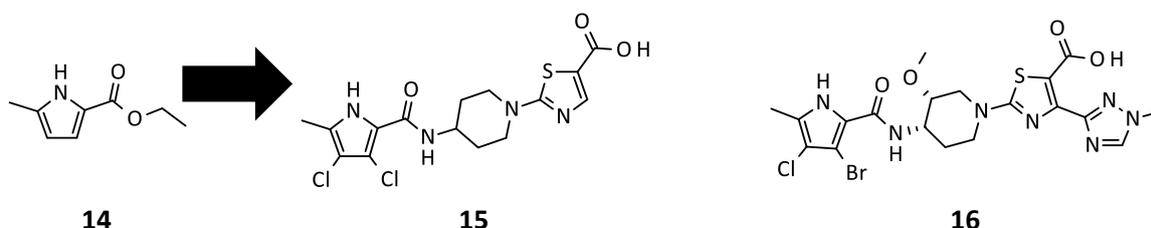


Abbildung 1.12 – Entwicklung von Pyrrolamiden. Das gefundene Fragment (**14**) wurde zu in unterschiedlichen Spezies aktiven Inhibitoren (**15** und **16**) weiterentwickelt.

Während der Untersuchung vom marinen Alkaloid Hymenidin **17** (s. Abbildung 1.13) fiel TOMAŠIČ *et al.* 2015 die Ähnlichkeit dieser Verbindungen zu den zuvor beschriebenen Pyrrolamiden auf (s. Abbil-

dung 1.12) [81]. Daher untersuchten sie ihre synthetisierten Analoga auf ihre inhibitorische Wirkung gegenüber der ATPase der Gyrase. Durch intensive SAR und Dockingexperimente konnte schließlich die Substanz **18** mit guter Aktivität entwickelt werden. In einem weiteren Schritt konnte mithilfe strukturbasierter Methoden diese noch weiter zur Substanz **19** verbessert werden [82].

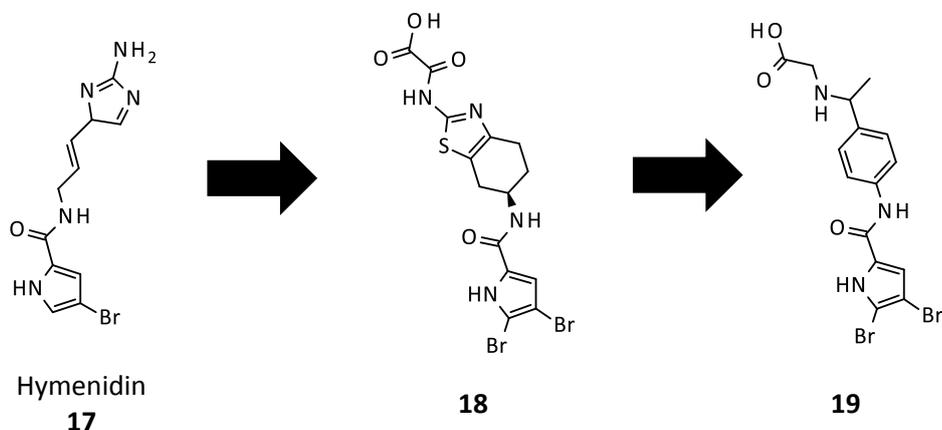


Abbildung 1.13 – Entwicklung von neuen Pyrrolamiden aus dem maritimen Alkaloid Hymenidin.

In einem fragmentbasierten kristallographischen Screening fand die Firma Trius Therapeutics 2013 Pyrrolpyrimidine (**20**) und entwickelte diese weiter (**21**; s. Abbildung 1.14) [83]. Durch weiterführende Untersuchungen der SAR mithilfe von Röntgenkristallstrukturen konnten neue Substanzen mit einem größeren antibakteriellen Spektrum entwickelt werden. Dabei wurde die Grundstruktur zu einem tricyclischen System (**22**) erweitert [84].

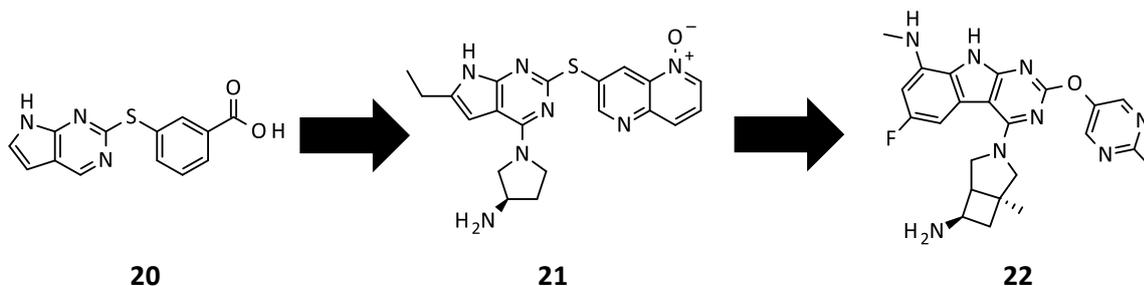


Abbildung 1.14 – Das von Trius Therapeutics gefundene Fragment (**20**) wurde über **21** zu einer tricyclischen Substanz (**22**) mit einem breiten antibakteriellen Spektrum weiterentwickelt.

– Synthetische Inhibitoren: *High Throughput (virtual) Screenings* (HT(v)S)

In einem HTS fand die Firma Vertex Pharmaceuticals 2008 ein Benzimidazolcarbamat-Derivat (**23**) als Hitstruktur (s. Abbildung 1.15) [85]. Mithilfe kristallographischer Untersuchungen wurden intensive SAR-Analysen durchgeführt und neue Derivate mit erheblich höherer Affinität gegenüber der Gyrase synthetisiert. Dabei waren insbesondere die Benzimidazolharnstoffderivate (**24**) überlegen. Die wichtige Hauptinteraktion dieser Stoffklasse findet über den Harnstoff-Rest und den stickstoffhaltigen, anellierten Fünfring statt. Dieses Pharmakophor ist bei sehr vielen Leitstrukturen von verschiedenen Firmen verbreitet (Stand: 2009; s. Abbildung 1.16) [86].

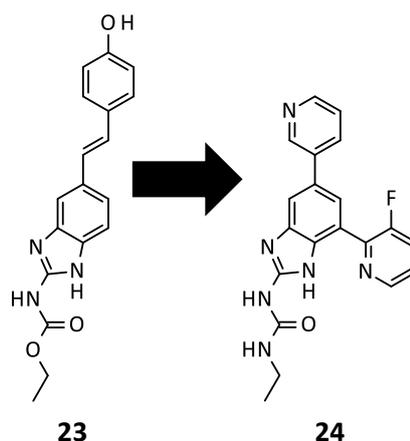


Abbildung 1.15 – Hitstruktur (**23**) in einem HTS-Programm von Vertex und die Weiterentwicklung (**24**).

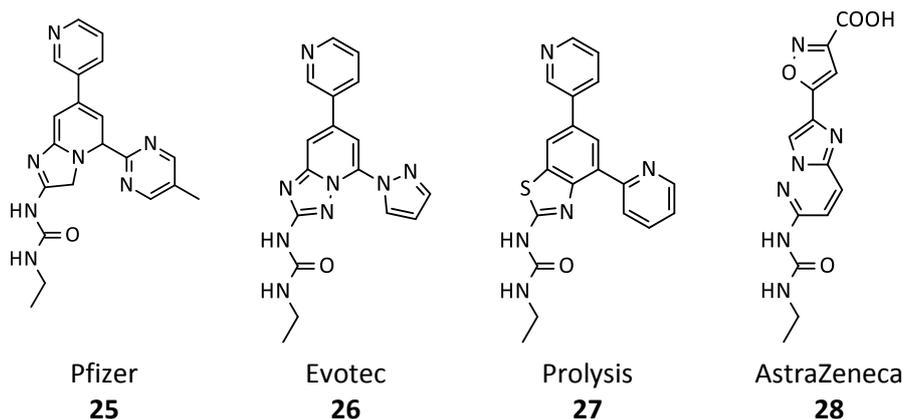


Abbildung 1.16 – Viele der sich in der Industrie in der Entwicklung befindlichen Gyrase-Inhibitoren ähneln sich in ihrer Struktur.

Ausgehend von dem Interaktionsmuster der zuvor dargestellten Benzimidazolharnstoff-Derivate entwickelte die Firma AstraZeneca 2013 Pyridylharnstoffe [87]. Diese besaßen prinzipiell ein ähnliches Interaktionsmuster für den Harnstoffrest und den aromatische Stickstoffatom, waren aber in der weiteren Substitution wesentlich flexibler. Es wurde damit also ein *Scaffold Hopping* (engl. für Gerüstwechsel) für den Benzimidazol-Körper aus **24** durchgeführt. Die daraus entwickelte Substanz **29** wies auch in *in vivo*-Experimenten in einem Krankheitsmodell mit Mäusen eine gute Aktivität auf.

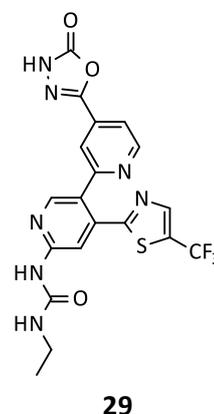


Abbildung 1.17 – Ein aus **24** durch *Scaffold Hopping* entwickeltes Pyridylharnstoffderivat.

In einem HTS-Programm fand die Firma Vertex Pharmaceuticals 2010 weiterhin Pyrazolthiazolderivate (**30**), die, wie bei den Benzimidazolen (s. o.), mithilfe kristallographischer Analysen weiterentwickelt wurden [88]. Dabei erwies sich insbesondere die Carbamat-Substitution am Pyrazolring (z. B. Substanz **31**) wegen seiner hohen Aktivität als sehr vorteilhaft. Die Pyrazolderivate wurden auch noch von anderen Arbeitsgruppen aufgenommen und 2013 (**32**) [89] sowie 2014 (**33**) [90] variiert.

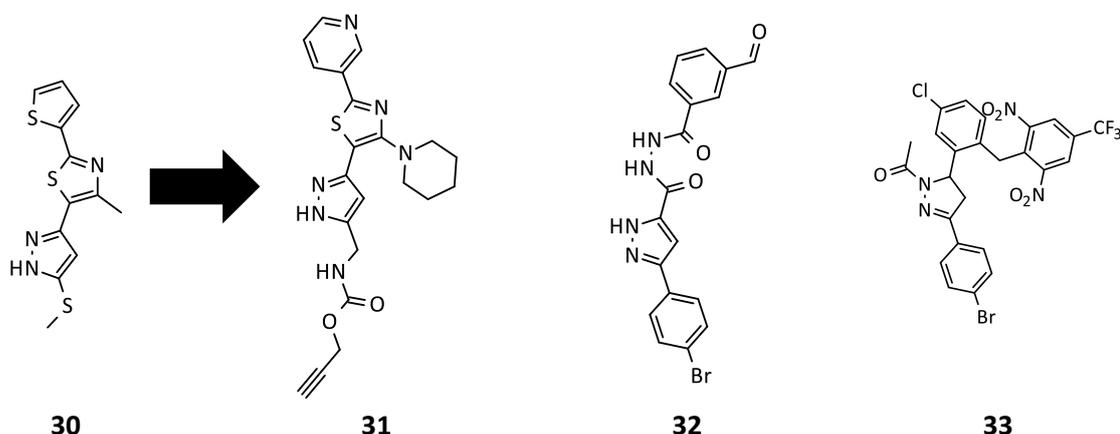


Abbildung 1.18 – Hitstruktur (**30**) in einem HTS-Programm von Vertex und die Weiterentwicklungen (**31**) sowie weitere entwickelte Pyrazolderivate (**32** und **33**).

Mittels eines HTS, das auf Mycobakterien fokussiert war, wurden 2012 erstmals Aminopyrazinamide von AstraZeneca beschrieben [91]. Dabei wurden intensivere SAR-Analysen wieder strukturbasiert über Röntgenkristallstrukturen durchgeführt, um das schließlich bei *M. tuberculosis* sehr potente Derivat (**35**) zu synthetisieren.

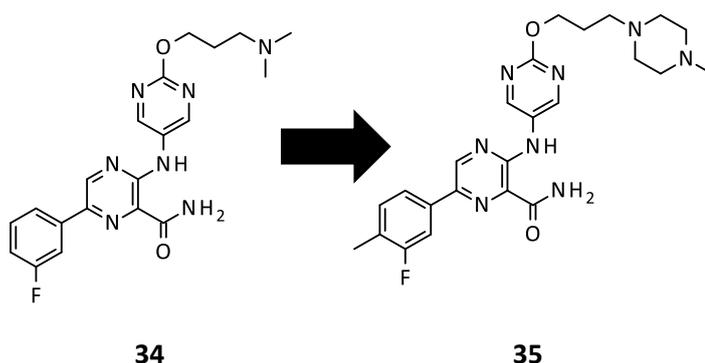


Abbildung 1.19 – Hitstruktur (**34**) und Weiterentwicklung (**35**) eines Inhibitors für *M. tuberculosis*.

BRVAR *et al.* verwendeten die bekannten SAR von Cyclothialidinen (**5**) in einem strukturbasiertem Ansatz 2010 [92]. Dafür filterten sie eine Datenbank von Substanzen pharmakophorbasiert nach verschieden substituierten Phenolen, da die Resorcinstruktur bei den Cyclothialidinen (**5**) eine wichtige Interaktion mit dem Enzym eingeht. Diese Phenolderivate wurden anschließend gedockt und die erfolgversprechenden Kandidaten getestet. Dabei wurden Dihydroxyphenylthiazole (**36**) als neue Substanzklasse entdeckt (s. Abbildung 1.20). Durch Verwendung eines ligandbasierten Screening nach neuen Derivaten für den Thiazolkörper und erneuten strukturbasierten Dockings konnten 2012 auch Dihydroxyrhodaninderivate (**37**) als neue Substanzklasse identifiziert werden [93].

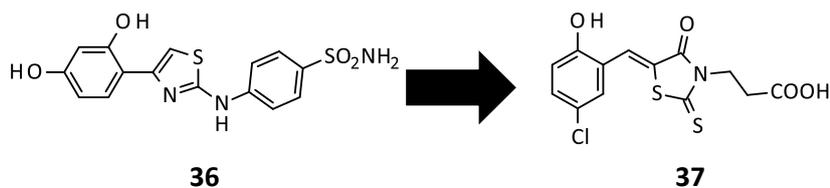


Abbildung 1.20 – Das in einem strukturbasiertem Ansatz gefundene Dihydroxyphenylthiazolderivat (**36**) und das weiterentwickelte Dihydroxyrhodaninderivat (**37**).

In einem ähnlichen Ansatz untersuchte die gleiche Arbeitsgruppe um BRVAR *et al.* 2012 die Aminocumarine: Anhand eines Aminocumarins wurde ein pharmakophorbasiertes Screening von Liganden durchgeführt. Dafür wurden 5 Mio. Substanzen anhand eines Pharmakophors, das aus dem Aminocumarin entwickelt wurde, auf 400 Substanzen reduziert. Diese Hits wurden in einem strukturbasiertem Docking erneut bewertet [94]. Nach der visuellen Auswahl konnten 4,5'-Bithiazole (**38**) als Hitstrukturen gefunden und experimentell bestätigt werden (s. Abbildung 1.21).

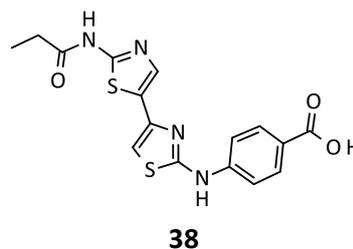


Abbildung 1.21 – Das in einem virtuellen Screening identifizierte 4,5'-Bithiazol.

Zusammengefasst konnten schon eine Reihe diverser kompetitiver Inhibitoren der ATPase der GyrB-Untereinheit der Gyrase entwickelt werden. Dabei hat insbesondere das Interesse innerhalb der Industrie in den letzten Jahren erheblich zugenommen. Einerseits zeigt dies, dass auch dort die Notwendigkeit für die Entwicklung von neuen Antibiotika wahrgenommen und diese Zielstruktur als vielversprechend angesehen wird. Andererseits helfen die Erkenntnisse, wie SAR oder neue Röntgenkristallstrukturen, bei der weiteren Entwicklung von neuen Leitstrukturen. Bisherige Inhibitoren wurden jedoch hauptsächlich durch fragmentbasierte Screenings oder HTS bestimmt. Bis 2011 führten nur BRVAR *et al.* eine strukturbasierte Leitstruktursuche durch (s. o.). Diese war allerdings aufgrund der Beschränkung auf substituierte Phenole stark eingeschränkt. Auch das von der gleichen Arbeitsgruppe 2012 durchgeführte HTVS wurde zunächst ligandenbasiert durchgeführt und nur ein kleiner Anteil an Substanzen letztendlich strukturbasiert gedockt (s. o.).

Trotz der Bemühungen und der bereits gefundenen Inhibitoren befinden sich in der Industrie, soweit bekannt, lediglich drei Klassen in der Weiterentwicklung: die Benzimidazolderivate von Vertex (s. S. 20), die auch von zahlreichen Firmen durch bioisostere Strukturvariationen aufgegriffen wurden, die tricyclischen Strukturen von Trius Therapeutics (s. S. 19) und die Pyrrolamide von AstraZeneca (s. S. 18) [95]. Bis zum jetzigen Zeitpunkt, Sommer 2015, befindet sich allerdings keine Substanz, die die ATPase der GyrB-Untereinheit kompetitiv inhibiert, in einer klinischen Prüfung.

1.4 Zielsetzung

Das Ziel dieser Arbeit war die Planung und Durchführung einer strukturbasierten Leitstruktursuche (s. Abschnitt 1.2 auf S. 9) nach einem kompetitiven Inhibitor der ATPase der GyrB-Untereinheit von *Escherichia coli* (s. Abschnitt 1.1 auf S. 4). Die Umsetzung dieser Aufgabe erfolgte mit der in Abschnitt 1.2 auf S. 11 vorgestellten strukturbasierten Leitstruktursuche. Bei der Durchführung sollte ein besonderer Fokus auf den strukturbasierten Ansatz gelegt werden. Deswegen sollte ein Docking den ersten starken Reduktionsschritt für die Substanzbibliothek darstellen. Die Entwicklung und Etablierung der biologischen Testsysteme für die Leitstruktursuche stellten neben ihrer Durchführung einen integralen Bestandteil dieser Arbeit dar.

Die abschließenden Ziele waren die Identifizierung und Validierung neuer chemischer Entitäten, die dazu geeignet sind, zu Leitstrukturen für Inhibitoren der ATPase der GyrB-Untereinheit weiterentwickelt zu werden. Die Zielsetzung ließ sich in folgende Abschnitte gliedern (s. Abbildung 1.22):

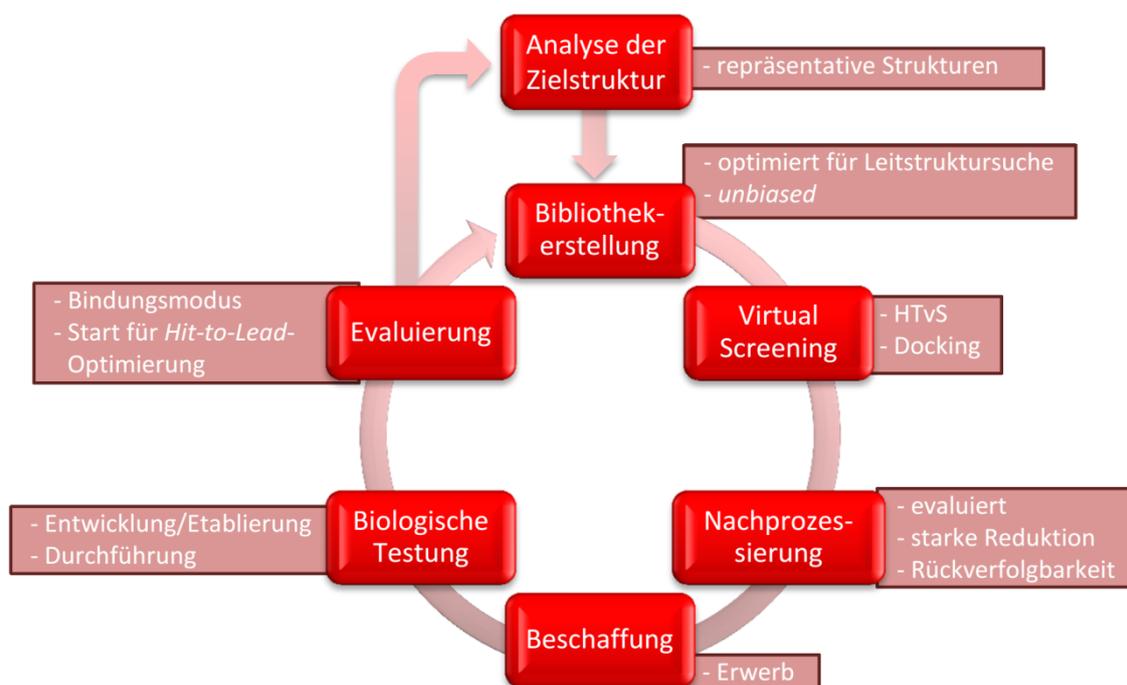


Abbildung 1.22 – Workflow für die Durchführung der strukturbasierten Leitstruktursuche. Angegeben sind die zentralen Elemente eines jeden Arbeitsschrittes.

– Analyse der Zielstruktur

Im Rahmen dieser Arbeit sollte ein virtuelles Screening mithilfe eines Docking-Tools durchgeführt werden (s. u.). Da die biologische Testung mit der Gyrase des Bakteriums *Escherichia coli* (*E. coli*) erfolgte (s. u.), mussten Röntgenkristallstrukturen dieser Spezies untersucht und vorbereitet werden.

Dafür sollten verschiedene Kristallstrukturen untersucht und mögliche Unterschiede in der Bindetasche analysiert werden. Durch die Auswahl geeigneter Proteinstrukturen sollte ein möglichst großer konformativer Raum der Bindetasche abgebildet werden, der für das virtuelle Screening verwendet wurde.

– Erzeugung der Bibliothek

Die Substanzbibliothek, die für das Docking verwendet wurde, sollte aus einem möglichst diversen Datensatz bestehen. Es wurde daher auf pharmakophore oder strukturelle Filter, die von bekannten Inhibitoren abgeleitet werden können, verzichtet. Bisher wurde noch kein rein strukturbasiertes Design von Inhibitoren für die ATPase der Gyrase in der Literatur beschrieben (s. Abschnitt 1.3 auf S.14). Durch den hier gewählten Ansatz können neue chemische Entitäten, die möglicherweise völlig andere Strukturelemente und -eigenschaften als bekannte Inhibitoren besitzen, gefunden werden. Dies bedeutete jedoch nicht, ungefilterte Substanzen im Screening zu verwenden. Für die Generierung der Bibliothek mussten die Substanzen dennoch Kriterien erfüllen, die bspw. für die Entwicklung zu Leitstrukturen wichtig waren (s. Abschnitt 1.2 auf S. 9). Es sollte jedoch explizit keine „Verzerrung“ (engl. *Bias*) hin zu bereits bekannten Inhibitoren stattfinden.

– Virtuelles Screening

Aufgrund der großen Substanzbibliothek standen sehr viele Substanzen zur Verfügung. Daher handelte es sich bei dem Screening-Prozess um ein HTvS. Dieses sollte mithilfe eines Dockings erfolgen, in dem Substanzen in die ausgewählten und vorbereiteten Bindetaschen des Enzyms unter Verwendung der dort aufgestellten Hypothesen (z. B. Schlüsselinteraktionen, Wassermoleküle) des ersten Schrittes gedockt wurden. Damit wurden die Substanzen im ersten entscheidenden Reduktionsschritt rein strukturbasiert ausgewählt; also basierend darauf, ob sie mit der Bindetasche des Proteins ausreichende Interaktionen durchführen können.

– Nachprozessierung

Die Nachprozessierung musste sehr drastische Reduktionsschritte enthalten, da im HTvS eine hohe Anzahl an Ergebnissen auftrat, aber letztendlich nur wenige Substanzen erworben werden konnten. Daher musste der Prozess sorgfältig evaluiert und geplant werden. Die Evaluation sollte durch die Mitverwendung von bekannten Inhibitoren im Prozess gesteuert werden. Methodisch konnte hierbei ein weites Spektrum an Computermethoden (z. B. Moleküldynamik-Simulationen, Rescorings, Redockings) eingesetzt werden, diese orientierten sich vor allem an der Durchführbarkeit (bspw. Anzahl der Substanzen). Trotz der starken Reduktion sollte dabei gewährleistet werden, dass gewonnene Informationen aus vorherigen Schritten nicht verloren gingen. Dafür sollte ein ausgewähltes Molekül jeweils eine Gruppe von Substanzen repräsentieren. In den folgenden Schritten konnten dann diese Repräsentanten mithilfe aufwendigerer Methoden weiter analysiert werden und die restlichen

Gruppenmitglieder bei Problemen (z. B. beim Erwerb) oder der Hit-Erweiterung berücksichtigt werden.

- Beschaffung

Ausgewählte Substanzen sollten nach der Nachprozessierung käuflich erworben werden. Daraus ergab sich die Anforderung an die Substanzbibliothek, aus erwerbbaaren Substanzen zu bestehen.

- Biologische Testung

Für die biologische Testung wurden zwei zueinander orthogonale Assays benötigt. Da es noch keine etablierten Methoden dafür gab, mussten beide Assays zunächst entwickelt werden. Für die Umsetzung stand die Gesamtgyrase von *E. coli* zur Verfügung. Einer der Assays sollte für die Testung aller Substanzen verwendet werden und daher einen höheren Durchsatz ermöglichen. Der Assay für die Hit-Validierung konnte anspruchsvoller sein, da hier eine kleinere Anzahl an Substanzen getestet wurde. Natürlich richtete sich die Wahl der Assays auch nach der Durchführbarkeit und den zur Verfügung stehenden Geräten. Das Ziel bestand in der sicheren Testung und Evaluierung der zuvor ausgewählten Substanzen.

- Evaluierung

Zum Abschluss sollten die Ergebnisse ausgewertet werden. Zum einen sollte der Gesamtprozess evaluiert werden, um bspw. die Ausgangshypothesen der Interaktionsmöglichkeiten zu bewerten. Zum anderen sollten die einzelnen Hits analysiert und damit die erste Hypothese für den Bindungsmodus erstellt werden. Diese beiden Auswertungen sind notwendig, um eine erfolgreiche Weiterentwicklung (*Hit to Lead*), die aber nicht mehr Teil dieser Arbeit ist, zu gewährleisten.

2 Analyse der Zielstrukturen

2.1 Ausgangsstrukturen

Für die Analyse der Bindetasche wurde die *Protein Data Bank* [58] nach Röntgenkristallstrukturen (im Folgenden Kristallstruktur genannt) der ATPase der GyrB-Untereinheit von *Escherichia Coli* (*E. Coli*) durchsuchtⁱⁱⁱ. In Tabelle 2.1 sind die Ergebnisse dargestellt.

Tabelle 2.1 – Röntgenkristallstrukturen von der Gyrase in der *Protein Data Bank*. Angegeben sind der PDB-Code, die Auflösung (R), der Organismus, der kokristallisierte Ligand (Abk. s. Text), die Fragmentgröße, fehlende Aminosäuren (AS), Besonderheiten und die Quelle.

PDB-Code	R (Å)	Organismus	Ligand	Fragment	Fehlende AS	Besonderheiten	Quelle
1EI1	2.30	<i>E. coli</i>	ANP	43 kDa A: 2 – 392, B: 402 – 792	–	Mutation Y5S	[35]
1AJ6	2.30	<i>E. coli</i>	NOV	24 kDa, A: 12 – 217	83 – 87, 105 – 111	Mutation R136H	[69]
1KZN	2.30	<i>E. coli</i>	CBN	24 kDa A: 15 – 219	83 – 87, 101 – 115	–	[96]
1KIJ	2.30	<i>T. thermophilus</i>	NOV	43 kDa A1: 9 – 392, A2: 9 – 392	–	vollständig aufgelöst	[97]
3G7E	2.20	<i>E. coli</i>	B46	24 kDa A: 15 – 217	–	–	[88]

Grundsätzlich konnten drei verschiedene Arten von kokristallisierten Liganden gefunden werden: ein substratähnlicher Ligand (PDB-Code: 1EI1), verschiedene Naturstoffe (1AJ6, 1KZN und 1KIJ) und ein rational entwickelter Inhibitor (3G7E). Bei dem Liganden der ersten Struktur 1EI1 handelte es sich um ANP (5'-Adenylyl- β - γ -imidodiphosphat), ein Analogon des endogen vorkommenden Substrats ATP. In der zweiten Kristallstruktur 1AJ6 band das bereits in der Therapie eingesetzte Aminocumarin Novobiocin (NOV). Neben Novobiocin war noch die Kristallstruktur 1KZN mit Clorobiocin (CBN), ebenfalls ein Aminocumarin, in *E. coli* vorhanden. Bei beiden Strukturen fehlten jedoch die gleichen Aminosäuren in unmittelbarer Umgebung der Bindetasche. Dafür war eine vollständig aufgelöste Kristallstruktur mit Novobiocin von *Thermus thermophilus* (*T. thermophilus*) vorhanden (1KIJ). In der letzten Kristallstruktur 3G7E besetzte ein synthetisch und rational entwickeltes Pyrazolderivat die Bindetasche.

Da für das Screening ein möglichst großer konformativer Raum der Zielstruktur abgebildet werden sollte, wurde für jede Art von Ligand eine Röntgenkristallstrukturen untersucht: 1EI1, 1AJ6 und 3G7E.

ⁱⁱⁱ Die Analyse fand im Frühjahr 2011 statt. Alle nachträglich veröffentlichten Kristallstrukturen wurden, soweit wie möglich, berücksichtigt.

Schon der Vergleich der Ausdehnung des Bindungsbereiches der verschiedenen Liganden verdeutlicht die große Diversität an Konformationen, die die Bindetasche annehmen kann (s. Abbildung 2.1). Während sich ANP (blau) sehr weit nach rechts ausdehnt, zeigt Novobiocin (gelb) eine ausgeprägte Bindung im Bereich des Ausgangs der Bindetasche und das Pyrazolderivat (rot) eine sehr tiefe Bindung in die Bindetasche selbst. Neben diesen deutlichen Unterschieden gibt es aber auch Bereiche, die sich bei allen Strukturen überlagern. Dieser in der Mitte links befindliche Bereich wird beim ANP von der Adeninsubstruktur besetzt. Im Folgenden wird jede Kristallstruktur einzeln untersucht. Für die Auswertung der Interaktionen wurden die von BISSANTZ *et al.* ermittelten Geometrien verwendet [99].

– 1E11

Der natürliche Ligand dieser Bindetasche ist ATP. Es wird von der GyrB-Untereinheit der Gyrase hydrolysiert. Ohne die GyrA-Untereinheit, wenn sich folglich also kein Tetramer ausbildet (s. Abschnitt 1.1 auf S. 4), zeigt die GyrB-Untereinheit selbstständig ATPase-Aktivität [63]. Daher konnte die GyrB-Untereinheit bisher nicht mit ATP auskristallisiert werden. 2014 gelang jedoch STANGER *et al.* die Kristallisation des Übergangszustandes von ADP mit dem abgespaltenen Phosphatprodukt (PDB-Code: 4PRX) [100]. Obwohl die Hydrolyse eine große Änderung in der Konformation der gesamten Gyrase bedeutete, war die ATP-Bindetasche weitestgehend identisch mit 1E11.

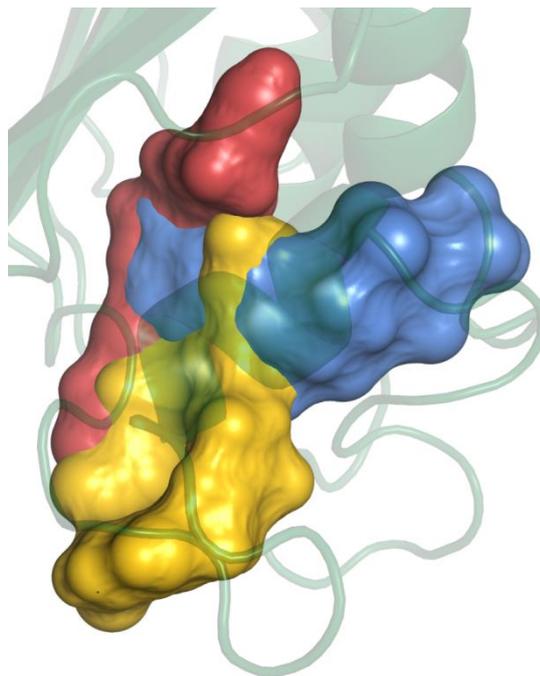


Abbildung 2.1 – Vergleich der verschiedenen Liganden. In dieser Darstellung sind die Oberflächen von ANP (blau), Novobiocin (gelb) und das Pyrazolderivat (rot) der verschiedenen Kristallstrukturen nach dem *Alignment* der Proteine dargestellt. Zur Orientierung ist die Proteinstruktur von 1E11 (grün) transparent gezeigt.^{iv}

^{iv} Alle dreidimensionalen Strukturbilder wurden mit PyMOL [98] erzeugt (s. Experimenteller Teil 11.3.1 auf S. 202).

Abbildung 2.2 zeigt die Bindetasche von 1E11 mit einer transparenten Oberfläche des Enzyms. Strukturell lässt sich die Bindetasche in zwei Bereiche unterteilen: Auf der einen Seite befindet sich der Adeninkörper des ANP mit drei Wassermolekülen, auf der anderen Seite der Phosphatrest mit vier Wassermolekülen. Die beiden Bereiche sind über die Ribose verbunden. Auffällig ist das Wassermolekül, welches sich in der Abbildung ganz links neben den Adeninrest aufhält (s. orangener Kreis). Dieses befindet sich innerhalb des nach Standardparametern berechneten vdW-Radius des Proteins. Dies stellt schon einen ersten Hinweis für eine sehr starke Bindung des Wassers an das Protein dar.

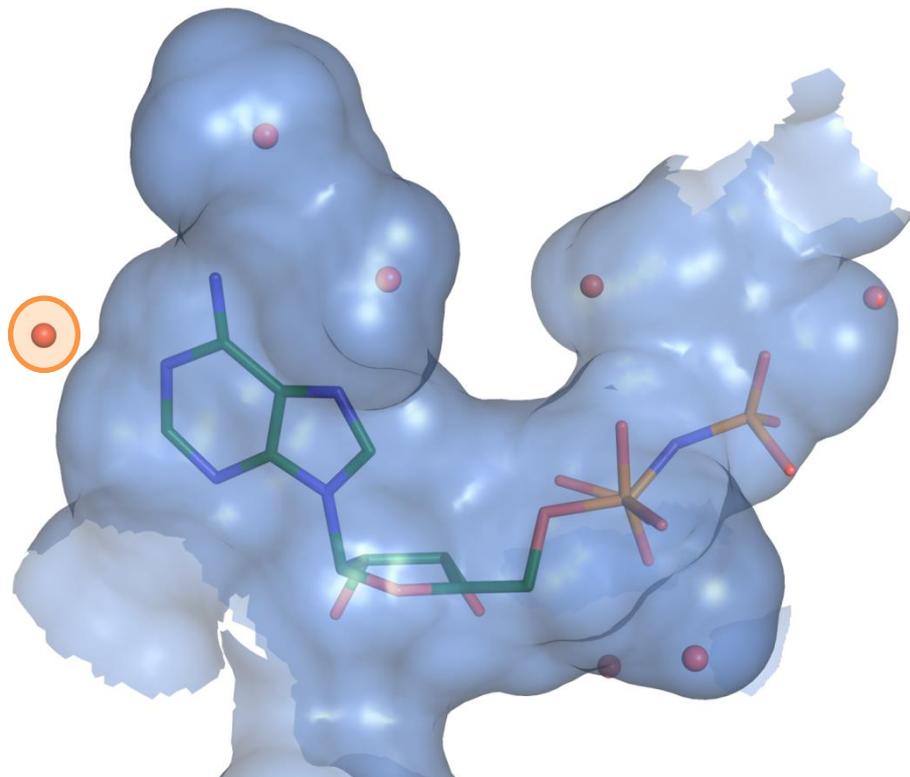


Abbildung 2.2 – Übersicht der Bindetasche von 1E11. Die Bindetasche lässt sich in zwei Bereiche unterteilen. Auf der linken Seite befindet sich der Adeninbereich, rechts die Ribose- und Phosphattasche. Ein Wassermolekül befindet sich innerhalb des nach Standardparametern berechneten vdW-Radius des Proteins (orangener Kreis).

In Abbildung 2.3 sind die wichtigen Aminosäuren für den Bindungsbereich des Adenins dargestellt. Ile78 liegt in einer guten Entfernung für hydrophobe Interaktion mit dem Purinkörper. Das Wasser links neben dem Adeninkörper befindet sich in einem sehr engen Interaktionsmuster. Es geht eine direkte Wasserstoffbrückenbindung mit dem Liganden ein (Entfernung: 2.5 Å). Weiterhin interagiert dieses Wasser mit dem Backbone-Stickstoff von Gly77 (2.7 Å), mit dem Carboxylat vom Asp73 (2.6 Å) und der Alkoholfunktion des Thr165 (2.8 Å). Damit befinden sich alle Interaktionen an der unteren Grenze der optimalen Wasserstoffbrückendbindungslänge von 2.6 – 3.0 Å. Den einzigen direkten hydrophilen Kontakt des Adeninkörpers mit einer Aminosäure stellt Asp73 ebenfalls über das Carboxylat (2.6 Å) her. Die anderen beiden Wassermoleküle in diesem Bereich befinden sich in einer hydrophoben, von außen nicht zugänglichen Tasche. Diese wird unter anderem von Val71, Val43 und Val167 ausgebildet. Das hinter Val167 befindliche Wasser besitzt keine guten Interaktionen, sowohl Asp73 als auch das Backbone-Carbonyl von Val71 liegen außerhalb guter Interaktionsabstände. Das zweite Wasser vermittelt eine Interaktion von Asn46 (3.0 Å) zum Adeninkörper (2.8 Å), besitzt sonst jedoch keine weiteren Interaktionen.

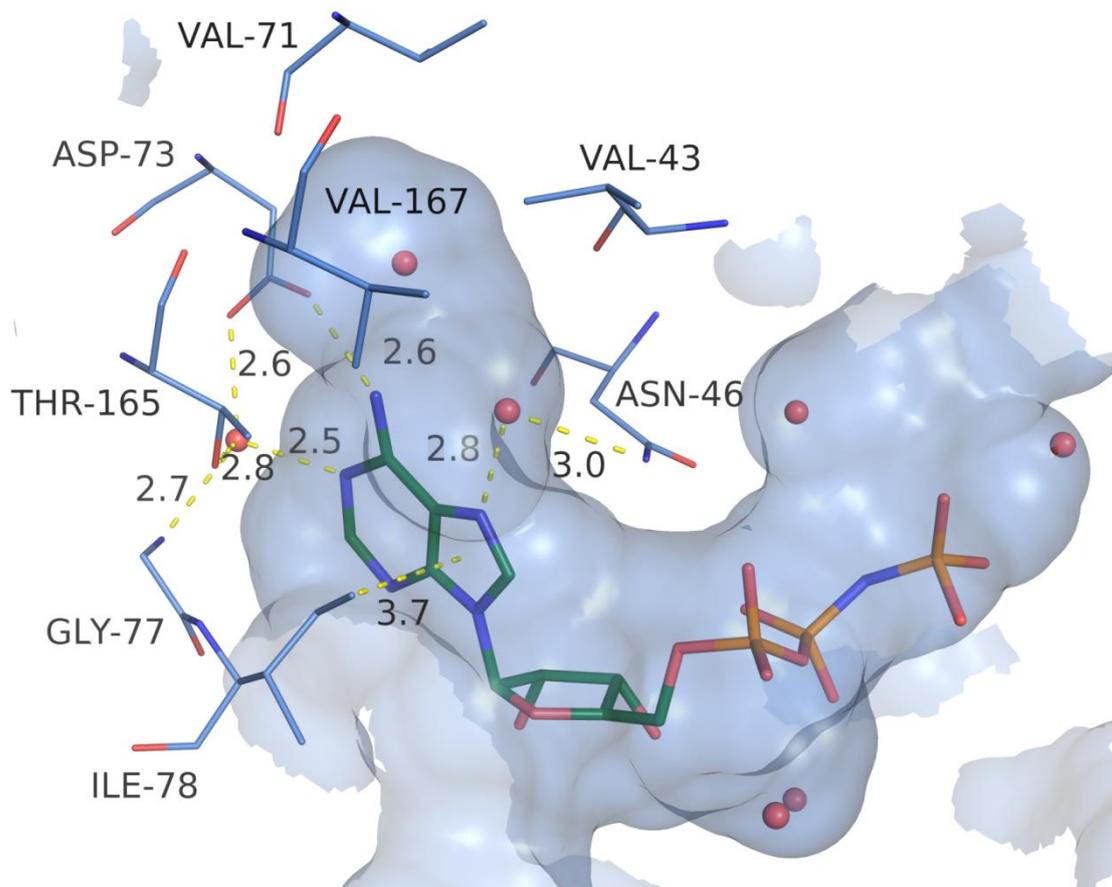


Abbildung 2.3 – Bindungsbereich des Adeninkörpers in 1E11. Wichtige Aminosäuren sind als Linien dargestellt.

Bei der Bindung des Phosphatrestes sind aufgrund der vielen Ladungen am Phosphat und der Hydroxylgruppen der Ribose zahlreiche hydrophile Interaktionen und Wassermoleküle vorhanden (s. Abbildung 2.4). An dieser Stelle findet auch die katalytische Abspaltung des γ -Phosphats statt [100]. Zugunsten der Übersicht sind diese Interaktionen nicht eingezeichnet. Viele Aminosäuren sind über Wasserstoffbrückenbindungen an der Bindung beteiligt: Glu42, Ile94, Ala100, Gly102, Lys103, His116, Gly117, Gly119, Val120, Gln335 und Lys337.

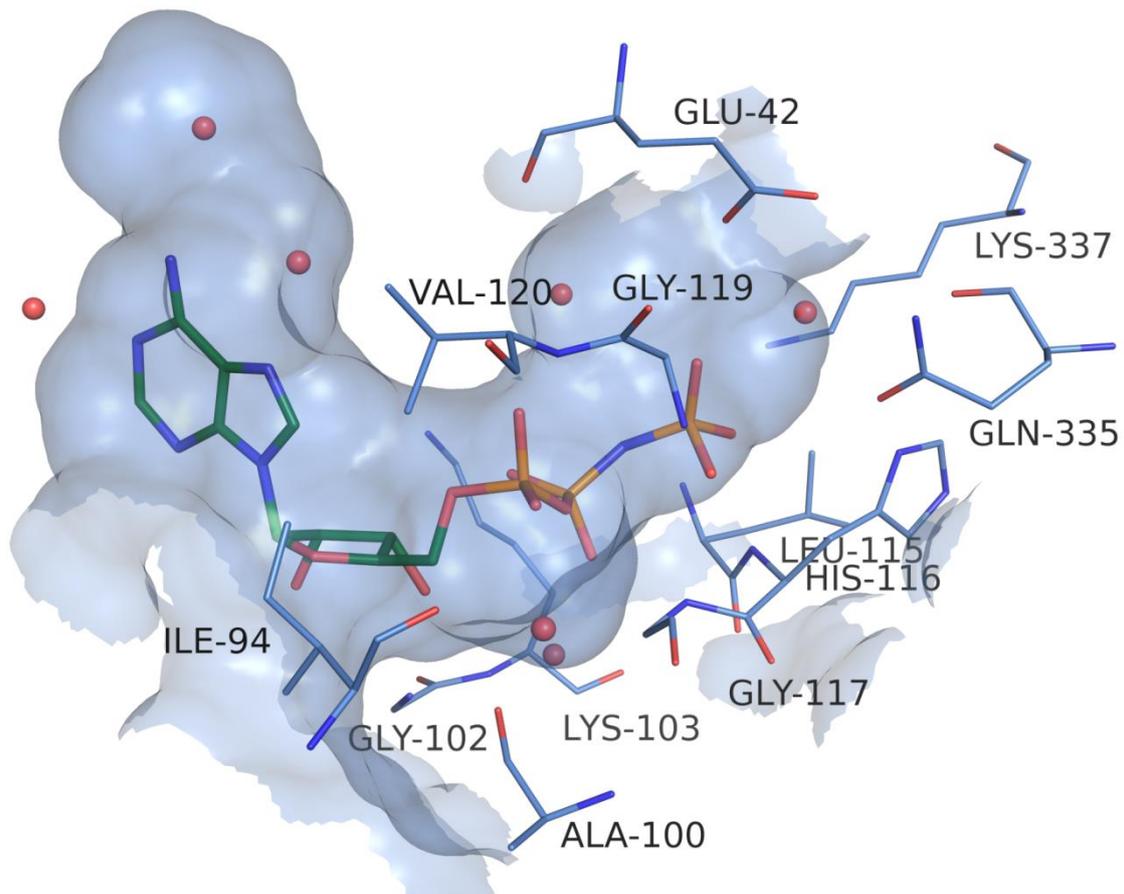


Abbildung 2.4 – Bindungsbereich der Ribose und des Phosphats in 1E11. Wichtige Aminosäuren sind als Linien dargestellt.

– 1AJ6

Einen weiteren interessanten Inhibitor stellt das Aminocumarin Novobiocin dar (s. Abschnitt 1.3 auf S. 14). Bei der einzig verfügbaren Kristallstruktur von Novobiocin mit *E. coli* handelt es sich um eine Mutante, die dem Bakterium eine Resistenz gegenüber Novobiocin verleiht [69]. Weiterhin sind wichtige Aminosäuren in der Nähe der Bindungstasche nicht aufgelöst. Dies betrifft sowohl den Loop A (Aminosäuren 78–86) als auch den Loop

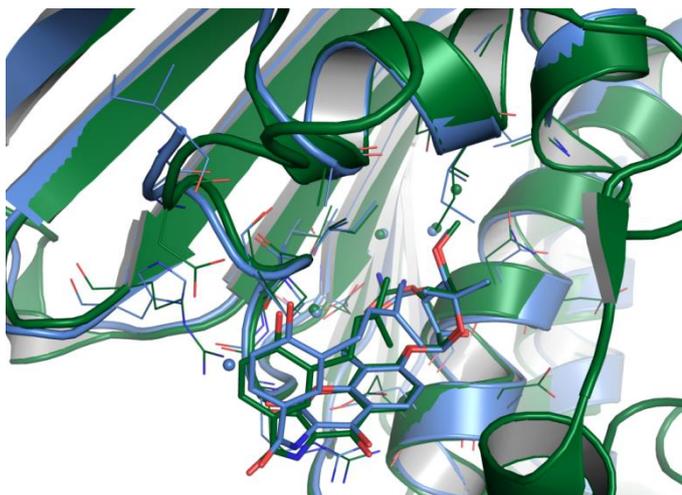


Abbildung 2.5 – Vergleich der Kristallstruktur von 1AJ6 (blau, *E. coli*) und 1KIJ (grün, *T. thermophilus*).

B (Aminosäuren 98 – 118) [96]. Bei der Mutation handelt es sich um eine Punktmutation: R136H. Der Austausch dieser Aminosäure verändert den Bindungsmodus von Novobiocin nicht, führt aber zu einer Einbindung eines zusätzlichen Wassermoleküls, wodurch die Dissoziationskonstante des Novobiocin-Komplexes aufgrund entropischer Kosten auf das 40fache ansteigt (s. auch Abschnitt 2.3.2 auf S. 50)[69]. SAÍZ-URRA *ET AL.* entwickelten ein Protokoll, um die native Aminosäuresequenz und Konformation wiederherzustellen [101]. Damit bestand die Möglichkeit, diesen Liganden in *E. coli* zu untersuchen. Zur Analyse der Bindetasche wurde jedoch zunächst die Kristallstruktur von Novobiocin mit *T. thermophilus* (1KIJ) verwendet [97]. Diese Röntgenkristallstruktur wies große Ähnlichkeit zur Struktur von *E. coli* auf (s. Abbildung 2.5) und ist im Bindungsbereich kristallographisch vollständig aufgelöst.

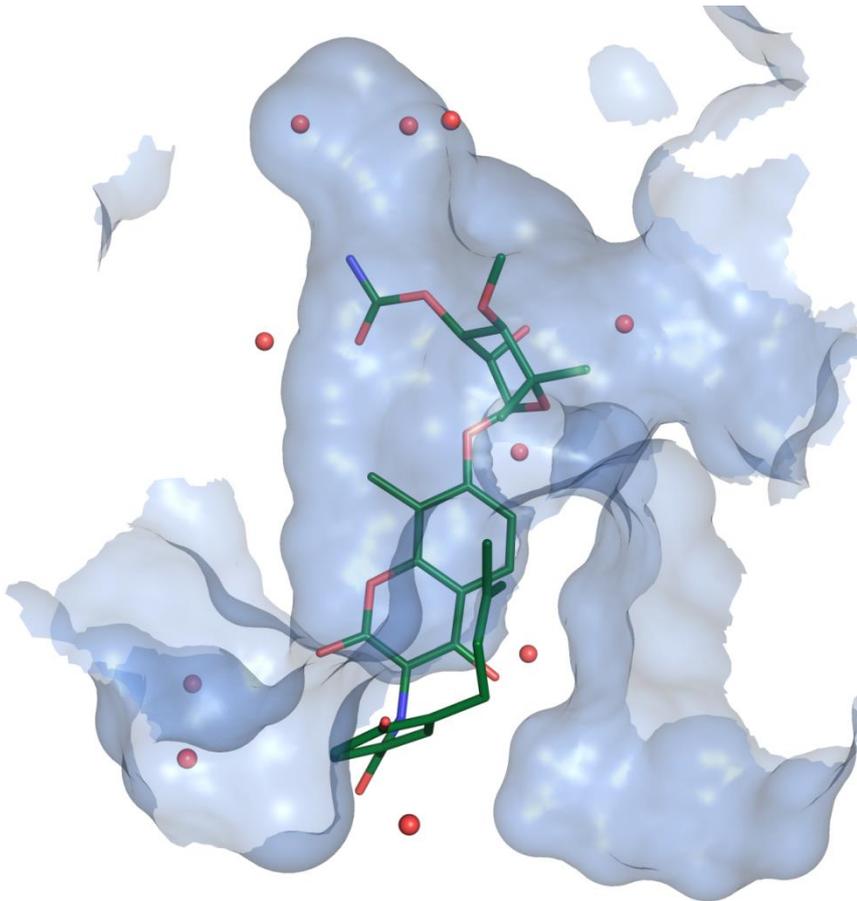


Abbildung 2.6 – Übersicht der Bindetasche von *T. thermophilus* mit Novobiocin (PDB-Code: 1KIJ).

Abbildung 2.6 zeigt die Bindetasche von 1KIJ. Im Vergleich zu 1EI1 unterscheidet sich der Bindungsbereich deutlich. Übereinstimmung gibt es lediglich im Bereich des Carbamats und der Noviose, dem Zucker des Novobiocins. Im weiteren Verlauf bindet Novobiocin in Richtung des Ausgangs der Bindetasche. Dafür bleibt der gesamte Bindetaschenbereich, in dem das Phosphat des ANP bindet, frei. Tief in der Tasche eingeschlossen befinden sich wieder mehrere Wassermoleküle. Das auf der linken Seite befindliche Wassermolekül ist wieder sehr eng an die Proteinstruktur gebunden und befindet sich innerhalb des bestimmten vdW-Radius des Enzyms. In der tieferen Bindetasche finden sich noch drei weitere Wassermoleküle.

Über das Carbamat findet die sehr enge Wechselwirkung des Liganden mit dem Wasser, Asp72, Gly76 und Thr166 statt, die in dieser Form in allen Strukturen auftritt (s. Abbildung 2.7). Obwohl es sich hier um eine andere Spezies handelt, ist dieses Interaktionsmuster einschließlich der beteiligten Aminosäuren mit der Kristallstruktur von *E. coli* identisch. Novobiocin interagiert auch mit dem katalytisch wichtigen Asn45 (entspricht Asn46 in *E. coli*), diese wird aber nicht mehr über die Seitenkette sondern über das Backbone-Carbonyl zur freien Hydroxylgruppe des Zuckers vermittelt. Hydrophobe Interaktionen finden über die Methoxygruppe mit Ala119 und den Dimethylgruppen mit Val117 und Ile93 statt. Damit sind hydrophobe Interaktionen in der Seitentasche vorhanden, die bei dem natürlichen Liganden mit dem Phosphat besetzt ist. Der Zucker interagiert weiterhin mit seinem glykosidisch gebundenen Sauerstoff über ein Wassermolekül mit Glu49. Die drei in der tiefen Tasche liegenden Wassermoleküle nehmen an keiner Interaktion mit dem Liganden teil.

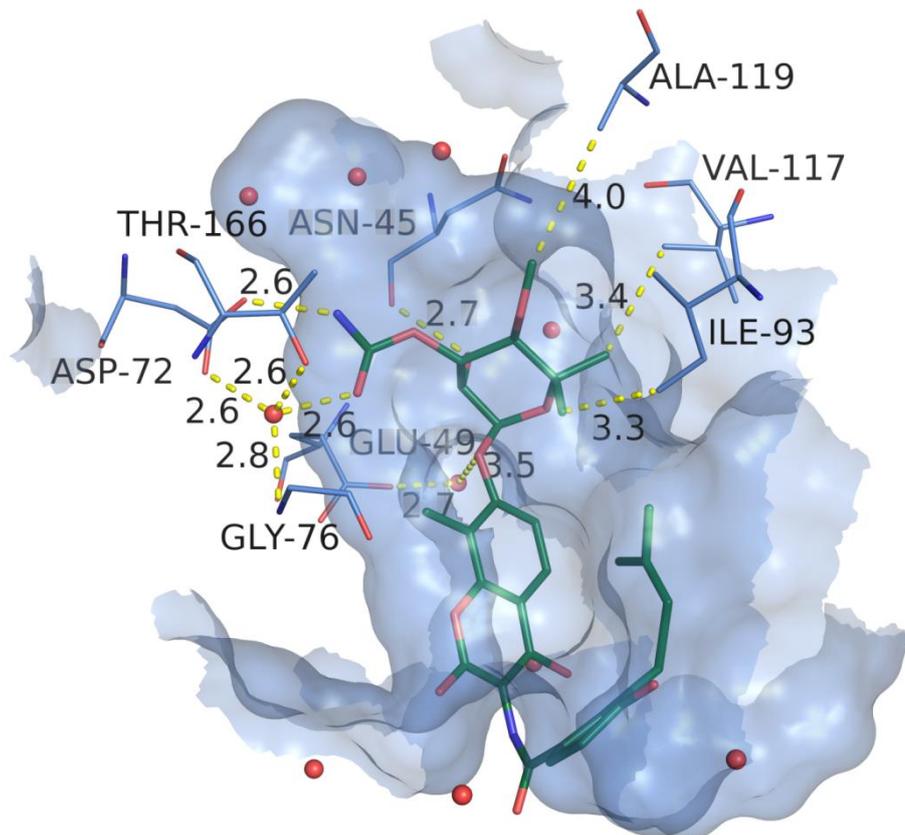


Abbildung 2.7 – Analyse der tiefen Bindetasche von 1KIJ.

Im äußeren Bereich findet ein π -Kationen-Stacking zwischen Arg75 und dem Cumarinring von Novobiocin statt (s. Abbildung 2.8). Arg135 kann eine direkte Wasserstoffbrückenbindung mit dem Carbonyl-Sauerstoff des Cumarinringgrundgerüsts eingehen. Hierbei handelt es sich um das in der Kristallstruktur 1AJ6 mutierte Arginin. Weiterhin ist auch die Hydroxylfunktion vom Cumarin in einer guten Entfernung für eine Wasserstoffbrücke mit Lys102 und einem Wassermolekül hinter dem Novobiocin. Der Isobutylrest befindet sich in einer sehr hydrophoben Umgebung und kann zahlreiche Interaktionen mit den verschiedenen Seitenketten durchführen: Ala89, Ile93, Phe103.

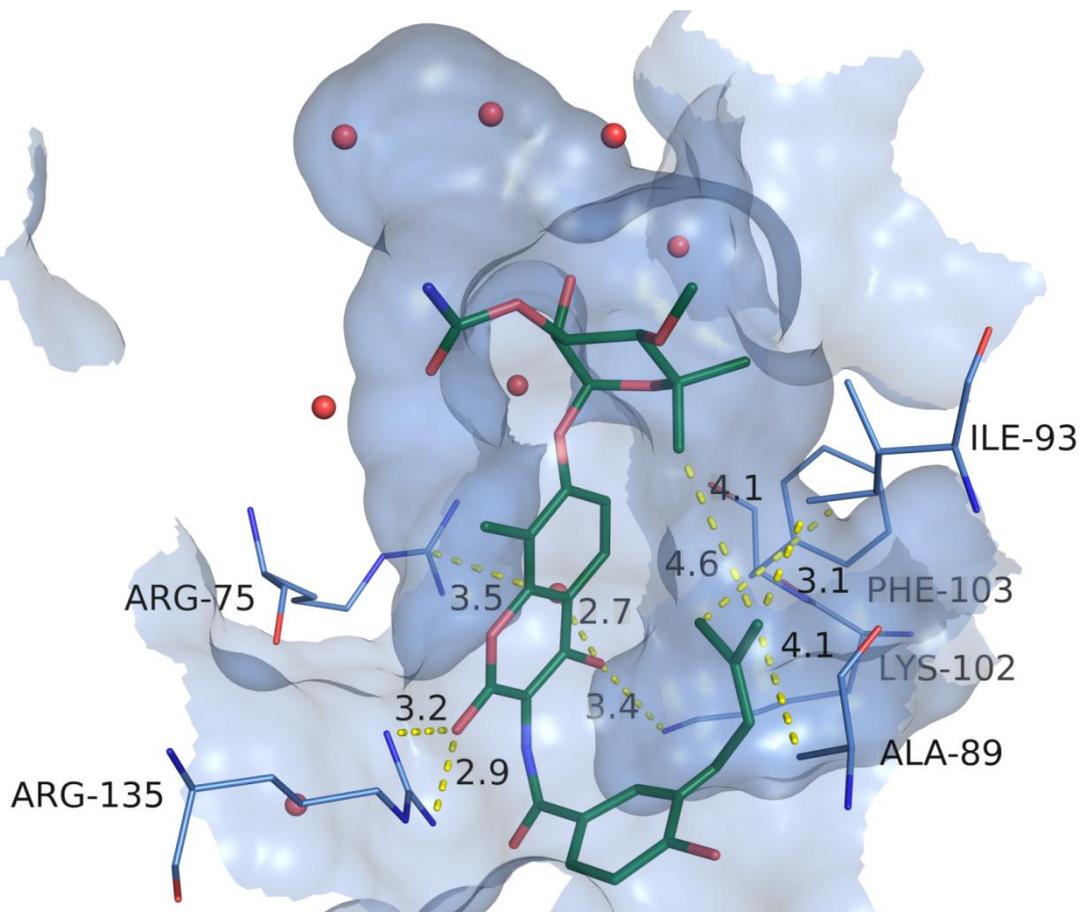


Abbildung 2.8 – Analyse des Bindetaschenausgangs von 1KIJ.

– 3G7E

RONKIN *et al.* führten eine strukturbasierte Optimierung von Pyrazolthiazolderivaten zu nanomolaren Inhibitoren durch [88]. Eine dieser optimierten Substanzen wurde als Kristallstruktur (PDB-Code: 3G7E) veröffentlicht. Der Ligand in der PDB 3G7E ist mit dem IUPAC-Namen „Prop-2-yn-1-yl {[5-(4-piperidin-1-yl-2-pyridin-3-yl-1,3-thiazol-5-yl)-1*H*-pyrazol-3-yl]methyl}carbamate“ angegeben. Die Substanz ist auf der Homepage der Protein Data Bank^v entsprechend des Namens dargestellt. In der Veröffentlichung von RONKIN *et al.* wird jedoch für alle Substanzen das 2*H*-Pyrazol-Tautomer gezeigt [88]. Dies ist für die Analyse der Kristallstruktur nicht entscheidend, da die Protonen nicht enthalten sind. Für spätere Optimierungsschritte müssen aber die richtigen Protonierungen verwendet werden. Der Ligand von 3G7E zeigt eine ähnliche Ausdehnung wie Novobiocin in seiner Kristallstruktur (s. Abbildung 2.9). Allerdings sind die einzelnen Bereiche unterschiedlich stark ausgeprägt. Der Ligand bindet auch in Richtung des Ausgangs der Bindetasche, ragt jedoch nicht aus der Tasche hinaus. Dafür bildet sich eine neue Seitentasche tief in der Bindetasche aus. Im Ribose-/Phosphatbereich der Bindetasche befindet sich in dieser Kristallstruktur ein Piperidinrest.

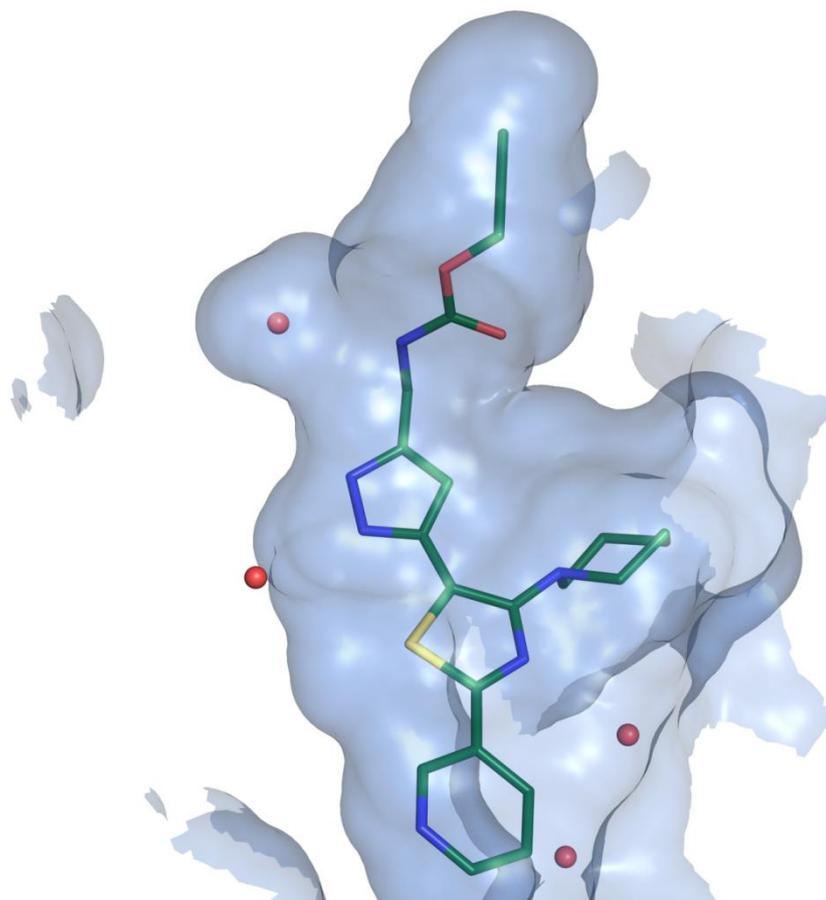


Abbildung 2.9 – Übersicht der Bindetasche der Kristallstruktur 3G7E.

^v <http://www.rcsb.org/pdb/explore/explore.do?structureId=3g7e>, zuletzt aufgerufen: Mai 2015

In Abbildung 2.10 ist der obere Bindungsbereich der Kristallstruktur dargestellt. Ile78 besitzt eine gute Distanz für einen hydrophoben Kontakt mit dem Pyrazolring des Liganden. Das untere Wassermolekül zeigt, wie in den anderen Strukturen, ein sehr enges Interaktionsnetz mit dem Liganden und dem Protein. Asp73, Gly77, Thr165 und der Inhibitor weisen alle sehr gute Entfernungen für Wasserstoffbrücken auf (Distanz: 2.8 Å). Weiterhin gibt es auch wieder die direkte Bindung vom Asp73 zum Liganden (2.7 Å). In der tieferen Bindungsregion gibt es jedoch deutliche Unterschiede zu den bisher gezeigten Kristallstrukturen. Das Stickstoffatom der Carbamat-Gruppe befindet sich in einer optimalen Entfernung für eine Wasserstoffbrückenbindung zu einem weiteren Wassermolekül (2.8 Å). Dieses Wasser befindet sich auch in einer guten Entfernung für eine Wasserstoffbrücke zum Backbone-Carbonyl von Val71. Die tiefe, neu ausgebildete Tasche, wird von hydrophoben Aminosäuren dominiert. Val167, Leu130 und Met95 zeigen gute Entfernungen für hydrophobe Interaktionen mit dem dort liegenden Ethinyl-Rest des Liganden (ca. 3.8 Å). Der Stickstoff des Amids von Asn46 besitzt mit 3.1 Å einen guten Abstand für eine Wasserstoffbrückenbindung zum Carbonylsauerstoff der Carbamoyl-Gruppe.

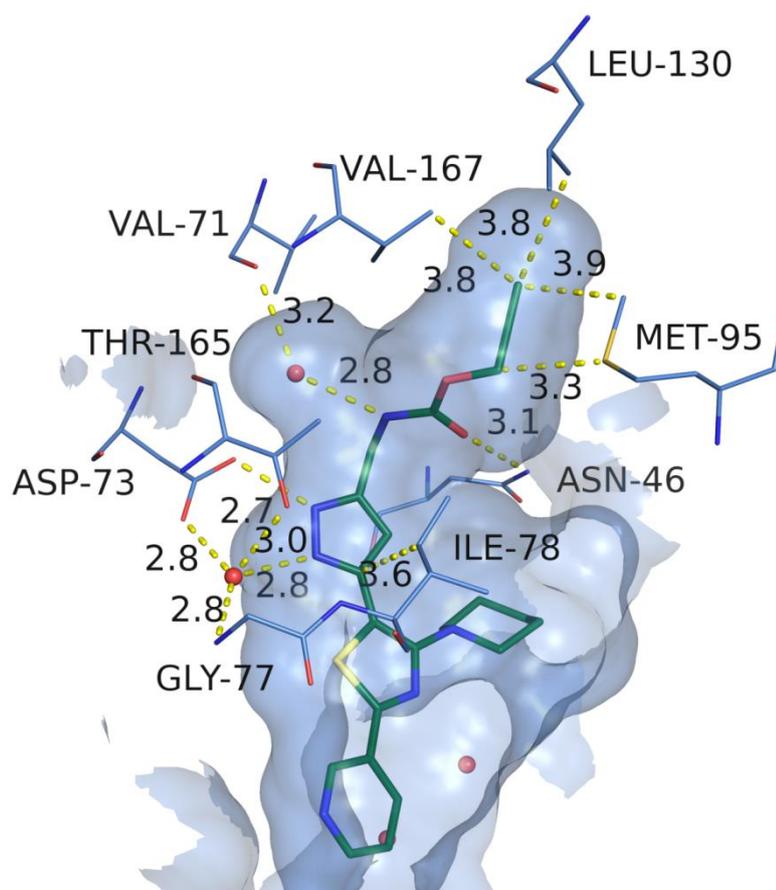


Abbildung 2.10 – Tiefer Bindungsbereich der Kristallstruktur 3G7E.

Zusätzlich gibt es zahlreiche Interaktionen mit dem nach außen zeigenden Pyridinrest (s. Abbildung 2.11). Arg76 befindet sich parallel zum Pyridin in einer Entfernung von 3.7 Å, was eine gute π -Kationen-Wechselwirkung darstellt. Mit dem Stickstoff des Pyridins kann Arg136 möglicherweise mit der Beteiligung von Asp106 Wasserstoffbrückenbindungen ausbilden. Phe104 bildet über sein Backbone-Carbonyl und ein Wassermolekül eine Wasserstoffbrücke zum Thiazol-Stickstoff des Liganden aus.

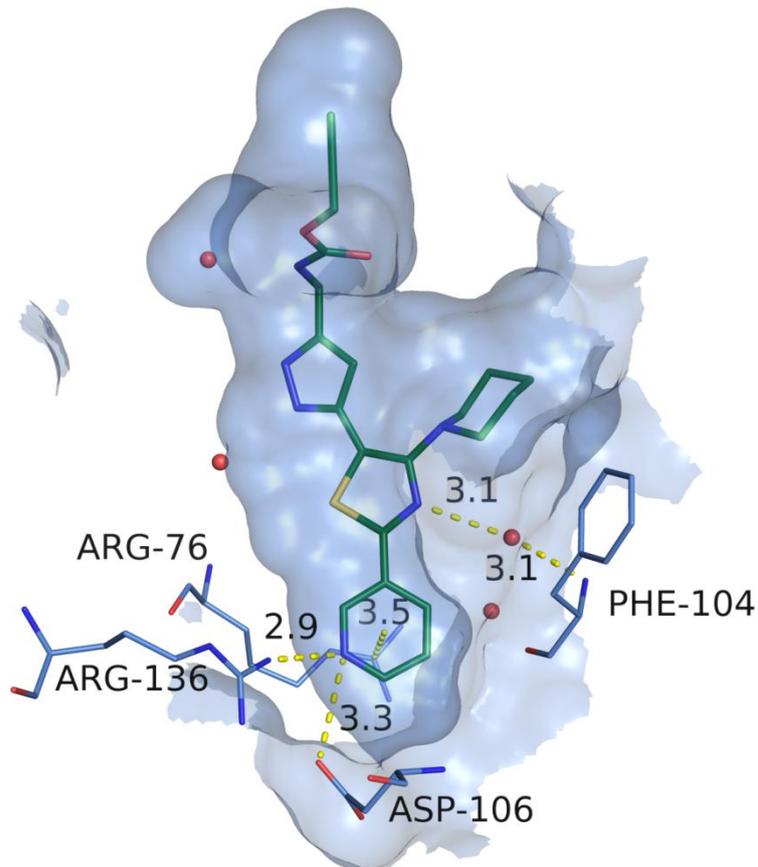


Abbildung 2.11 – Ausgangsbereich der Kristallstruktur 3G7E.

Im Vergleich aller Kristallstrukturen werden sowohl die Gemeinsamkeiten als auch Unterschiede deutlich. Zur Untersuchung der Bindetasche wurden die einzelnen Bereiche wie folgt definiert (s. Abbildung 2.12): Der zentrale Bereich der Bindetasche, bei dem alle Liganden eine Interaktion mit dem Asp73 und dem dort vorhandenen Wassermolekül durchführen, wurde als Hauptpharmakophor (HP) angesehen. Über diesem zentralen Interaktionspunkt befindet sich das Wat B (s. Abschnitt 2.3.2 auf S. 50). Dieser Bereich wurde als W definiert. Die tiefe Seitentasche, die bei 3G7E sehr dominant durch den Ethinyl-Rest vom Liganden ausgebildet wird, wird im Folgenden E-Bereich genannt. Der sich in Abbildung 2.12 nach rechts ausbildende Bereich, die Bindungsstelle des Phosphatrests bei 1E11, ist der P-Bereich. Schließlich wurde der Ausgangsbereich als A bezeichnet.

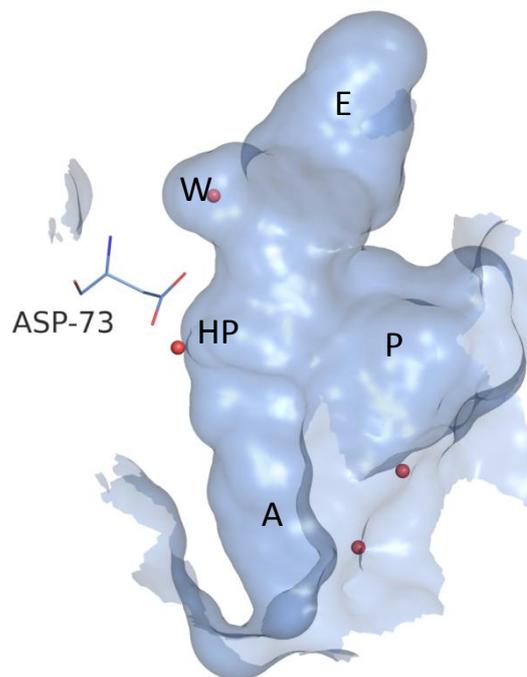


Abbildung 2.12 – Definition der Bindetaschenbereiche anhand der Oberfläche von 3G7E.

Die größte Übereinstimmung gab es in der Bindetasche im HP-Bereich. Bei allen Strukturen wurde das gleiche Interaktionsmuster unter Beteiligung der gleichen Aminosäuren und Wassermoleküle gefunden. Dies trat sogar speziesübergreifend auf. Das dort kokristallisierte Wassermolekül war sehr stark in das Wasserstoffbrückennetzwerk eingebunden. Im B-Bereich kam in allen Strukturen ebenfalls ein Wassermolekül vor, das lediglich bei 1E11 leicht versetzt war. Im Vergleich der verschiedenen Strukturen trat aber nur bei 3G7E eine gute Interaktion mit diesem Wassermolekül auf, da sowohl ANP als auch Novobiocin nicht sehr tief in diesem Bereich banden. Der E-Bereich war sehr stark in 3G7E ausgebildet. In den beiden anderen Strukturen war dieser Bereich zwar deutlich schwächer ausgeprägt, dennoch waren in den verschiedenen Strukturen noch unterschiedliche Wassermoleküle vorhanden. Dies war insofern auffällig, da in diesem Bereich der Anteil an hydrophoben Aminosäuren überwog. Der P-Bereich wurde vornehmlich von ANP in 1E11 adressiert. In 3G7E wies der Ligand dort einen vergleichsweise hydrophoben (möglicherweise sogar positiv geladenen) Piperidinrest auf. Novobiocin zeigte dort keine Interaktion. Im A-Bereich wiederum wies ANP in 1E11 keine Interaktion auf. Der Ligand von 3G7E baute dort ein gutes Interaktionsmuster mittels eines Pyridins auf. Novobiocin zeigte demgegenüber eine sehr ausgeprägte Bindung bis zum Loop A.

Damit wiesen alle Strukturen sehr unterschiedliche Charakteristika auf, die während eines Screeningvorgangs verwendet werden konnten, um die Flexibilität des Proteins zu erfassen (s. Abschnitt 4 auf

S. 79). Bevor die Strukturen jedoch für ein Screening verwendet werden konnten, mussten sie noch vorbereitet werden:

1. Vorbereitung der Strukturen: Keiner der verwendeten Strukturen wies eine Auflösung auf, die hoch genug war, um Wasserstoffatome zu lokalisieren. Daher mussten diese für alle Strukturen hinzugefügt werden. Bei 3G7E musste noch zusätzlich ermittelt werden, welches Tautomer am Pyrazol vorlag (s. o.). Bei 1AJ6 mussten eine Aminosäure zurückmutiert und die beiden flexiblen Loops in der Nähe der Bindetasche modelliert werden.
2. In allen Strukturen waren Wassermoleküle vorhanden. Für ein späteres Screening war es von elementarer Bedeutung vorherzusagen, welche Wassermoleküle essentiell für die Ligandenbindung sind.

2.2 Vorbereitung der Röntgenkristallstrukturen

2.2.1 Material und Methoden

2.2.1.1 Structure Prediction Tool

Für die Strukturvorhersage eignet sich das *Structure Prediction Tool* von Prime [102]. Die bei der Optimierung benötigten Energien werden mithilfe des OPLS_2005-Kraftfelds [103] und Surface Generalized Born Lösungsmodells [104] berechnet.

Zunächst werden alle Backbone-Atome und konservierten Aminosäuren von der Vorlage in die Zielstruktur kopiert. Anschließend werden die nicht-konservierten Aminosäuren aufgebaut. Dafür wird eine Bibliothek verwendet, die nicht-redundante Rotamere von Seitenketten enthält. Das erste Rotamer, welches keinen Clash verursacht, wird weiterverwendet. In einem ähnlichen Ansatz werden auch fehlende oder veränderte Backbone-Atome mithilfe einer Torsionswinkel-Bibliothek aufgebaut. Für größere Lücken werden Loop-Konformationen gesucht, die diese Lücke schließen können. Es findet jedoch keine weitere Optimierung für den Loop statt, so dass noch ein *Loop Refinement* (s. Abschnitt 2.2.1.2) nötig ist. Abschließend können zumindest die Seitenketten nochmal optimiert und minimiert werden.

2.2.1.2 Loop Refinement

Für das *Loop Refinement* (engl. für „Schleifenverfeinerung“) verwendet Prime [102] einen hierarchischen Ansatz [105–107]: Zunächst wird der gesamte konformere Raum des Loops mithilfe einer auf Dihedralwinkel basierenden Methode aufgebaut. Da die Anzahl der möglichen Loops mit der Länge sehr stark zunimmt, gibt es in Abhängigkeit der Länge verschiedene *Sampling*-Methoden: *Default Sampling* (≤ 5 Aminosäuren), *Extended Sampling* (6 – 11) und *Ultra Extended Sampling* (≥ 12). Diese

sorgen für eine angemessene Anzahl an Loops, die weiterverarbeitet werden können [106]. Die Platzierung der *Backbone*-Atome kann durch falschplatzierte Seitenketten anderer Aminosäuren blockiert werden. Daher steigt die Genauigkeit der Vorhersage für die Loops durch das Zulassen von vdW-Überschneidungen [107]. Die Anzahl der Loops wird im Anschluss mit iterativen Schritten durch Clustering-Methoden reduziert. Für die repräsentativen Clustermitglieder findet in der Folge eine intensive Seitenkettenoptimierung statt. Im letzten Schritt werden die Loops minimiert und ihre Energie berechnet. Die Energieberechnung wird bei jedem Schritt mithilfe des OPLS-AA Kraftfeld [108] und dem Surface Generalized Born Lösungsmodell [104] durchgeführt.

2.2.1.3 Molecular Dynamics Simulationen

Molecular Dynamics (MD, engl. für Moleküldynamik)-Simulationen ermöglichen die Untersuchung von biologischen und chemischen Systemen auf atomarer Ebene über die Zeit. Die Zusammenhänge und Berechnungen sind sehr komplex. Hier soll nur ein grober Überblick gegeben werden. Für weitergehende Information wird auf die Literatur verwiesen [109, 110].

Im Rahmen dieser Arbeit wurden kraftfeldbasierte MD-Simulationen in einem expliziten Lösungsmittel durchgeführt, d. h. ein Protein wird in eine Box, die aus Wassermolekülen besteht, platziert und die Bewegung aller Atome wird unter Berücksichtigung der physikalischen Gesetze mit Kraftfeldern berechnet. Kraftfelder beinhalten Funktionen und die dazugehörigen Parameter, um die potentielle Energie eines Systems zu berechnen. Bei so genannten *all atom* Kraftfeldern, sind die Funktionen und Parameter für jedes einzelne Atom definiert.

$$V(r) = \sum_{\text{Bind}} K_b (b - b_0)^2 + \sum_{\text{Winkel}} K_\theta (\theta - \theta_0)^2$$

Formel 2.1

$$+ \sum_{\text{Torsion}} \frac{V_n}{2} (1 + \cos[n\varphi - \delta]) + \sum_{\text{keinBind } ij} \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{12}} + \frac{q_i q_j}{r_{ij}}$$

Formel 2.1 zeigt die fundamentale Gleichung, auf der das Kraftfeld von AMBER11 [111] basiert. Die potentielle Gesamtenergie des Systems ($V(r)$) wird über die Aufsummierung von Bindungs-, Winkel-, Torsionswinkel- und nicht-Bindungsenergien bestimmt. Alle Berechnungen finden dabei auf atomarer Ebene statt, d. h. in Form des klassischen Kugel-Feder-Modells für Moleküle. Dies stellt eine elementare Vereinfachung dar, da quantenmechanische Berechnungen bei einer großen Anzahl an Wechselwirkungen zu anspruchsvoll wären.

Während einer MD-Simulation wird mithilfe des Kraftfeldes – mit einigen Vereinfachungen – für jedes Atom die Kraft, die auf das Atom wirkt, berechnet. Unter Berücksichtigung der Newtonschen Bewegungsgleichung kann damit die Bewegung für jedes einzelne Atom berechnet werden. Da die Anzahl an Wassermolekülen sehr hoch ist und die einzelne Bewegung für jedes Atom sehr aufwendig

zu berechnen wäre, werden häufig vereinfachte Wasser-Modelle verwendet. Weiterhin werden die freien Bindungsschwingungen von Wasserstoffbindungen eingeschränkt (SHAKE-Algorithmus [112]). Dies ermöglicht während der Simulation größerer Zeitschritte und beschleunigt die Berechnung.

Ein weiterer wichtiger Einflussfaktor auf die atomare Bewegung ist die Temperatur. Diese wird durch ein so genanntes Thermostat reguliert, wodurch die Energie des Systems erhöht wird und die Bewegung erheblich zunimmt. Zusätzlich muss ein äußerer Druck vorhanden sein, der verhindert, dass das gesamte System auseinander diffundiert. In diesem Zustand wäre die geschlossene Wasser-Box für sich stabil.

Für eine Simulation muss jedoch noch bestimmt werden, welche Kräfte auf die Atome am Rand des Systems wirken. Dafür wird von dem System in jeder Raumebene ein Spiegelbild von sich selbst erzeugt (Periodische Randbedingungen, PBC (engl. *Periodic boundary conditions*)). Überschreitet ein Atom diese Grenze wird es, wie bei einem Spiegelbild, auf der gegenüberliegenden Seite wieder eingefügt. Unter Berücksichtigung einiger Einschränkungen (ein Atom darf bspw. nicht mit sich selbst interagieren) und der weitreichenden elektrostatischen Interaktionen kann so eine freie MD-Simulation stattfinden.

Im Rahmen dieser Arbeit wurden MD-Simulationen mit AMBER11 [111] durchgeführt. Als Kraftfeld wurde für die Ligandenatome das „general Amber force field“ (GAFF) [113] und für die Proteinatome das ff03SB Kraftfeld [114] verwendet. Das genaue Vorgehen ist im Experimentellen Teil 11.3.4 auf S. 221 beschrieben.

2.2.2 Röntgenkristallstrukturen

– 1EI1

Beim ANP-Derivat wurde davon ausgegangen, dass die Phosphate deprotoniert sind. Daher war der Ligand vierfach negativ geladen. Es war nicht notwendig, weitere Anpassungen während der Standardprozedur (s. Experimenteller Teil 11.3.2 auf S. 220) durchzuführen.

– 3G7E

Für die Struktur von 3G7E musste zunächst geklärt werden, welches Tautomer des Liganden (**31**) vorlag. Der angegebene IUPAC-Name der PDB-Struktur (*1H*-Tautomer) stimmte nicht mit den dargestellten Strukturen der Veröffentlichung überein (*2H*-Tautomer, s. Abbildung 2.13). Zunächst wurde für beide Tautomere die Standardprozedur durchgeführt. Das jeweils erstellte Wasserstoffbrückennetzwerk ist in Abbildung 2.14 dargestellt. Bei beiden optimierten Strukturen waren beide Wasser in einem sehr engen Wasserstoffbrückennetzwerk eingebunden. Die beiden Strukturen unterschieden sich auf der Seite des Proteins darin, dass das Thr165 beim *1H*-Tautomer (in Abbildung 2.14 links) ein Wasserstoffbrückenbindungsakzeptor war, während das Threonin beim *2H*-Tautomer (rechts) als Wasserstoffbrückenbindungsdonor agierte. Das Thr165 kann unabhängig von der Orientierung keine weitere Interaktion mit einer weiteren Aminosäure eingehen. Das Proton der Alkoholfunktion beim *1H*-Tautomer oder das Elektronenpaar beim anderen Tautomer zeigt in Richtung der Seitenkette von Ile78. Bei der zweiten wichtigen Interaktion des Asp73 mit dem Liganden war der Unterschied deutlicher: Da Asp73 beim *1H*-Tautomer (links) als Wasserstoffbrückenbindungsdonor diente, wurde es während der Optimierung protoniert. Das *2H*-Tautomer interagiert mit der deprotonierten Asparaginsäure.

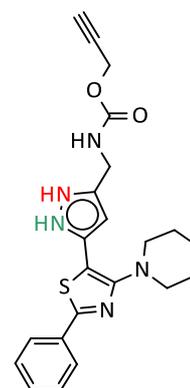


Abbildung 2.13 – Beide möglichen Tautomere des Liganden (**31**) von 3G7E. Das *1H*-Tautomer ist grün, das *2H*-Tautomer rot dargestellt. Für jedes Tautomer ist der jeweils andere Stickstoff unprotoniert.

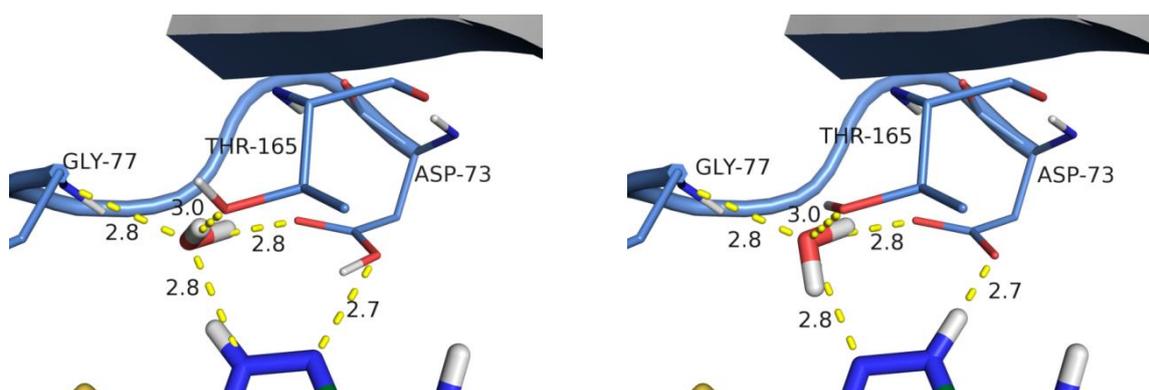


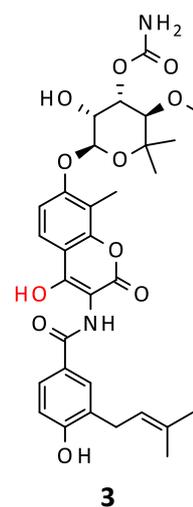
Abbildung 2.14 – Vergleich der Wasserstoffbrückennetze der beiden möglichen Tautomere des Pyrazolderivats von 3G7E. Links: Bei der Struktur mit dem *1H*-Tautomer ist das Wassermolekül in vier Wasserstoffbrückenbindungen involviert. Das Asp73 ist protoniert und agiert als Wasserstoffbrückenbindungsdonor gegenüber dem Pyrazol. Rechts: Beim *2H*-Tautomer ist das Wassermolekül ebenfalls an vier Wasserstoffbrückenbindungen beteiligt. Das Asp73 ist deprotoniert und in diesem Fall ein Wasserstoffbrückenbindungsakzeptor.

Insgesamt war die Protonierung des *2H*-Tautomers damit wahrscheinlicher. Zum einen wurde dieses in der Veröffentlichung von RONKIN *et al.* dargestellt [88]. Zum anderen ist die Protonierung der Seitenkette von der Asparaginsäure unter physiologischen Bedingungen bzw. bei neutralem pH-Wert ein seltener beobachtetes Phänomen, auch wenn es bekannte Ausnahmen gibt [115]. In allen ande-

ren bekannten Kristallstrukturen der GyrB-Untereinheit ist die Asparaginsäure jedoch deprotoniert und die Wechselwirkung zum Wasser findet auf der Seite des Liganden immer als Wasserstoffbrückenbindungsdonor statt. Aus chemischer Sicht sind beide Tautomere für den Liganden plausibel, so dass sich das Tautomer ausbilden wird, welches bei der Bindung energetisch günstigere Eigenschaften aufweist. Aus diesen Gründen wurde ausschließlich das *2H*-Tautomer weiterverwendet.

– 1AJ6

Bei der Kristallstruktur von 1AJ6 musste das His136 zu einem Arginin zurückmutiert und die beiden flexiblen Loops, mit den fehlenden Aminosäuren 83 – 87 und 105 – 111, rekonstruiert werden. Zunächst wurde jedoch die Standardprozedur zum Hinzufügen der Wasserstoffatome und der Optimierung des Wasserstoffbrückennetzwerkes durchgeführt. Bei dem Aminocumarin-Derivat im Novobiocin (**3**) handelt es sich um eine schwache Säure, weil sich die Alkoholfunktion des Cumarins in vinyloger Position zum Ester befindet (s. Abbildung 2.15). Der experimentell bestimmte pK_a -Wert beträgt 3.6 [116]. Da die Alkoholfunktion in Richtung der Lösung zeigt, ist Novobiocin (**3**) bei einem physiologischen pH-Wert höchstwahrscheinlich deprotoniert.



3
Abbildung 2.15 – Die azide Hydroxylgruppe von Novobiocin (**3**) ist rot markiert.

Zur Modellierung der nicht aufgelösten Bereiche in der Nähe der Bindetasche wurde nach einem ähnlichen Protokoll wie bei Saíz-URRA *et al.* verfahren [101]. In diesem führten sie ein strukturelles *Alignment* von 1AJ6 mit 1E11 und 1KIJ mithilfe des Dali-Servers [117] durch. Anschließend überführten Saíz-URRA *et al.* die Koordinaten für die Aminosäuren 77 – 90 (Loop A) von 1E11 und 100 – 117 (Loop B) von 1KIJ in 1AJ6. Bei 1KIJ handelte es sich um die Kristallstruktur von *T. thermophilus*. Daher mussten im nächsten Schritt alle Aminosäuren dieses Loops mutiert werden, die sich von *E. coli* unterschieden. Zusätzlich wurde das His136 zum Arginin zurückmutiert [101].

Um verlässlichere Konformationen für die Loops zu erhalten, wurde das Verfahren im Rahmen dieser Arbeit optimiert. Zunächst wurde das *Structure Prediction Tool* von Prime zur Positionierung der Aminosäuren verwendet: Dafür wurde ein *Sequenzalignment* von 1AJ6, 1E11 und 1KIJ durchgeführt. Für die Vorhersage von Loop A wurde die Aminosäuren 83 – 87 von 1E11 verwendet. Da sich die räumliche Struktur der beiden Kristallstrukturen in diesem Bereich sehr ähnelten, wurden ausschließlich die fehlenden

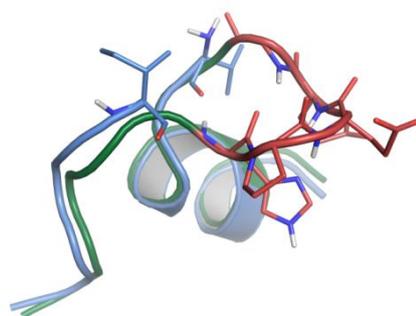


Abbildung 2.16 – Überlagerung der Struktur von 1AJ6 (blau) und 1E11 (grün) für Loop A. Die für die Vorhersage verwendeten Aminosäuren von 1E11 sind rot dargestellt.

Aminosäuren vorhergesagt (s. Abbildung 2.16). Mithilfe der Kristallstruktur von 1KIJ wurden Arg136 und die Aminosäuren 97 – 119 in 1AJ6 (entspricht den Aminosäuren 96 – 118 in *T. thermophilus*) bestimmt. Für erstere Aminosäure ist die Kristallstruktur von 1KIJ besser geeignet als 1E11, da das Arginin mit dem Novobiocin interagiert. Deshalb war es plausibel die Kristallstruktur von 1KIJ mit *T. thermophilus* zu verwenden, weil dort Novobiocin mit dieser Interaktion vorlag. Für Loop B wurden mehr Aminosäuren vorhergesagt, als in der Kristallstruktur von 1AJ6 fehlten. Dies lag in der schlechten Auflösung und der Größe des zu modellierenden Bereichs begründet (s. Abbildung 2.18). Daher wurde der komplette Loop neu vorhergesagt. Von den 23 Aminosäuren waren 17 identisch und zwei Paare ähnlich zueinander (s. Abbildung 2.17).

Nr. (<i>E. Coli</i>)	97	100																					110
<i>E. Coli</i>	V	L	H	A	G	G	K	F	D	D	N	S	Y	K	V	S	G	G	L	H	G	V	G
<i>T. thermophilus</i>	T	L	H	S	G	G	K	F	E	Q	G	A	Y	K	V	S	G	G	L	H	G	V	G

Abbildung 2.17 – Sequenzalignment von *E. Coli* und *T. thermophilus* für die Strukturvorhersage von Loop B. Identische Aminosäuren sind fett, ähnliche Amine sind kursiv dargestellt.

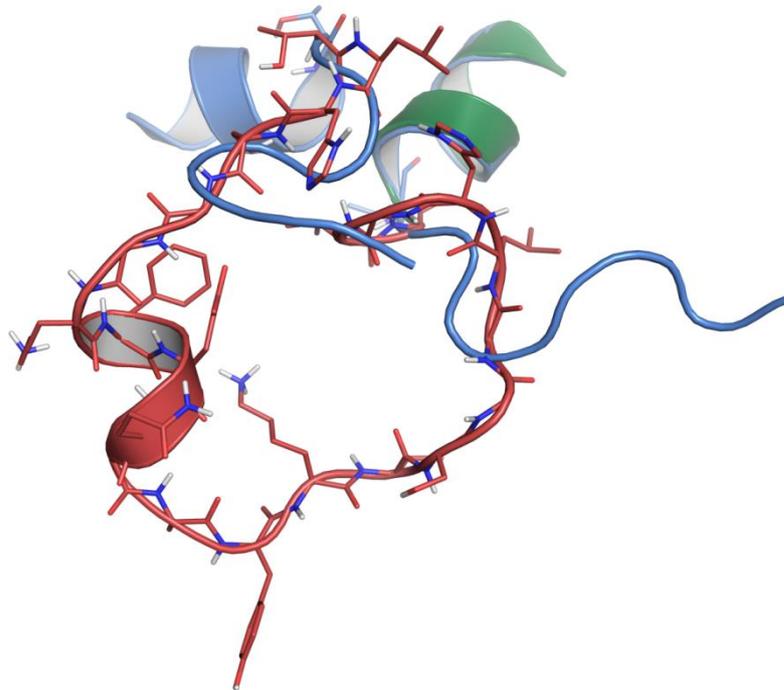


Abbildung 2.18 – Überlagerung der Struktur von 1AJ6 (blau) und 1KIJ (grün) für Loop B. Die für die Vorhersage verwendeten Aminosäuren von 1KIJ sind rot dargestellt.

Im nächsten Schritt wurde für die beiden vorhergesagten Loops ein *Loop Refinement* mit Prime (s. Experimenteller Teil 11.3.3 auf S. 220) durchgeführt. Bei beiden Loops war die energiegunstigste Konformation akzeptabel. Zum Abschluss wurde mit der erhaltenen Struktur eine MD-Simulation über 1 ns durchgeführt (s. Experimenteller Teil 11.3.4 auf S. 221). Die MD-Simulation war im gesamten Verlauf stabil (s. Abbildung 2.19). Der letzte *Snapshot* der MD-Simulation wurde schließlich minimiert (s. Experimenteller Teil 11.3.5 auf S. 224) und diente als Ausgangsstruktur für alle weiteren Untersuchungen. Abbildung 2.20 fasst alle Anpassungen zusammen und zeigt die final erhaltene Struktur.

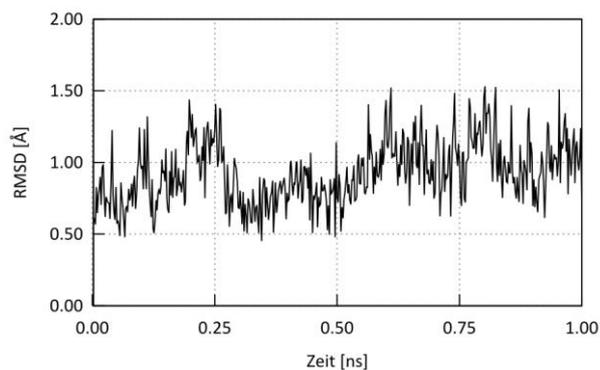


Abbildung 2.19 – RMSD-Plot vom Novobiocin zur Pose der Kristallstruktur. Die Pose blieb stabil und während der gesamten Simulation unter einem RMSD-Wert von 2 Å.

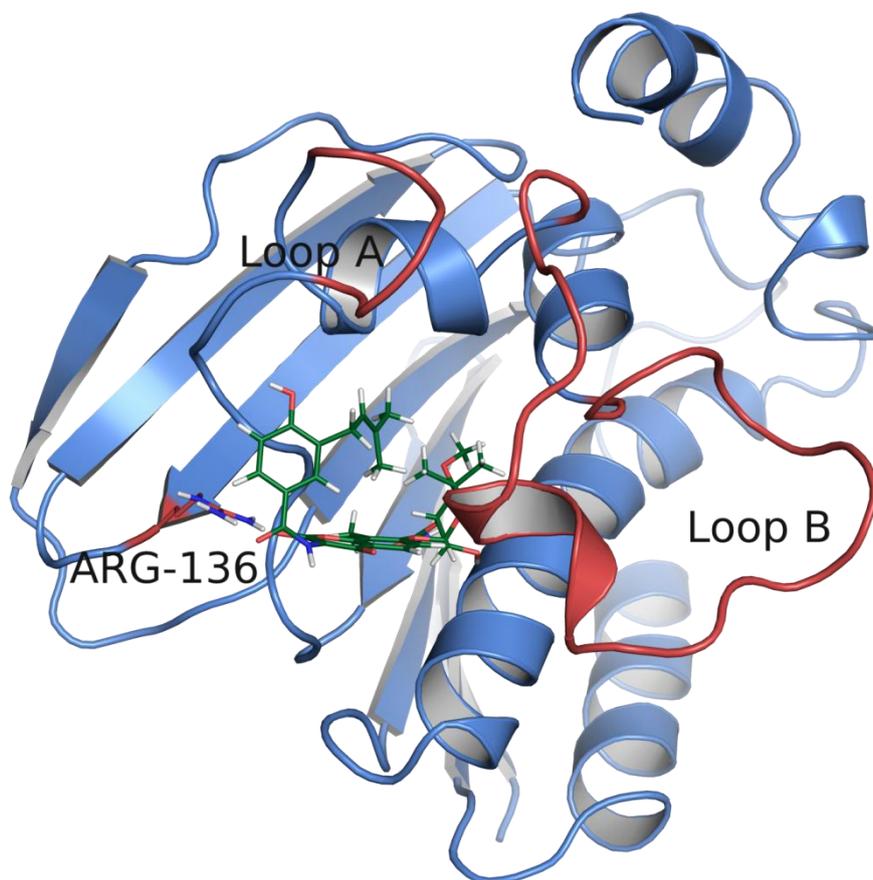


Abbildung 2.20 – Final verwendete Struktur von 1AJ6. Alle veränderten oder generierten Abschnitte sind rot markiert. Zur Orientierung ist Novobiocin (grün) ebenfalls dargestellt.

Damit lagen drei verschiedene Strukturen vor, von denen jede Struktur einzigartige Merkmale aufwies, die den vielfältigen konformativen Raum des Enzyms im Bereich der Bindetasche beschrieben und die für ein HTVS geeignet waren.

2.3 Analyse der Bedeutung der Bindetaschen Wassermoleküle

Wassermoleküle spielen für die Bindung von Liganden in einer Bindetasche eine wichtige Rolle [69]. Aufgrund der Wasserstoffbrückenbindungsdonor- und Wasserstoffbrückenbindungsakzeptoreigenschaften eignet sich Wasser sehr gut für hydrophile Interaktionen mit dem Liganden und dem Protein. Dieser enthalpische Gewinn für die Bindung muss jedoch gegen den entropischen Verlust durch die Einschränkung der Beweglichkeit des Wassermoleküls aufgewogen werden [118]. Weiterhin können Wassermoleküle für die strukturelle Stabilität des Proteins verantwortlich sein, bspw. zur Stabilisierung von Sekundärstrukturen [119]. Diese und weitere gegensätzlichen Einflüsse können dazu führen, dass die Verdrängung eines Wassermoleküls durch den Liganden entweder einen positiven, keinen oder sogar einen negativen Einfluss auf die Freie Bindungsenthalpie hat [118]. Diese ambivalenten Effekte wurden bereits ausführlich diskutiert und hier sei nur auf zwei aktuelle, umfassende Übersichtsartikel über die Eigenschaften des Wassers von BALL [119] und die thermodynamischen Einflüsse auf die Bindung der Liganden von KLEBE [120] verwiesen.

– Evaluierung von Wassermolekülen

Aufgrund der komplexen Zusammenhänge, wie Wassermolekülen die Freie Bindungsenthalpie beeinflussen können, ist die Bewertung sehr aufwendig. Es gibt zahlreiche Tools und Methoden, die Wassermoleküle und deren Einfluss bestimmen können. Dabei werden sehr unterschiedliche Ansätze verfolgt: von geometriebasierten über empirischen (bspw. B-Faktor bei Röntgenkristallstruktur), wissensbasierten, kraftfeldbasierten (MM-GB/PBSA) bis hin zu quantenmechanischen Berechnungen. Dabei ist der rechnerische Aufwand für die verschiedenen Methoden sehr unterschiedlich. Tools, die auf geometriebasierten, empirischen und/oder wissensbasierten Ansätzen beruhen, bewerten Wassermoleküle in wenigen Sekunden bis Minuten. Dazu gehören z. B.: Aquarius (Veröffentlichung: 1991) [121] Consolv (1997) [122], SuperStar (2001) [123], HINT und Rank (2008) [124], WaterClassifier (2012) [125], wPMF (2013) [126] oder EDIA (2015) [127]. Bei den aufwendigeren Methoden werden Monte Carlo-Simulationen (2007) [118], MD-Simulationen (2013) [128], MM-GB/SA-Berechnungen (2012) [129] oder, wie bei WaterMap (2008) [130], Kombinationen (MD-Simulationen, Wasser-Clustering, statistische Thermodynamik) verwendet, um die Freie Bindungsenthalpie der einzelnen Wassermoleküle zu berechnen und letztendlich ihren Einfluss zu bestimmen.

Für die Untersuchungen im Rahmen dieser Arbeit wurde der Fokus auf effiziente Methoden gelegt. Da schon einige Vorkenntnisse über die strukturellen Eigenschaften der Bindetasche bestanden (s. u.), wurden keine aufwendigen Energieberechnungen durchgeführt. Zunächst wurde die Umgebung der Wassermoleküle in der Bindetasche selbst untersucht. Zur Überprüfung der Stabilität der Komplexe MD-Simulationen (s. Abschnitt 2.2.1.3 auf S. 41) mit verschiedenen Wassermolekülen ana-

lysiert. Für die Bewertungen wurden schließlich SuperStar, WaterClassifier und HINT und Rank verwendet, die im Folgenden kurz erläutert werden.

2.3.1 Material und Methoden

2.3.1.1 Superstar

SuperStar verfolgt einen wissensbasierten Ansatz, um die „Neigung“ abzuschätzen, mit der eine spezifische Funktionalität (so genannte *Probe*, engl. für Sonde) in einer bestimmten chemischen Umgebung (hier Bindetasche) platziert wird [123, 131]. Die Wissensbasis für SuperStar stellte die IsoStar-Datenbank dar [132]. Diese Datenbank enthält Verteilungsdichten funktioneller Gruppe um eine andere, zentrale Gruppe. Um diese Verteilungsdichten zu bestimmen, wurden Kristallstrukturen von der Cambridge Structural Database (CSD) und der Protein Data Bank (PDB) ausgewertet [133]. SuperStar fragmentiert die Bindetasche in zentrale Gruppen und bestimmt die Dichteverteilung der entsprechenden Sonde für diese Gruppe. Anschließend wird die Bindetasche wieder zusammengesetzt und überlappende Verteilungen kombiniert. Die final entstandene Dichteverteilung gibt darüber Auskunft, wieviel häufiger die untersuchte Sonde an dieser Stelle platziert ist, als es bei einer zufälligen Verteilung zu erwarten wäre. Der Wert zwei bedeutet dann beispielsweise, dass die Sonde doppelt so häufig an dieser Stelle vorkommt wie bei einer zufälligen Verteilung. Damit wird im Grunde keine Interaktion oder deren Stärke bewertet, sondern nur die relative Häufigkeit in Kristallstrukturen, welche indirekt für eine höhere Stabilität und stärkere Interaktion steht [131]. VERDONK *et al.* konnten jedoch bereits zeigen, dass die Wahrscheinlichkeit der Platzierung von einer Wasser-Sauerstoff-*Probe* mit dem B-Faktor der dort vorhandenen Sauerstoffatomen der aufgelösten Wasser korreliert: Gut definierte Wassermoleküle mit niedrigem B-Faktor zeigen scharfe Peaks für die Wasser-Sauerstoff- Sonde und umgekehrt [123]. Es werden jedoch keine Angaben gemacht, wie ein scharfer Peak definiert wird. Das von GOODFORD *et al.* entwickelte Tool GRID analysiert ähnlich wie SuperStar, allerdings mit einem empirischen Ansatz, die Interaktionsmöglichkeiten einer Bindetasche mit Sonden [134]. Für GRID konnte gezeigt werden, dass es strukturell wichtigen Wassermoleküle in einem HIV-1-Protease-Komplex erfolgreich vorhersagen konnte [135]. Damit wurde der prinzipielle Nutzen dieser Form der Untersuchung weiter belegt.

2.3.1.2 WaterClassifier

WaterClassifier ist ein Tool, um die Wahrscheinlichkeit zu bestimmen, mit der ein Wassermolekül ersetzt werden kann. Zusätzlich bietet es die Möglichkeit abzuschätzen, mit welcher Wahrscheinlichkeit diese Stelle von einer hydrophilen oder hydrophoben Gruppe belegt wird [125]. Dafür entwickelten Ross *et al.* ein Modell, welches aus einem neugewichteten Wasserstoffbindungsterm der Scoring-Funktion von AutoDock Vina [136] und zwei selbst entwickelten Termen für die hydrophile und lipo-

phile Umgebung des Wassermoleküls bestand. Damit konnten sie in einem Testset mit 1,004 Wassermolekülen in verschiedenen Proteinstrukturen 75 % der Wasser korrekt als ersetzbar oder konserviert vorhersagen. Da WaterClassifier eine Wahrscheinlichkeit für die Substitution angibt, steigt die Vorhersagekraft mit steigendem Wert. Bei einem höheren Wahrscheinlichkeits-Cutoff von 0.8 sind beispielsweise 94 % aller Wassermoleküle korrekt als ersetzbar klassifiziert worden [125].

2.3.1.3 HINT und Rank

AMADASI *et al.* entwickelten ein Bewertungsschema mithilfe des HINT-Kraftfeldes und des geometrie-basierten Rank-Protokolls [124]. HINT ist ein nicht-kovalentes Kraftfeld. Es basiert auf $\text{LogP}_{\text{O/W}}$ -Werte und berechnet über diese die Interaktionsstärke zwischen zwei Molekülen [137]. Rank verwendet geometrische Eigenschaften, um Wassermoleküle zu evaluieren, indem es anhand der Länge und der Winkel der Wasserstoffbrückenbindungen bewertet wird. Je näher sich diese am Optimum befinden, desto besser wird das Wassermolekül beurteilt [138]. Mithilfe einer Statistikanalyse entwickelten AMADASI *et al.* ein Bewertungsschema, das den HINT-Score und die Rank-Bewertung zur Vorhersage der Konservierung von Wassermolekülen kombiniert. In ihrem Testset mit 68 Wassermolekülen konnten 56 (87 %) der Wassermoleküle erfolgreich als konserviert oder ersetzbar vorhergesagt werden. Wurden nur hochaufgelöste Kristallstrukturen (Auflösung $\leq 2.0 \text{ \AA}$) bewertet, stieg die Erfolgsrate auf 92 % [124]. Damit wird die Abhängigkeit von der Qualität der experimentellen Daten verdeutlicht, da auch die Möglichkeit besteht, dass die Wasser in der Röntgenkristallstruktur aufgrund der Auflösung nicht bestimmbar sind.

– Vergleich der Methoden

Die verschiedenen Methoden besitzen unterschiedliche Vor- und Nachteile. Die MD-Simulation eignet sich sehr gut, um Aussagen über die allgemeine Stabilität des Gesamtkomplexes und den Einfluss der Wassermoleküle auf diese zu beschreiben. Zeigen die Komplexe ohne die entsprechenden Wasser noch eine gute Stabilität, sind die Auswirkungen auf die Liganden und die Wichtigkeit der Wassermoleküle gering zu bewerten. Sind die Wassermoleküle aber für die Stabilität der Bindung notwendig, kann im Umkehrschluss nicht davon ausgegangen werden, dass diese nicht ersetzbar sind. Ein anderer Ligand könnte beispielsweise an dieser Position selbst eine Interaktion mit dem Protein ausbilden und deswegen eine stabile Bindung eingehen.

Bei SuperStar liegt der Vorteil darin, dass die Vorhersage und Evaluierung der Wassermoleküle allein strukturbasiert vorgenommen wird. Die Wasser müssen vorher nicht experimentell bestimmt worden sein und die Auswertung wird nicht durch einen definierten Liganden verzerrt. Dafür sind die Ergebnisse schwieriger zu interpretieren und es kann nur eine relative Aussage getroffen werden, da die Methode nicht explizit für die Analyse von Wassermolekülen konzipiert wurde: Die Vorhersage

für ein bestimmtes Wassermolekül in einer Kristallstruktur kann korrekt sein, da das Wassermolekül in der entsprechenden Kristallstruktur auch vorkommt. Allerdings trifft das jedoch keine direkte Aussage über die Ersetzbarkeit.

Anders ist dies bei dem empirisch entwickelten Tool WaterClassifier sowie der HINT und Rank-Methode. Diese sind wiederum sehr stark von den zur Erstellung verwendeten Datensätzen abhängig. Sind bestimmte Konfigurationen unterrepräsentiert, werden die Wassermoleküle falsch bewertet. Der Vorteil bei allen zur Evaluierung vorgestellten und verwendeten Tools bestand darin, dass diese schnelle Auswertungen ermöglichten.

2.3.2 Wassermoleküle in der ATP-Bindetasche der Gyrase

Für die ATPase der GyrB-Untereinheit zeigt sich die Bedeutung der Wassermoleküle schon anhand der Röntgenkristallstruktur von 1AJ6. Wie in Abschnitt 2.1 auf S. 27 beschrieben, handelt es sich bei 1AJ6 um eine Mutante, die eine Novobiocinresistenz vermittelt. Dabei geht durch den Austausch des Arg136 zu einem Histidin keine Interaktion zum Novobiocin verloren, da die Wasserstoffbrücke über ein Wassermolekül weiterhin auch zur kürzeren Histidin-Seitenkette vermittelt wird. Die Einbindung des Wassers ist jedoch entropisch ungünstig und lässt, wie oben beschrieben, die Dissoziationskonstante des Novobiocin-Protein-Komplexes auf das 40fache ansteigen. Dies konnte auch durch Berechnung der Freien Energie der Wassermoleküle nachgewiesen werden [139].

Während der Analyse der Kristallstrukturen (s. Abschnitt 2.1 auf S. 27) traten Wassermoleküle an vergleichbaren Positionen wiederholt auf. Diese Wassermoleküle befinden sich in vier verschiedenen Bereichen der Bindetasche, die alle auch in der Kristallstruktur von 1KIJ auftreten (s. Abbildung 2.21).

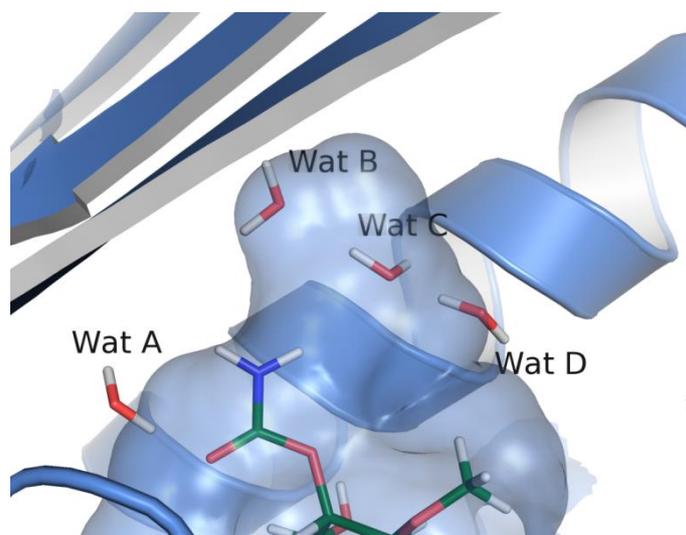


Abbildung 2.21 – Position und Bezeichnung der in allen Strukturen wiederkehrenden Wassermoleküle. Dargestellt ist die Kristallstruktur von *T. thermophilus* mit Novobiocin (grün) als Liganden (1KIJ). Wasserstoffe wurden nach der Standardprozedur (s. Experimenteller Teil 11.3.2 auf S. 220) hinzugefügt.

In der Kristallstruktur mit dem Pyrazolderivat von 3G7E waren zwei Wassermoleküle (Wat A und Wat B; s. Abbildung 2.21) enthalten. In der in dieser Arbeit verwendeten Struktur von 1AJ6 mit Novobiocin lagen die Wassermoleküle Wat A, Wat B und Wat C vor. Bei 1E11 mit dem ATP-Derivat ANP waren Wat A, Wat B und Wat D wiederzufinden. Außerdem wies eine erst 2014 veröffentlichte Röntgenkristallstruktur von STANGER *et al.* mit ADP und P_i (PDB-Code: 4PRX^{vi}) ebenfalls vier Wassermoleküle auf [100]. Diese befanden sich an den gleichen

Positionen wie bei 1KIJ (s. Abbildung 2.21).

Möglicherweise war die Auflösung bei 1E11 nicht ausreichend, um alle Sauerstoffatome der Wassermoleküle zu identifizieren. 1E11 besaß eine Auflösung von 2.3 Å, 4PRX 1.8 Å.

Tabelle 2.2 fasst nochmal das Auftreten aller Wassermoleküle zusammen.

Tabelle 2.2 – Wassermoleküle in den Bindetaschen der Kristallstrukturen. Vorhandene Wasser sind mit einem Haken (✓), nicht-vorhandene mit einem Kreuz (x) markiert.

Kristallstruktur	Wat A	Wat B	Wat C	Wat D
1E11	✓	✓	x ^a	✓
1AJ6	✓	✓	✓	x
3G7E	✓	✓	x	x

^a: Wassermolekül in einer ähnlichen Kristallstruktur mit besserer Auflösung enthalten.

Für Novobiocin (**3**) wurden schon weitergehende Untersuchungen für die Wassermoleküle durchgeführt, um zu bestimmen, welche wichtigen Einfluss auf die Bindungen haben, als essentiell gelten oder ersetzt werden können. Diese Untersuchungen wurden dadurch unterstützt, dass mit Clorobiocin (**39**) ein Derivat von Novobiocin (**3**) existiert, welches das Wassernetzwerk beeinflusst. Der entscheidende Unterschied zwischen Novobiocin (**3**) und Clorobiocin (**39**) besteht im Austausch des Stickstoffatoms des Carbamats mit einem Pyrrol-Derivat (s. Abbildung 2.22). Dadurch werden die beiden Wassermoleküle Wat B und Wat C, die in der Kristallstruktur von 1AJ6 vorkommen, verdrängt (s. Abbildung 2.23). Bemerkenswert dabei ist, dass die Positionen der

Wassermoleküle von hydrophoben Kohlenwasserstoffen besetzt sind. Enthalpisch betrachtet erhöht sich daher die Freie Bindungsenthalpie. Dennoch hat Clorobiocin (**39**) eine 37fach höhere Bindungsaffinität als Novobiocin (**3**). Dies führten Yu *et al.* auf den entropischen Gewinn durch die Freisetzung der zuvor geordneten Wassermoleküle zurück [139]. Für das Wat A gibt es, nach unserem Kenntnisstand, bisher keinen Liganden, der dieses Wasser verdrängen kann. Das Wasser ist sehr eng an die Struktur gebunden und befindet sich innerhalb der unter Standardbedingungen berechneten Pro-

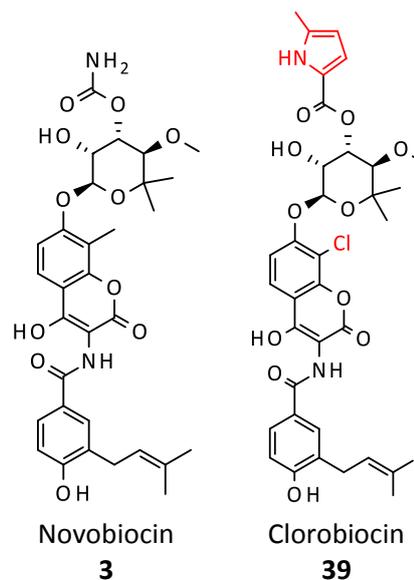


Abbildung 2.22 – Vergleich von Novobiocin (**3**, links) und Clorobiocin (**39**, rechts). Die Unterschiede beim Clorobiocin sind rot markiert.

^{vi} Da diese Röntgenkristallstruktur erst nach Durchführung des HTvS verfügbar war, konnte sie in diesen Untersuchungen nicht berücksichtigt werden.

teinoberfläche. Die ausgebildete Tasche für das Wat A ist daher auch sehr eng und tief. Daher erfordert es hohe sterische Anforderungen an den möglichen Liganden, sollte er dieses Wasser verdrängen. Möglicherweise ist Wat A auch mit an der strukturellen Integrität des Proteins beteiligt. Bei diversen Screening-Experimenten wurde dieses Wasser daher als konserviert angesehen [73, 76, 94, 140].

Damit kann in der Kristallstruktur von 1AJ6 zwischen verschiedenen Wassermolekülen und deren Bedeutung unterschieden werden: Wat B und Wat C sind ersetzbar, möglicherweise sogar durch einen hydrophoben Rest des Liganden. Bei Wat A handelt es sich um ein strukturelles Wassermolekül, das hochkonserviert ist. Diese Kategorisierung ist auch für die anderen Kristallstrukturen interessant. Da es jedoch, abgesehen von der Umgebung des Wat A, große strukturelle Unterschiede bei den verschiedenen Kristallstrukturen gab (s. Abschnitt 2.1 auf S. 27), war es erforderlich, die Wassermoleküle in den anderen beiden Strukturen genauer zu eva-

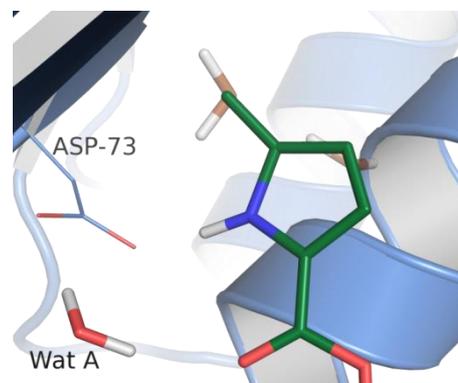


Abbildung 2.23 – Kristallstruktur von Clorobioicin in GyrB (PDB-code: 1KZN). Die ersetzten Wasser von 1AJ6 sind transparent dargestellt. Alle nicht-polaren Wasserstoffe sind ausgeblendet.

luieren. Dafür wurden zunächst die einzelnen Wassermoleküle analysiert und anschließend bewertet.

– Wat A

Wat A kommt in allen Kristallstrukturen vor. Nach einem *Alignment* der *Backbones* der Proteine beträgt die Abweichung der Positionen zwischen den verschiedenen Strukturen für das Sauerstoffatom nur 0,6 – 0,7 Å (s. Abbildung 2.24). Es sind immer die gleichen Aminosäuren an der Bindung beteiligt (auch speziesübergreifend; s. Abschnitt 2.1 auf S. 27): Asp73, Gly77, Thr165 (*E. coli*-Nummerierung). Eine Bindung findet zusätzlich zu den Liganden statt. Die ausgebildeten Wasserstoffbrücken sind teilweise sehr kurz (bis zu 2.5 Å). Aus den experimentellen Daten des Novobiocins ist schon bekannt, dass Wat A sehr wahrscheinlich nicht ersetzbar ist (s. o.). Da sich die verschiedenen Proteinstrukturen in diesem Bereich nur wenig unterscheiden, ist dieses Wasser voraussichtlich in allen Proteinen essentiell.

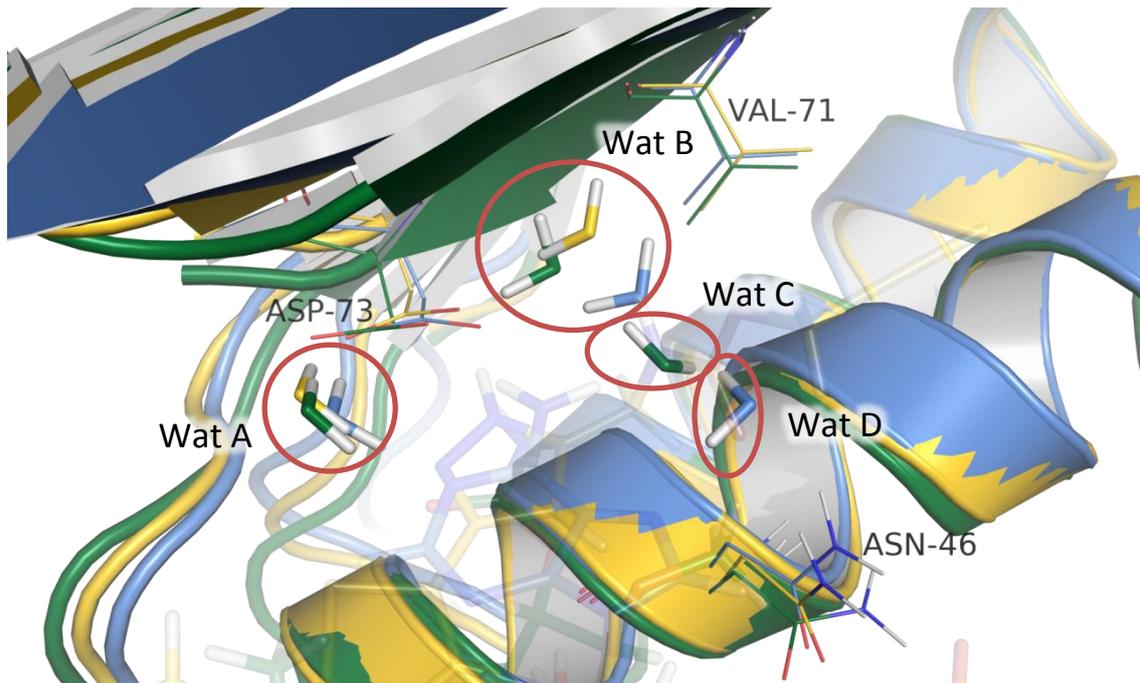


Abbildung 2.24 – Überlagerung der verschiedenen vorbereiteten Kristallstrukturen und deren Wassermoleküle. Dargestellt sind 1E11 (blau), 1AJ6 (grün) und 3G7E (gelb). Die Sauerstoffatome der Wasser sind entsprechend der zur Kristallstruktur zugehörigen Kohlenstoffatome gefärbt. Zur Orientierung sind die Liganden transparent dargestellt.

– Wat B

Auch Wat B lässt sich in allen Strukturen wiederfinden. Es befindet sich in einer Seitentasche, die in keiner Struktur von einem Liganden besetzt wird, aber deutlich größer ist als beim Wat A. Die Position des Wat B ist insgesamt variabler als das Wat A. Nach dem *Alignment* über den Backbone sind die Sauerstoffatome von den Wassern der verschiedenen Strukturen zwischen 0.8 und 1.7 Å zueinander versetzt, damit aber noch immer näher beieinander platziert als die Auflösung der einzelnen Kristallstrukturen groß ist. Insbesondere Wat B von 1E11 (blau in Abbildung 2.24) ist leicht verschoben. Während der Analyse mit SuperStar (s. auch Anhang 12.2 auf S. 230) konnte an der Position des Wassers auch kein Maximum gefunden werden (s. Abbildung 2.25). Retrospektiv konnte jedoch der Vergleich mit 4PRX zeigen, dass die Vorhersage von SuperStar mit dieser Struktur, die auch eine bessere Auflösung als 1E11 aufwies, übereinstimmte. Bei allen Strukturen ist Asp73 direkt an der Bindung beteiligt, welche aufgrund des ionischen Charakters die stärkste Interaktion darstellt. Dies ist die einzige Wasserstoffbrückenbindung, die Wat B in 1E11 mit dem Protein oder dem Liganden eingehen kann. In der Struktur von 1AJ6 ist noch eine Interaktion mit dem Backbone-

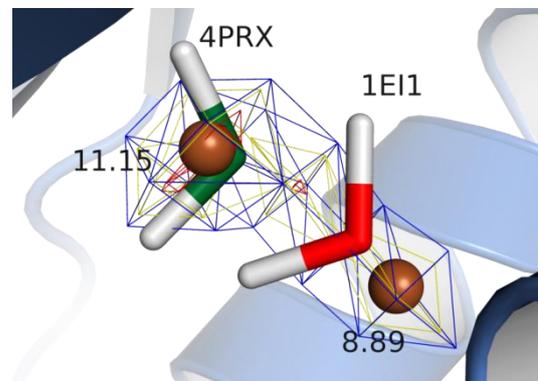


Abbildung 2.25 – Vergleich der Wat B-Position von 1E11 und 4PRX. Das Sauerstoffatom von 1E11 ist rot, 4PRX grün dargestellt. Die Kontur zeigt die Häufigkeitsanalyse von SuperStar (1E11): blau > 2, gelb > 4, rot > 8. Die braunen Sphären markieren die Position der Maxima. Die Zahlen zeigen die Werte der Maxima.

Carbonylsauerstoff von Thr165 (3.0 Å) möglich. Bei 3G7E befindet sich das Wat B einem deutlich engeren Interaktionsnetzwerk: In 3.2 Å Entfernung befindet sich das Carbonyl des Val71 und 2.8 Å entfernt das Stickstoffatom des Carbamats vom Liganden. Durch Clorobiocin ist bekannt, dass dieses Wasser im Komplex mit Novobiocin von einem hydrophoben Rest ersetzt werden kann. Im Vergleich dazu ist insbesondere die Struktur von 3G7E interessant, da sich das Wassermolekül dort in einem deutlich engeren Interaktionsnetz befindet und näher am Protein gebunden ist. Im Unterschied zu Wat A treten bei diesem Wasser jedoch nur drei der vier möglichen Wasserstoffbrückenbindungen auf.

– Wat C und Wat D

Wat C tritt nur bei 1AJ6 auf. Wie in Abschnitt beschrieben, befindet sich dieses Wasser, ähnlich wie Wat D, in einer hydrophoben Umgebung. Zum Protein weist es nur eine Interaktion mit dem Sauerstoff des Backbones von Val43 auf. Eine weitere Interaktion mit dem Stickstoff der Carbamatstruktur des Novobiocins wäre bei einer Entfernung von 3.0 Å möglich, der Winkel für die Wasserstoffbrückenbindung weicht aber mit 143 ° deutlich von den optimalen 180 ° ab. Sehr ähnlich verhält sich das Wat D in der Kristallstruktur von 1EI1. Es steht in einem Wasserstoffbrückennetzwerk mit dem ANP (Entfernung 2.8 Å) und Asn46 (3.0 Å). Bei diesen Interaktionen sind alle Winkel akzeptabel.

– Bewertung der Wassermoleküle

Zunächst erfolgt die Bewertung der MD-Simulationen, die einen allgemeinen Einblick über die Stabilität geben. Anschließend werden die anderen Tools gemeinsam ausgewertet.

Alle Protein-Ligand-Komplexe waren während der MD-Simulation instabil, wenn Wat A nicht vorhanden war. Die Liganden bewegten sich aus ihrem ursprünglichen Bindungsmodus heraus und versuchten neue Wasserstoffbrücken mit den Proteinen in diesem Bereich auszubilden. Bei 3G7E rutschte in jeder MD-Simulation ohne Wat A der Pyrazolkörper in die Wat A-Tasche hinein (s. Abbildung 2.26).

Bei 1E11 und 1AJ6 verschoben sich entweder die Liganden in den Bereich oder Wassermoleküle wanderten aus der Bindetasche oder von der

Lösung in die Wat A-Tasche hinein. Im

zweiten Fall verließen die Liganden zuvor die Bindetasche (s. Anhang 12.1 auf S. 229). In keiner Simulation konnte beobachtet werden, dass die Wat A-Tasche unbesetzt blieb. Dies betonte die strukturelle Wichtigkeit dieses Wassermoleküls

und die Bedeutung für die Stabilität der Komplexe. Bei der umgekehrten Simulation – das Wat A war vorhanden, es fehlten jedoch alle anderen Wasser der Bindetasche – verhielten sich die Komplexe

meist stabil. Dies konnte auch bei 3G7E, bei dem sich das Wat B im Vergleich zu

den anderen Kristallstrukturen in einem engeren Interaktionsnetz befand, beobachtet werden. Auch wenn das Wat B fehlte, verschob sich der Ligand nur wenig, so dass der RMSD (*Root-Mean-Square Deviation*) zur Kristallstruktur nach der MD-Simulation für den Liganden lediglich 0.81 Å betrug (s. Abbildung 2.26).

Bei der umgekehrten Simulation – das Wat A war vorhanden, es fehlten jedoch alle anderen Wasser der Bindetasche – verhielten sich die Komplexe meist stabil. Dies konnte auch bei 3G7E,

bei dem sich das Wat B im Vergleich zu den anderen Kristallstrukturen in einem engeren Interaktionsnetz befand, beobachtet werden. Auch wenn das Wat B fehlte, verschob sich der Ligand nur wenig, so dass der RMSD (*Root-Mean-Square Deviation*) zur Kristallstruktur nach der MD-Simulation für den Liganden lediglich 0.81 Å betrug (s. Abbildung 2.26).

Bei der umgekehrten Simulation – das Wat A war vorhanden, es fehlten jedoch alle anderen Wasser der Bindetasche – verhielten sich die Komplexe meist stabil. Dies konnte auch bei 3G7E,

bei dem sich das Wat B im Vergleich zu den anderen Kristallstrukturen in einem engeren Interaktionsnetz befand, beobachtet werden. Auch wenn das Wat B fehlte, verschob sich der Ligand nur wenig, so dass der RMSD (*Root-Mean-Square Deviation*) zur Kristallstruktur nach der MD-Simulation für den Liganden lediglich 0.81 Å betrug (s. Abbildung 2.26).

Bei der umgekehrten Simulation – das Wat A war vorhanden, es fehlten jedoch alle anderen Wasser der Bindetasche – verhielten sich die Komplexe meist stabil. Dies konnte auch bei 3G7E,

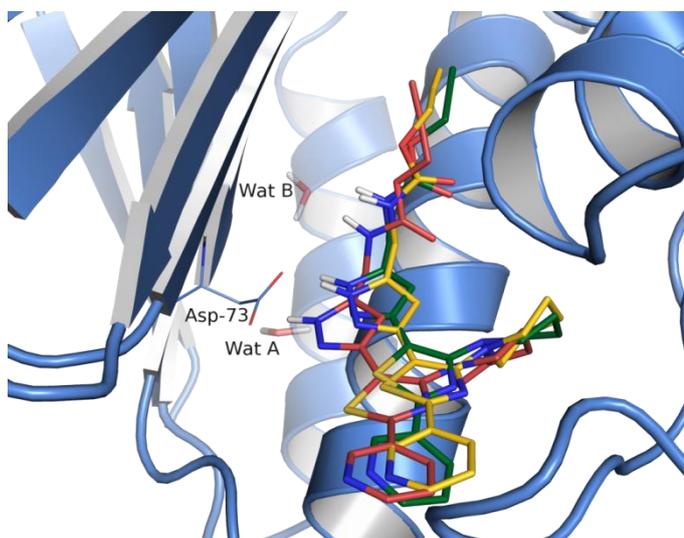


Abbildung 2.26 – Stabilität der MD-Simulationen von 3G7E mit verschiedenen Wassermolekülen. Die Kristallstruktur des Liganden von 3G7E ist grün dargestellt. Weiterhin sind der Ligand vom letzten Frame der MD-Simulation ohne Wat A (rot) und ohne Wat B (gelb) gezeigt. Nicht-polare Wasserstoffe sind ausgeblendet. Die Wassermoleküle sind transparent abgebildet.

In Tabelle 2.3 sind die Bewertungen der verschiedenen Tools für alle Strukturen gezeigt. Die Ergebnisse der SuperStar-Berechnungen sind im Anhang dargestellt (s. Anhang 12.2 auf S. 230).

Tabelle 2.3 – Bewertungen der Wassermoleküle in der Bindetasche von allen Kristallstrukturen durch SuperStar, HINT und Rank und WaterClassifier. Die Bedeutung der Werte sind im Abschnitt 2.3.1 auf S. 48 erklärt. Beim WaterClassifier sind zusätzlich die Wahrscheinlichkeiten angegeben, mit der eine hydrophobe Substruktur des Liganden dort binden könnte (in Klammern).

Kristallstruktur	Wat A			Wat B			Wat C	Wat D
	1E11	1AJ6	3G7E	1E11	1AJ6	3G7E	1AJ6	1E11
SuperStar	51.29	37.44	63.03	11.15	7.55	60.63	5.19	6.28
HINT und Rank	1.00	1.07	1.05	-0.25	0.83	0.88	-0.73	-0.67
WaterClassifier (hydrophob)	0.90 (0.34)	0.80 (0.10)	0.90 (0.34)	0.24 (0.68)	0.34 (0.78)	0.37 (0.79)	0.23 (0.47)	0.22 (0.58)

Wat A wies bei allen Bewertungen herausragende Werte auf. In SuperStar wurden hohe zweistellige Werte bestimmt. Sowohl HINT und Rank als auch WaterClassifier konnten hohe Wahrscheinlichkeiten für die Konservierung der Wassermoleküle ermittelt werden. Beispielsweise waren alle Werte beim WaterClassifier ≥ 0.80 , ein Bereich bei dem im Testdatensatz von Ross *et al.* über 94 % der Wassermoleküle richtig vorhergesagt wurden [125].

Wat B wurde in Abhängigkeit von den Tools und den Strukturen sehr unterschiedlich beurteilt. Auffällig ist die schlechte Bewertung des Wat B in der Struktur von 1E11. Die auf dem Wasser selbst basierenden Tools HINT und Rank sowie WaterClassifier zeigten nur eine sehr geringe Wahrscheinlichkeit für die Konservierung des Wassers. Wie schon oben beschrieben (s. S. 53), befand sich das Wassermolekül bei der SuperStar-Analyse zwischen zwei Maxima. Doch auch das nächste Maximum ist mit 11.15 noch deutlich niedriger als beim Wat A. Für 1AJ6 waren die Ergebnisse nicht einheitlich. Die SuperStar- und WaterClassifier-Werte wiesen auf ein nicht konserviertes und ersetzbares Wasser hin. Der WaterClassifier fand sogar eine fast 80 %ige Wahrscheinlichkeit für einen hydrophoben Rest, der dort platziert werden könnte. HINT und Rank berechnete jedoch noch eine hohe Wahrscheinlichkeit für ein konserviertes Wassermolekül. Bei 3G7E wiederum war das Wat B nach der Bewertung von SuperStar und HINT und Rank konserviert, während WaterClassifier dieses Wasser, sehr ähnlich wie bei 1E11, als ersetzbar bewertete.

Bei Wat C und Wat D waren die Beurteilungen wieder bei allen Strukturen und von allen Tools einheitlich. In der Bewertung mit SuperStar konnten nur einstellige Werte ermittelt werden. Darüber hinaus erachteten auch die beiden anderen Tools diese Wassermoleküle als ersetzbar. Ob diese Tasche von einem hydrophoben oder hydrophilen Rest besetzt werden sollten, ließ sich mithilfe des WaterClassifiers nicht klar vorhersagen, da er für diesen Bereich eine ca. 50 %ige Wahrscheinlichkeit für hydrophobe Reste ermittelte.

Mithilfe dieser Untersuchungen ließen sich drei Wasserpositionen sicher evaluieren. Für das Wat A bewiesen die MD-Simulationen die wichtige strukturelle Funktion des Wassermoleküls. Alle Tools stimmten darin überein, dass dieses Wassermolekül nicht ersetzbar ist. Es sind keine Inhibitoren bekannt, die dieses Wasser verdrängen.

Für Wat C und Wat D waren die Ergebnisse ebenfalls eindeutig. Sie besaßen keinen komplexstabilisierenden Effekt und wurden konsistent als ersetzbar bewertet. Mit Clorobiocin ist auch ein Derivat bekannt, welches Wat C in 1AJ6 mit einem hydrophoben Rest substituiert. Der Ligand von 3G7E selbst übt hydrophile Interaktionen an Positionen des Wat C und Wat D aus. Aus diesen Gründen schien die Bewertung mit einer 50 % Wahrscheinlichkeit für hydrophobe Reste, die der WaterClassifier vorhersagte, nachvollziehbar, da hier beide Fälle auftraten.

Bei Wat B war die Bewertung über die verschiedenen Komplexe nicht einheitlich. Dieses Wasser wurde in der Struktur von 1E11 als nicht-konserviert bewertet. SuperStar konnte dieses Wasser nicht eindeutig vorhersagen. Bei 1AJ6 stimmten SuperStar und der WaterClassifier mit der Zuordnung als ersetzbares Wasser überein. Die Kristallstruktur mit Clorobiocin belegte auch experimentell diese Einschätzung. Lediglich in der Struktur von 3G7E sind die Ergebnisse nicht eindeutig interpretierbar: Sowohl HINT und Rank als auch SuperStar beurteilten das Wasser als sehr wahrscheinlich konserviert. Dies war insbesondere bei SuperStar interessant, weil es das Wat B bei 1AJ6 als nicht konserviert bewertete, welches von der Umgebung und Struktur sehr ähnlich war. SuperStar bewertete die Positionen weiterhin rein strukturbasiert. Die gute Interaktion, die zu dem Liganden möglich war, hatte keinen Einfluss auf die Analyse. Trotzdem erreichte dieses Wasser höhere Werte als das Wat A bei 1E11 und 1AJ6. Dem Ergebnis stand die Bewertung vom WaterClassifier gegenüber. Dieser bewertete das Wasser nicht nur als nicht-konserviert sondern prognostizierte sogar hydrophobe Reste in diesem Bereich. In 1AJ6 wurde, wie oben beschrieben, dieses Wasser auch von einem hydrophoben Rest ersetzt.

Damit ließen die Tools verschiedene Schlussfolgerungen zu: Strukturell war dieser Bereich stärker in Richtung eines Wassermoleküls geprägt und die kraftfeld- und geometriebasierende Bewertung von HINT und Rank stimmte damit überein. Die MD-Simulationen, Bewertungen von WaterClassifier und die Analogie zum Clorobiocin sprachen für einen geringeren Einfluss des Wassers. Dies führte zur folgenden Hypothese: Aus struktureller und energetischen Sicht bestand die Möglichkeit, dieses Wasser zu ersetzen. Dies musste jedoch nicht bedeuten, dass die Verdrängung des Wassers die Freie Bindungsenthalpie verbessert. Für ein Screening wurde dieses Wassermolekül daher als optional angesehen: Es wurde sowohl ein Screening mit als auch ohne dieses Wassermolekül durchgeführt. Die Hypothese, dass das Wat B ersetzbar sei, wurde retrospektiv auch in einer im August 2012 veröffentlichten Röntgenkristallstruktur von BRVAR *et al.* nachgewiesen (PDB-Code: 4DUH) [94]. Die Grup-

pe führte eine strukturbasierte Leitstruktur-Suche nach GyrB-Inhibitoren durch und fanden Bithiazole, die den Wat B-Bereich mit einer Ethyl-Gruppe besetzten.

Nach der Analyse sowie Aufarbeitung der Röntgenkristallstrukturen und einer genauen Evaluierung der verschiedenen Wassermoleküle in der Bindetasche standen damit vier verschiedene Proteinstrukturen für das Screening zur Verfügung. Beim Docking in 1EI1 und in das vervollständigte 1AJ6 wurde nur das Wat A berücksichtigt. Für das Docking in 3G7E wurden zwei Strukturen verwendet. Eines enthielt, wie die anderen Strukturen, ausschließlich Wat A (im Folgenden 3G7E_1 genannt) und das zweite enthielt sowohl Wat A als auch Wat B (3G7E_2). Die Ergebnisse sind in Tabelle 2.4 zusammengefasst.

Tabelle 2.4 – Übersicht der für das Screening verwendeten Strukturen.

Struktur	Bemerkung
1EI1	Vorbereitung: Standardprozedur; ANP (vierfach negativ geladen), enthält Wat A
1AJ6	Vorbereitung: Standardprozedur, Loop A mit 1EI1, Loop B und H136R mit 1KIJ gemodelliert, MD-Simulation, Minimierung; Novobiocin (einfach negativ geladen), enthält Wat A
3G7E_1	Vorbereitung: Standardprozedur; Pyrazolderivat (2H-Tautomer), enthält Wat A
3G7E_2	Vorbereitung: Standardprozedur; Pyrazolderivat (2H-Tautomer), enthält Wat A und Wat B

3 Bibliothekerstellung

Die Erstellung einer Substanzbibliothek für ein virtuelles Screening stellt einen wichtigen Schritt für ein effizientes Vorgehen dar. Die Anzahl der möglichen Substanzen ist derart hoch, dass die Generierung oder Verarbeitung aller denkbaren Verbindungen *in silico* bisher nicht möglich ist. Da ein virtuelles Screening aufwendige Rechenschritte erfordert, kann der Zeitaufwand stark reduziert werden, wenn kleinere und optimierte Datensätze verarbeitet werden. Einfache und schnell anzuwendende Filterkriterien können die Bibliothek reduzieren, so dass aktive Substanzen, die präferierte Eigenschaften, wie ein günstiges ADME-Profil, besitzen, angereichert werden.

3.1 Chemischer Raum

Die Grundlage für die Erstellung einer Screening-Bibliothek setzt zunächst das Verständnis des chemischen Raumes voraus. Der chemische Raum stellt dabei einen sehr abstrakten Begriff dar, der wie folgt von LIPINSKI *et al.* beschrieben wurde [141]:

„Chemical space can be viewed as being analogous to the cosmological universe in its vastness, with chemical compounds populating space instead of stars.“

Nach dieser und anderen Definitionen [45, 142] wird unter dem Begriff also die Gesamtheit aller möglichen, organisch-chemischen Verbindungen verstanden – folglich eine unvorstellbar große Zahl von Substanzen. Aus medizinisch-chemischer Sicht werden mit diesem Ausdruck auch häufig alle als Wirkstoffe infrage kommenden Substanzen bezeichnet (z. B. Möglichkeit der Synthese) [143]. Damit wird auch Rücksicht auf die Eigenschaften und Zugänglichkeit der Verbindungen genommen. Ausgehend von diesen Definitionen gibt es verschiedene Schätzungen über die Größe des chemischen Raums: 10^{24} – 10^{60} Substanzen [144, 145].

– Datenbanken

Im Vergleich zur theoretischen Größe des chemischen Raums ist der beschriebene oder synthetisierte deutlich kleiner. Die größte virtuelle Substanzdatenbank ist die CHEMICAL UNIVERSE DATABASE DB-17 mit über 166 Milliarden Molekülen [146]. Dazu erstellten RUDDIGKEIT *et al.* mittels eines kombinatorischen Ansatzes eine Datenbank aus allen prinzipiell möglichen Molekülen, die keinen chemischen Regeln widersprechen [146]. Sie besteht aus Substanzen mit bis zu siebzehn schweren Atomen (Kohlenstoff-, Stickstoff-, Sauerstoff-, Schwefelatomen oder Halogene). Die Begrenzung auf siebzehn Atome ist auf die hohen Anforderungen an Computerressourcen zurückzuführen. Dennoch stellt dies einen interessanten Bereich dar, da es 367 zugelassene Arzneistoffe mit einer Molekülmasse von < 350 Da gibt. Zudem befindet sich die Datenbank mit diesen Eigenschaften zumindest im Leitstrukturbereich [146]. Sie verdeutlicht auch, wie wenig der chemische Raum bisher erforscht ist.

Die größte frei zugängliche Datenbank von synthetisierten Substanzen, PUBCHEM, enthält 63 Mio. Verbindungen (Stand: November 2014) [147]. Davon bestehen nur 9.6 Mio. Substanzen aus bis zu siebzehn schweren Atomen. Des Weiteren gibt es noch Datenbanken, die speziell käuflich erwerbba-re Substanzen beinhalten. Deren größter Vertreter ist die ZINC-Datenbank mit knapp 23 Mio. Substanzen (Stand: Mai 2014)[148]. Diese Zahlen zeigen eindrucksvoll, wie unerschlossen der chemische Raum selbst bei Molekülen ist, die weniger

als siebzehn schwere Atome besitzen (s. Tabelle 3.1). Die größte Datenbank synthetisierter oder isolierter Verbindungen ist die CAS REGISTRYSM mit über 90 Mio. Substanzen (Stand November 2014)[149].

Tabelle 3.1 – Vergleich der Datenbankgrößen und Anteil an Verbindungen mit weniger als 17 schweren Atomen.

Datenbank	Größe (in Mio.)	≤ 17 schwere Atome (in Mio.)
GDB-17	166 000	166 000
PubChem	63	9.6
ZINC	23	2

Neben der Verwendung von einer Datenbank wäre es auch möglich, direkt *in silicio* neue Substanzen zu erstellen und zu analysieren, das so genannte „*computer-based de novo Design*“ [150]. Aber auch bei diesem Ansatz treten die gleichen Limitierungen auf, bspw. Beschränkung der Anzahl der Substanzen durch Computerressourcen. Der Vorteil besteht in einer freieren Wahl der Substanzen und der Optimierungsmöglichkeit. Dafür muss jedoch im Anschluss der Analyse die Substanz noch beschafft werden (z. B. mittels Synthese). Dieser möglicherweise sehr aufwendige Schritt entfällt, bei Verwendung einer Datenbank von bereits beschriebenen oder erwerbba-ren Substanzen.

– Drug-like Kriterien

Die Erforschung der idealen physikochemischen Parameter für Kandidaten der Arzneistoffsuche begann mit einer beachtenswerten Veröffentlichung von LIPINKSI *et al.* 1997 [151]. LIPINSKI *et al.* untersuchten eine Datenbank mit über 2,000 Substanzen von mindestens in klinischer Prüfung Phase II befindlichen Substanzen oder zugelassenen Arzneistoffen, die oral verfügbar sind. Nur 12 % der Sub-

stanzen hatten mehr als 10 Wasserstoffbrückenakzeptoren, 8 % mehr als 5 Wasserstoffbrückendonoren, 11 % eine höhere Molekülmasse als 500 Da und 10 % einen höheren cLogP als 5 [151]. Mithilfe dieser Daten formulierten sie die „Rule of five“ (Ro5), bekannt auch unter „Lipinski’s Rule of five“ oder „Pfizer’s Rules of five“: Für optimale ADME-Eigenschaften darf eine peroral bioverfügbare Substanz nicht mehr als

- 1) 10 Wasserstoffbrückenakzeptoren,
- 2) 5 Wasserstoffbrückendonoren,
- 3) 500 Da Molekülmasse und
- 4) einen cLogP von 5 besitzen.

Seit dieser Entdeckung gibt es zahlreiche Veröffentlichung, die sich mit diesem Thema beschäftigen und Subgruppen (z. B. Indikationen) oder andere Fragestellungen (wie Leitstruktursuche) untersucht haben [45, 152–154]. Dabei werden als weitere Parameter häufig noch die Anzahl der drehbaren Bindungen und die polare Oberfläche (PSA) verwendet. Dennoch bleiben die Ro5 eines der am meist genutzten Filterkriterien für Drug-Like-Kandidaten.

Die Anwendung physikochemischer Filterkriterien ist eine sinnvolle Maßnahme zur Reduktion großer Datensätze, um potentielle Hits anzureichern und Substanzen mit voraussichtlich schlechten Eigenschaften zu reduzieren. Lediglich die Wahl der Kriterien kann variieren und die anhaltende Diskussion darüber verdeutlicht, dass die Kriterien problemorientiert erarbeitet werden müssen.

3.2 Material und Methoden

3.2.1 ZINC-Datenbank

Die ZINC-Datenbank (Zinc Is Not Commercial) stellt eine gute Startauswahl für ein virtuelles Screening dar. Die Datenbank wurde ursprünglich für die Durchführung von strukturbasierten virtuellen Screenings erstellt [155]. Dafür werden die chemischen Strukturen von über 200 Anbietern (Stand: September 2012) extrahiert und verarbeitet. Alle eingelesenen Substanzen werden zunächst nach folgenden Kriterien gefiltert: Die Molekülmasse muss zwischen 50 und 500 Da betragen und es werden nur Verbindungen weiterverarbeitet, die aus folgenden Elementen bestehen: C, N, O, S, P, Cl, F, Br, I und H. Weiterhin werden bei Salzen die Gegenionen entfernt und bestimmte Funktionalitäten aussortiert bzw. beschränkt, die aufgrund ihrer Reaktivität (z. B. Carbonsäureanhydride) oder ihren ungünstigen physikochemischen Eigenschaften (z. B. maximal ein quartäres Ammoniumsalz) für ein SBVS ungeeignet sind. Alle Regeln sind in Tabelle 3.2 aufgeführt, gelten allerdings nicht für bereits bekannte bioaktive Strukturen, da diese ihre Aktivität bewiesen haben. Die Filterkriterien sorgen für einen Ausgangsdatensatz von ca. 20 Millionen unterschiedlichen, käuflich erwerblichen Strukturen.

Tabelle 3.2 – Regeln des Standardfilters. Alle Strukturen, die grundsätzlich in die ZINC-Datenbank aufgenommen werden, müssen mindestens diesen Regeln entsprechen.

Struktur	Anzahl	Struktur	Anzahl
Nicht-Wasserstoffatome	5 bis 40	Phosphene	0
Kohlenstoffatome	2 bis 40	Heptane	0
Stickstoff-, Sauerstoff-, Schwefelatome	0 bis 20	Perchlorate	0
Fluoratome	0 bis 7	Pentafluorphenylester	0
Halogenatome außer Fluoratome	0 bis 6	Paranitrophenylester	0
Di- und Triphosphate	0 bis 3	Hydroxybenzotriazolester	0
β -Oxoammonium	0 bis 2	Triflate	0
(Thio)-Epoxyde, Aziridine	0 bis 2	Lawessons Reagenz	0
Benzylammonium	0 bis 2	Phosphoramide	0
Thioester	0 bis 2	aromatische Azide	0
Hydroxamsäure	0 bis 2	Chloramidine	0
Nitrogruppen	0 bis 2	Isonitrile	0
Imine	0 bis 2	Triacyloximes	0
Propenale	0 bis 2	Acylcyanide	0
Kationische Kohlenstoff-, Chlor-, Iod-, Phosphor- oder Schwefelatome	0 bis 2	Sulfonylcyanide	0
Acylhydrazide	0 bis 2	Cyanophosphonate	0
Thiocyanate	0 bis 1	Azocyanamide	0
Iso(thio)cyanate	0 bis 1	Azoalkanale	0
Ammonium	0 bis 1	Carbazide	0
Schwefelsäureester	0 bis 1	Säureanhydride	0
Sulfonsäureester	0 bis 1	Peroxide	0
Sulfone	0 bis 1	Acrylnitril	0
Sulfonsäurehalogenide	0 bis 1	Phosphorane	0
Säurehalogenide	0 bis 1	Phosphorhalogenide	0
Alkylhalogenide	0 bis 1	Cyanhydrine	0

Neben den Standardkriterien gibt es noch zusätzlichen Filter. Dieser ist beim Aussortieren von reaktiven Verbindungen deutlich strenger als der Standardfilter. Beispielsweise ist beim Standardfilter ein Säurehalogenid erlaubt, welches wiederum über den zweiten Filter ausgeschlossen wird. Darüber hinaus wird noch auf weitere reaktive Strukturen, wie Michael-Akzeptoren, eingegangen. Ein mit diesen Filterkriterien bearbeiteter Datensatz wird als „Clean“ bezeichnet. Insgesamt werden letztendlich über 200 Filterregeln eingesetzt, wodurch die Substanzbibliothek auf ca. 12 Millionen Strukturen reduziert werden konnte.

Von diesem „Clean Dataset“ ausgehend gibt es zahlreiche Subsets, die die Untersuchung unterschiedlicher Fragestellungen erlauben (z. B. Leitstruktursuche oder fragmentbasierte Ansätze). Für die in dieser Arbeit durchzuführenden Untersuchungen bietet sich der „Drug-Like“-Filter an. Dieser besteht aus einer modifizierten Version der Ro5 (s. Tabelle 3.3). Dadurch wurde der Datensatz um ungefähr 1 Mio. auf 11 Mio. Substanzen reduziert. Diese Datenbank wird „Drug-Like Clean Dataset“ genannt.

Tabelle 3.3 – Filterkriterien für den „Drug-Like“-Filter. Regeln, die nicht den Ro5 entsprechen, sind kursiv dargestellt.

Filterkriterium	Grenze
Molekülmasse	150 bis 500 Da
berechneter logP	bis 5
Anzahl Wasserstoffbrückendonoren	bis 5
Anzahl Wasserstoffbrückenakzeptoren	bis 10
Anzahl drehbarer Bindungen	bis 7
<i>polare Oberfläche</i>	<i>bis 150 Å²</i>

3.2.2 Mona

Zum Filtern und zur Auswertung der Datenbank wurde das Programm MONA verwendet [156]. MONA ermöglicht es, große Moleküldatenbanken einzulesen, nach physikochemischen und chemischen Parametern zu filtern und bezüglich der Verteilung von Eigenschaften auszuwerten. Dabei erzeugt MONA automatisch eindeutige Tautomere, Protonierungszustände und Konformationen der Substanzen und eliminiert Duplikate, auch von verschiedenen Quellen, selbstständig. Zusätzlich kann mit Hilfe von SMARTS-Ausdrücken nach weiteren Strukturelementen gefiltert werden.

MONA basiert auf den Grundlagen von NAOMI [157], einem Tool zum Einlesen verschiedener Dateiformate. Es verarbeitet die Dateien in SQL-Datenbanken, lässt sich aber über ein GUI interaktiv steuern [156]. Es eignet sich aufgrund seiner hohen Geschwindigkeit beim Einlesen und Verarbeiten der Substanzen und seiner hohen Kapazität, um unterschiedliche Filterkriterien für die Datenbank zu testen, anzuwenden und direkt bezüglich ihrer Effektivität auszuwerten.

3.3 Filterkriterien für die ZINC-Datenbank

Ausgangspunkt für die Erarbeitung eigener Filterkriterien stellte der „Drug-Like Clean Dataset“ von ZINC dar. Für die Durchführung der strukturbasierten Leitstruktursuche gab es noch verschiedene weitere Kriterien, die berücksichtigt werden sollten und zu zusätzlichen Filterregeln führten. Diese waren nötig, um die Datenbank für das Vorgehen anzupassen. Dies schloss analytische Probleme (z. B. Anzahl an Chiralitätszentren), Erwägungen vom Dockingalgorithmus (Anzahl drehbarer Bindungen) und dem Verständnis der Bindetasche (Anzahl an Stickstoff- und Sauerstoffatomen) mit ein. Die Auswirkungen werden in den einzelnen Abschnitten diskutiert. Die hier bestimmten Parameter wurden nur zum Filtern des Datensatzes durch MONA verwendet. Daher kann z. B. eine Bindung während des Filterns als nicht-drehbar definiert werden, obwohl sie im Dockingprozess später dennoch rotiert wird.

Trotz der hohen Kapazität von MONA war es nicht möglich, den gesamten „Drug-Like Clean Dataset“ mit insgesamt 11 Mio. Substanzen zu verwenden. Für die Optimierung der Parameter wurde daher ein Testdatensatz aus dem „Drug-Like Clean Dataset“ erstellt, der im folgenden MONA-Testset genannt wird. Dieser wurde erzeugt, indem 100,000 Substanzen zufällig aus dem Gesamtdatensatz ausgewählt wurden. Zunächst wurde der Effekt jedes Filterkriteriums auf den MONA-Testset einzeln untersucht. Anschließend wurden die Kriterien kombiniert und mit NAOMI auf den gesamten Datensatz angewendet. In Abbildung 3.1 ist die Vorgehensweise schematisch dargestellt.

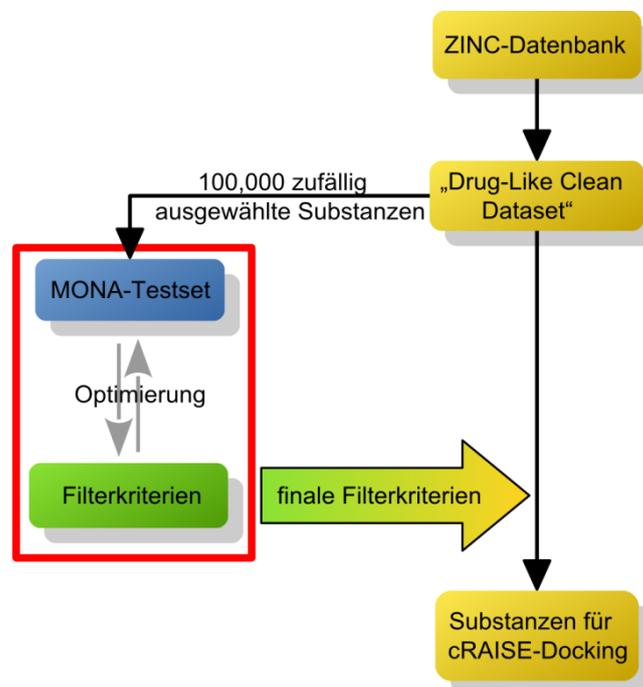


Abbildung 3.1 – Vorgehen für das Filtern der ZINC-Datenbank. Ausgehend von der ZINC-Datenbank wurde der „Drug-Like Clean Dataset“ verwendet. Von diesem wurden 100,000 Substanzen zufällig ausgewählt (Mona-Testset). Mit diesem Testset wurde die Optimierung der Filterkriterien erarbeitet (roter Kasten). Anschließend wurden diese Filterkriterien auf den gesamten „Drug-Like Clean Dataset“ angewendet.

– Anzahl an drehbaren Bindungen

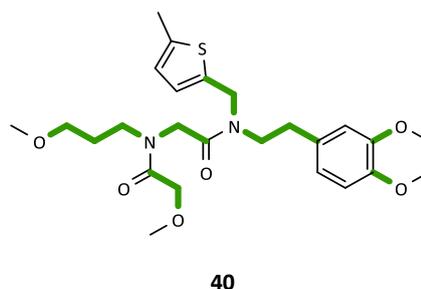
Die Definition von drehbaren Bindungen ist sowohl aus chemischer als auch algorithmischer Sicht nicht trivial. So gelten beispielsweise Amide wegen ihrer hohen Rotationsbarriere als pseudoplanar und können als nicht-drehbar betrachtet werden. Andererseits gibt es dann zwei verschiedene Substitutionsmuster am Amid (in Analogie zur cis-/trans-Isomerie, s. Abbildung 3.2). Von unterschiedlichen Algorithmen wer-



Abbildung 3.2 – Analogie der „Cis-/Trans-Isomerie“ bei Amid. Die nicht frei drehbare Bindung ist grün eingefärbt.

den Amide deswegen entweder als drehbar oder auch als nicht-drehbar bewertet. Aufwendiger wird dies bei Phänomenen, die schwieriger zu beurteilen sind. Ein Beispiel wäre die Drehbarkeit bei Atropisomerie. MONA benutzt eine sehr einfache Definition für drehbare Bindungen: Als drehbar werden Einzelbindungen angesehen, die nicht im Ring liegen oder zu einem terminalen schweren Atom gehen. Amidische C-N-Bindungen und Nitrile werden als nichtdrehbar betrachtet. Damit werden alle Bindungen zwischen verschiedenen Ringsystemen als drehbar definiert.

Die Anzahl der drehbaren Bindungen stellte ein sehr wichtiges Kriterium dar. Vor dem Docking werden von cRAISE alle möglichen plausiblen Konformere erzeugt und gespeichert (s. Abschnitt 4.1.1 auf S. 80). Rotierbare Bindungen lassen die Zahl der Konformere exponentiell ansteigen, so dass bei der Ligandenvorbereitung viel Rechen- und Speicherkapazität benötigt und der Dockingprozess deutlich verlangsamt werden würde. Auch aus der Perspektive der „Drug-



40

Likeness“ haben Substanzen mit weniger drehbaren Bindungen einen Vorteil. Ein Grund dafür könnte der Einfluss der Entropie auf sehr flexible Liganden sein [153]. Obwohl die Strukturen durch die Filter von ZINC auf maximal sieben drehbare Bindungen eingeschränkt wurden, wiesen im MONA-Testset noch über 17.2 % der Moleküle mehr als sieben drehbare Bindungen auf. Dies war ein überraschendes Ergebnis, da es sogar Strukturen gab, die weit über zehn drehbare Bindungen aufwiesen (s. Beispiel Abbildung 3.3), obwohl ZINC die gleiche Definition für drehbare Bindungen wie MONA hat und den Datensatz entsprechend vorher filterte.

Abbildung 3.3 – ZINC35798004 (**40**): Beispiel für ein Molekül mit 15 drehbaren Bindungen (grün markiert).

Um rechenaufwendige Moleküle für das Docking zu entfernen, wurden Moleküle mehr als acht rotierbare Bindungen herausgefiltert, wodurch sich Datensatz um ca. 8 % reduzierte.

Um rechenaufwendige Moleküle für das Docking zu entfernen, wurden Moleküle mehr als acht rotierbare Bindungen herausgefiltert, wodurch sich Datensatz um ca. 8 % reduzierte.

Wie in Abbildung 3.4 zu sehen, hatte der Großteil der Verbindungen 4 – 7 drehbare Bindungen.

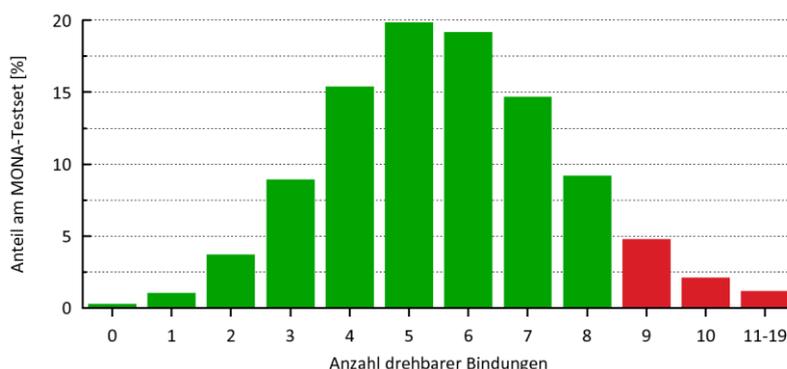


Abbildung 3.4 – Anzahl der drehbaren Bindungen im MONA-Testset. Die Werte sind nahezu normalverteilt mit einem Maximum bei 5 – 6. Nach den Filterkriterien der ZINC-Datenbank sollten nicht mehr als 7 drehbare Bindungen enthalten sein, MONA ordnet den Verbindungen jedoch deutlich höhere Werte zu (grün: weiter verwendete Substanzen; rot: herausgefilterte Substanzen).

– Molekülmasse

Die Molekülmasse ist ein kritischer Parameter für eine Hit-Struktur. Während des Optimierungsprozesses werden Moleküle sehr häufig schwerer [45]. Andererseits zeigen zu kleine Strukturen, eine unspezifische und unselektive Bindung. Gleichzeitig weisen sie häufig nur eine schwache Bioaktivität auf. Während des Docking-Prozesses werden daher auch größere Moleküle bevorzugt, weil enthalpische Effekte meist additiv berechnet werden (s. Abschnitt auf S. 100), und kleine Moleküle würden unnötigerweise prozessiert.

Wie in Abbildung 3.5 ersichtlich, war der MONA-Testset schon weitestgehend an diese Anforderungen angepasst: Über 70 % der Verbindungen besaßen eine mittelgroße Molekülmasse von 300 – 400 Da.

Aus diesen Gründen wurde der Bereich für die Molekülmasse geringfügig von 150 – 500 Da auf 200 – 450 Da verkleinert. Die Anpassung an die neuen Grenzen filterte knapp 8 % des MONA-Testsets heraus.

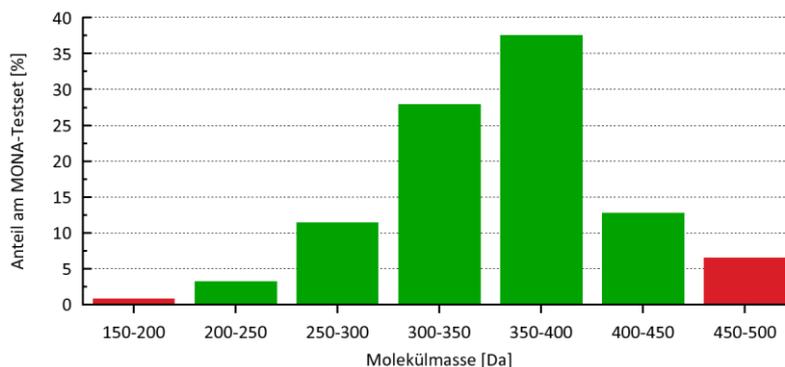


Abbildung 3.5 – Verteilung der Molekülmasse im MONA-Testset. Bis 400 Da steigt die Anzahl der Substanzen langsam an und fällt danach sehr steil ab. Damit waren vor allem kleine bis mittelschwere Substanzen im Datensatz enthalten (grün: weiter verwendete Substanzen; rot: herausgefilterte Substanzen).

– Anzahl an Stickstoffen und Sauerstoffen

Um das Minimum des im cRAISE-Docking gewählten Pharmakophors (s. Abschnitt 4.2.1 auf S. 87) zu erfüllen, muss mindestens ein guter Wasserstoffbrückendonator und -akzeptor vorhanden sein. Diese Eigenschaft wird in der Regel von einem Sauerstoff- oder Stickstoffatom erfüllt. Daher wurden alle Moleküle, die keines der beiden Atome enthielten, nicht weiter berücksichtigt. Dies betraf im MONA-Testset aber nur wenige Substanzen (29 von 100 000 Substanzen).

– Anzahl an Ringsystemen

Alle bekannten Inhibitoren besetzen die Bindetasche mit mindestens einem Ringsystem (s. Abschnitt 1.3 auf S. 14). Daher wurden alle Strukturen ohne Ringsystem herausgefiltert (280 Substanzen im MONA-Testset).

– Anzahl an Chiralitätszentren

Substanzen mit Chiralitätszentren sind in Screening-Experimenten problematisch. Dabei verursacht diese Eigenschaft nicht grundsätzlich Probleme bei der Inhibition oder dem virtuellen Screening. Diese sind eher in der Datengewinnung, -verarbeitung und der Akquisition zu sehen: Zunächst muss der Hersteller die Substanzen stereochemisch genau charakterisiert haben, um die Reinheit oder die Mischung beschreiben zu können. Anschließend muss die ZINC-Datenbank, die in automatisierten Prozessen die Substanzen von vielen verschiedenen Quellen aufarbeitet, diese Information sowohl korrekt übermittelt bekommen als auch richtig verarbeiten. Schließlich wird diese Information von der ZINC-Datenbank auch hier importiert und zugeordnet. Nach dem virtuellen Screening muss bei der Durchführung des Assays weiterhin sichergestellt werden, dass das vom Hersteller ausgewiesene Enantiomer erhalten wurde und rein vorliegt. Dadurch treten letztendlich auch Problemen bei der Analytik auf. Weitere Effekte, wie Racemisierungen, können zusätzlich auftreten.

Ein einfaches Beispiel für diese Problematik war die Substanz ZINC00164086 (**44**) (s. Abbildung 3.6) aus dem MONA-Testset. Diese Substanz besaß im Cyclopropylrest ein Chiralitätszentrum. Der verbreitetste Name war laut der ZINC-Datenbank: „1-[(2,2-dichlorocyclopropyl)methyl]benzene“. Es gab also keinen Hinweis auf die Konfiguration. Bei verschiedenen Anbietern war die Substanz meist ohne weitere Hinweise auf das Chiralitätszentrum aufgelistet. Bei dem Lieferanten MCULE^{vii} fand sich lediglich der Hinweis, dass es sich um eine chirale Substanz mit unbekannter Konfiguration handelte.

^{vii} ID: MCULE-5015716225

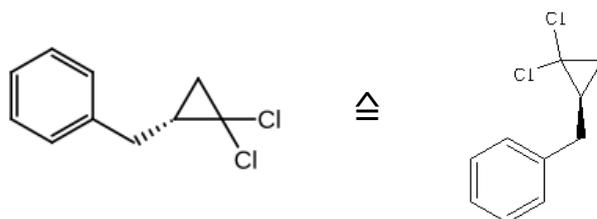


Abbildung 3.6 – ZINC00164086 (**44**): Originaldarstellung in MONA (links) und der ZINC-Datenbank (rechts). Bei beiden Strukturen ist im Cyclopropylrest ein chirales Zentrum enthalten, das als *R*-Enantiomer vorliegt.

Folglich gab es keine Informationen, in welcher Konfiguration diese Struktur vorlag. Sowohl in Mona als auch in der ZINC-Datenbank wurde aber jeweils nur das *R*-Enantiomer dargestellt. Dieser Fall zeigte, dass es unter Umständen gar nicht möglich war, eine Aussage darüber zu treffen, welche Enantiomere vorlagen. Dies hätte auch die Auswertung richtig verarbeiteter Strukturen erschwert. Daher wurden alle chiralen Substanzen aus dem Datensatz eliminiert.

Im MONA-Testset hatten fast 50 % aller Substanzen mindestens ein Chiralitätszentrum (s. Abbildung 3.7). Diese wurden herausgefiltert.

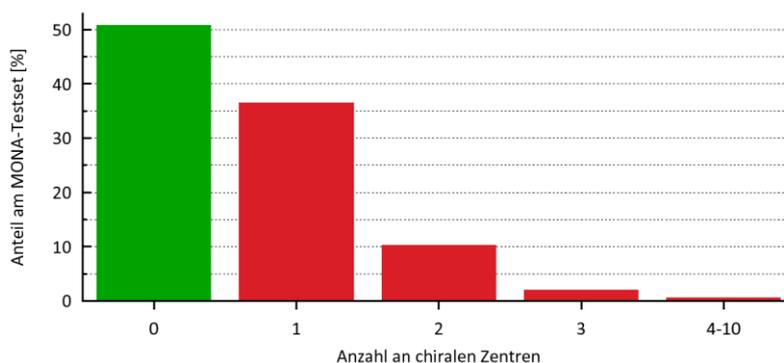


Abbildung 3.7 – Anzahl chiraler Zentren im MONA-Testset. Die Hälfte des Datensatzes hatte kein chirales Zentrum. Die Häufigkeit der chiralen Zentren nahm im weiteren Verlauf fast exponentiell ab (grün: weiter verwendete Substanzen; rot: herausgefilterte Substanzen).

– Anzahl an E/Z-Isomere

Bei der E/Z-Isomerie traten ähnliche Probleme wie bei den Chiralitätszentren auf. Die Bestimmung und Handhabung war problematisch. Der Anteil an dieser Isomerie war deutlich niedriger, so dass der Einfluss auf die Größe des Datensatzes gering war. Es gab im MONA-Testset lediglich 0.7 % Substanzen mit einem oder mehreren E/Z-Isomeren. Diese wurden ebenfalls entfernt.

– Weitere Parameter

Neben den bereits diskutierten Parametern wurden noch weitere Struktureigenschaften analysiert: der cLogP, die Gesamtladung und die Anzahl an Halogenen im Molekül. Diese haben ebenfalls einen

kritischen Einfluss auf die Substanzbibliothek, wurden jedoch wegen der hier dargestellten Gründe nicht weiter angepasst.

Wie schon aus den Ro5 deutlich wird, ist der cLogP-Wert ein wichtiger Parameter. Durch den „Drug-Like“-Filter waren die Substanzen auf einen cLogP-Wert < 5 beschränkt. Im MONA-Testset war der Parameter schon günstig verteilt (s. Abbildung 3.8). Ein Großteil der Substanzen hatte einen Wert, der deutlich kleiner als drei war. Nur 3.5 % der Substanzen besaßen einen ausgeprägten hydrophilen Charakter (cLogP < 0). Eine generelle Problematik stellten die Tautomerisierungs-

und Protonierungszustände dar. Für jede Substanz wurde zwar ein eindeutiger Zustand generiert, dieser musste jedoch nicht der energieärmste oder der im physiologischen System vorherrschende Zustand sein. In Abbildung 3.10 ist ein Beispiel für eine Struktur, bei der sowohl das Tautomer als auch die Protonierung starken Einfluss auf den cLogP-Wert hatten. Bei anderen möglichen Zuständen wäre der cLogP-Wert deutlich höher. Folglich ist es nicht möglich, den cLogP-Wert immer korrekt zuzuordnen. Für die Bibliothek war aber keine weitere Anpassung mehr erforderlich.

Der Protonierungszustand ist zusätzlich auch für die Gesamtladung entscheidend. Ein Vorteil des cRAISE-Tools liegt darin, dass diese während des Dockings flexibel angepasst wird (s. Abschnitt 4.1.1 auf S. 80). Daher war nicht entscheidend, welcher Protonierungszustand während der Vorbereitung der Substanzbibliothek angenommen wurde, und der Einfluss auf das Docking nicht vorhersagbar. Eine Anpassung erübrigte sich aus diesen Gründen. In den von MONA bestimmten Protonierungen hatten knapp 75 % der Substanzen keine, 21 % eine einfach positive und 3.5 % eine einfach negative Ladung (s. Abbildung 3.10).

Halogene sind in einem Molekül aus medizinisch-chemischer Sicht unproblematisch. Auch wenn es Arzneistoffe gibt, bei denen diese unerwünschte Wirkungen hervorrufen (z. B. Amiodaron), konnte gezeigt werden, dass Toxizitätsbedenken, selbst bei hyperhalogenierten Verbindungen, unbegründet sind [158]. Viel mehr gibt es Bestrebungen, Halogene häufiger, auch in Hinblick auf die kürzlich in Fokus geratenen Halogenbindungen, zu berücksichtigen [159]. Dennoch blieben Bedenken bei der

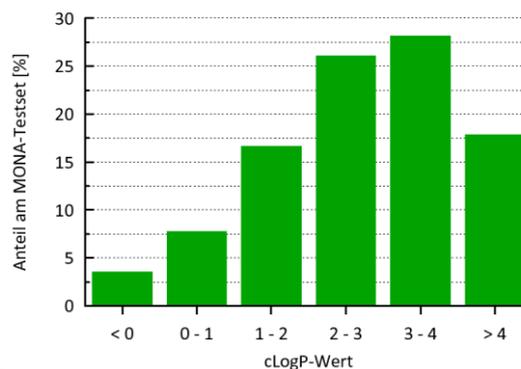
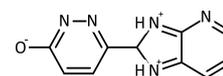


Abbildung 3.8 – cLogP-Werte im MONA-Testset. Es handelt es sich um eine linksschiefe Verteilung. Daher waren vor allem schwach- bis mittelhydrophobe Substanzen im Datensatz vorhanden.



42

Abbildung 3.9 – ZINC65356320: Beispiel aus dem MONA-Testset mit einem cLogP-Wert von -3.8.

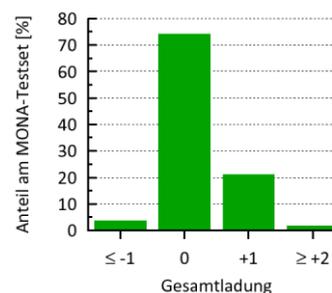


Abbildung 3.10 – Gesamtladung der Substanzen im MONA-Testset. Fast 75 % der Substanzen weisen keine Ladung auf.

synthetischen Erschließung von vielfach halogenierten Verbindungen. Im Testdatensatz hatten über 5 % der Verbindungen drei oder mehr Halogenatome (s. Abbildung 3.11). Diese Anreicherung lag darin begründet, dass 80 % dieser Verbindungen Trifluormethyl-Gruppen besaßen. Diese sollten aufgrund der guten Darstellung und Möglichkeit der Bioisosterie in der Weiterentwicklung keine Schwierigkeiten verursachen. Daher konnte auch hier auf einen weiteren Filter verzichtet werden.

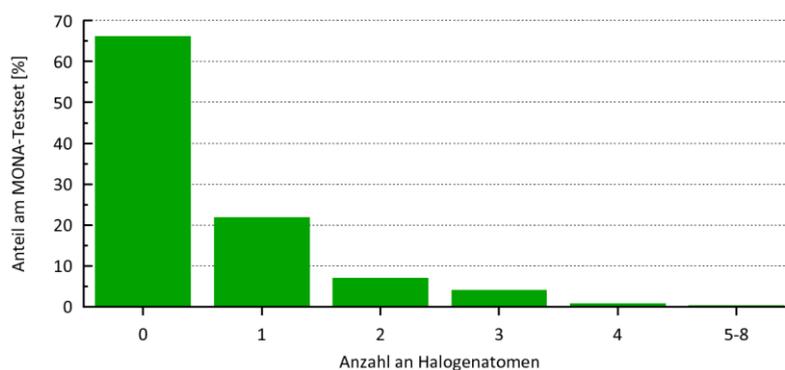


Abbildung 3.11 – Anzahl an Halogenatomen im Testdatensatz. Der Großteil der Verbindungen besaß kein oder nur ein Halogen.

– Zusammenfassung

Durch Anwendung aller erarbeiteten Filter (s. Tabelle 3.4) reduzierte sich der MONA-Testset von den ursprünglich 100,000 Substanzen auf 43,108 Substanzen. Dies entspricht einer Reduktion von 57 %. Bei einer rein additiven Betrachtung aller Filterkriterien betrug die Reduktion 64 %. Damit betrug die Überschneidung der verschiedenen Parameter 7 %. Diese war nachvollziehbar, da die Eigenschaften teilweise voneinander abhängig waren. Bei einer größeren Molekülgröße nehmen beispielsweise die Molekülmasse und die Anzahl der drehbaren Bindungen zu.

Tabelle 3.4 – Übersicht über die endgültigen Filterkriterien, mit denen das „Drug-Like“-Subset der ZINC-Datenbank gefiltert wurde.

Filterkriterium	Grenze
Anzahl drehbarer Bindungen	0 – 8
Molekulgewicht	200 – 450 Da
Anzahl an Stickstoff- oder Sauerstoffatomen	≥ 1
Anzahl an Ringsystemen	≥ 1
Anzahl an chiralen Zentren oder cis/trans-Isomerie	0

3.4 Anwendung des Filters auf den „Drug-Like Clean Dataset“

Schließlich wurde der „Drug-Like Clean Dataset“ der ZINC (Stand: September 2012) ebenfalls mit diesen Kriterien (s. Tabelle 3.4) gefiltert. Dabei reduzierte sich der Datensatz von ca. 11 Mio. Substanzen auf 3.4 Mio. Substanzen. Das stellte eine Reduktion auf knapp 31 % dar. Damit fiel die Reduktion bei der Datenbank größer aus als im MONA-Testset (69 % vs. 57 %). Dies lag zum einen an den Duplikaten, die noch in der gesamten Datenbank enthalten waren. Zum anderen könnten bei einer viel größeren Datenbank häufiger Variationen von größeren Molekülen auftreten, die alle herausgefiltert wurden. Letztendlich vereinfacht eine stärkere Reduktion spätere Schritte im Screening.

Wichtige Parametergruppen und deren Verteilung in der gefilterten Datenbank werden in den nachfolgenden Abschnitten diskutiert.

– Molekülgröße

In Abbildung 3.12 sind die Parameter dargestellt, die die „Größe der Substanzen“ beschreiben (Molekülmasse, Anzahl schwerer Atome und Anzahl drehbarer Bindungen).

Ein Großteil der Substanzen in der gefilterten Datenbank besaß eine Molekularmasse zwischen 300 und 400 Da. Dies stellte einen guten Bereich dar, weil der Fokus auf mittelgroße Substanzen lag (s. Abschnitt 1.2 auf S. 9). Interessanterweise reduzierte sich der Anteil der schweren Moleküle im Vergleich zum Ausgangsdatsatz nur marginal, obwohl die meisten eingesetzten Filterkriterien kleine Moleküle bevorzugen sollten. Im MONA-Testset lag der Anteil an schweren Molekülen (400 – 450 Da) bei 13.7 % im gefilterten Datensatz betrug dieser 11.2 % und lag damit auf einem vergleichbaren Niveau. Die Häufigkeit, dass zumindest ein chirales Zentrum in der Struktur vorliegt - das stärkste Filterkriterium bei den verwendeten Filtern - ist im „Drug-Like Clean Dataset“ für die einzelnen Moleküle folglich ähnlich hoch.

Die Anzahl der schweren Atome korrelierte erwartungsgemäß sehr stark mit der Molekülmasse. Die Verteilung war sehr eng und über 40 % aller Moleküle wiesen zwischen 23 und 26 schwere Atome auf.

Die Anzahl der drehbaren Bindungen zeigte eine ähnliche Verteilung der Werte, unterschied sich aber von den anderen Parametern, da sich auch ein größerer Anteil am oberen Grenzbereich mit sieben drehbaren Bindungen befand. Damit war die Datenbank mit komplexeren Substanzen angereichert und bestand nicht nur aus großen starren Aromaten oder unflexiblen Molekülen.

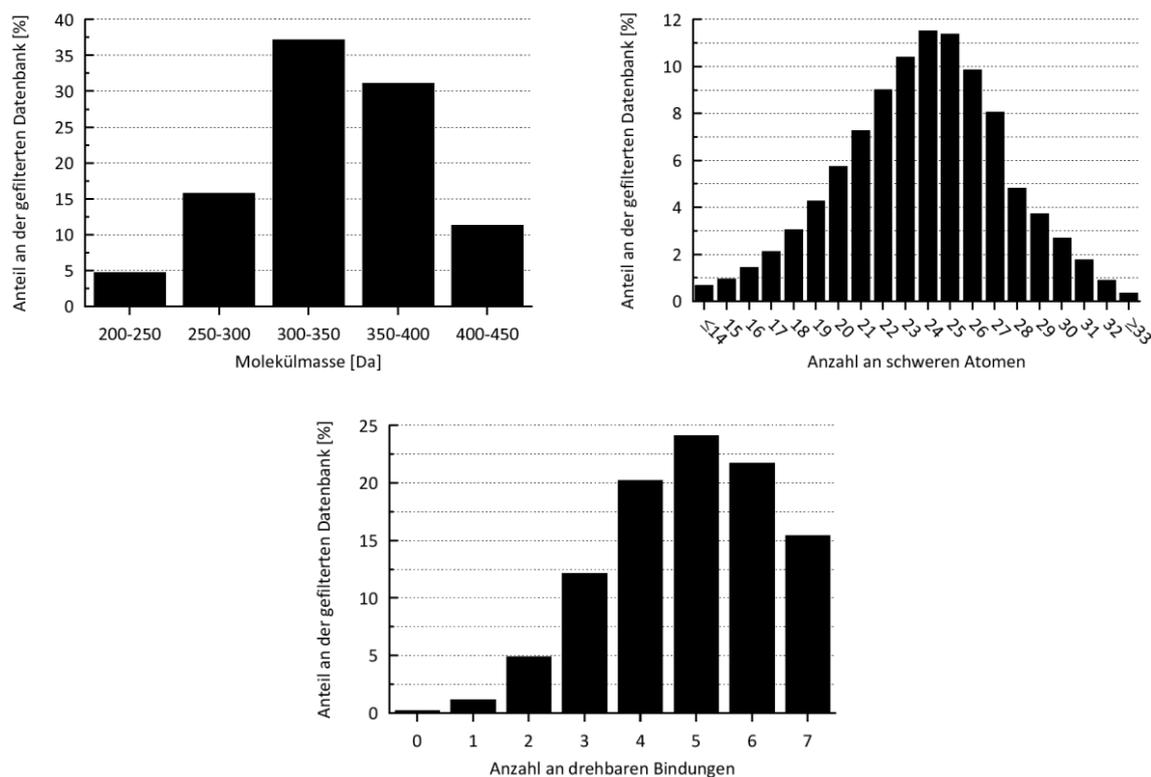


Abbildung 3.12 – Parameter für die Molekülgröße. Die Molekülmasse und Anzahl der schweren Atome verhalten sich in ihrem Verlauf sehr ähnlich und zeigen in dem unteren bis mittleren Bereichen eine gute Anreicherung. Die Anzahl der drehbaren Bindungen ist ähnlich wie im Ausgangsdatensatz verteilt und hat einen deutlich höheren Anteil an der oberen Grenze.

– Hydrophile Interaktionen

Wichtige Parameter für hydrophile Interaktionen sind die Anzahl der Wasserstoffbrückendonoren, -akzeptoren und die PSA (s. Abbildung 3.13).

In dieser Gruppe wurden die Vorgaben von den Ro5 von fast allen Substanzen deutlich unterschritten (≤ 5 Donoren, ≤ 10 Akzeptoren). Nur 0.3 % der Substanzen hatten vier oder mehr Wasserstoffbrückendonoren. Weiterhin gab es nur 1.6 % mit acht oder mehr Wasserstoffbrückenakzeptoren.

Ein weiterer Parameter, der eine ähnliche Aussage trifft, ist die PSA. Deren empfohlene Grenze liegt bei $\leq 140 \text{ \AA}^2$ [153]. Auch hier wurde der Wert größtenteils deutlich unterschritten.

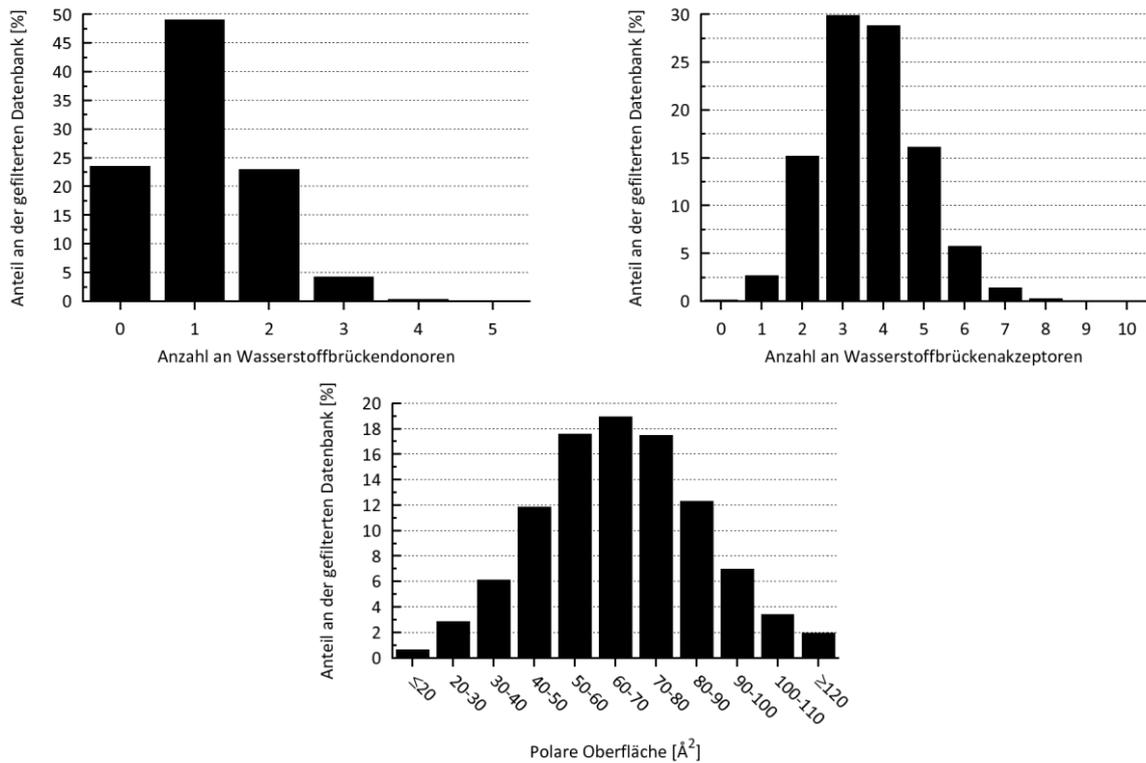


Abbildung 3.13 – Parameter für „hydrophile Interaktionen“. Alle Parameter sind sehr günstig verteilt. Die Maxima liegen deutlich unter den Grenzen für „Drug Like“-Substanzen. Fast die Hälfte aller Strukturen hat ein Donor, drei bis vier Akzeptoren und/oder eine PSA zwischen 50 und 80 Å².

– Lipophilie

Der letzte wichtige Parameter ist die Lipophilie, die mit dem cLogP-Wert beschrieben werden kann (s. Abbildung 3.14). Die Werte waren linksschief mit einem Maximum bei drei bis vier verteilt. Damit lagen die meisten Substanzen in einem guten bis sehr guten Bereich. Auffällig ist jedoch, dass in dem Datensatz nur wenige hydrophile Substanzen vorkamen: Es gab kaum Verbindungen (ca. 0.5 %), deren cLogP-Wert unter 0 lag. Bei der

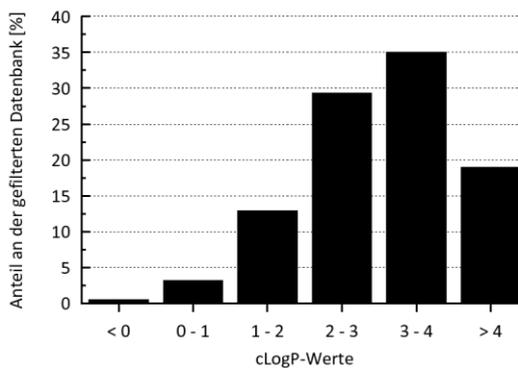


Abbildung 3.14 – cLogP-Werte der gefilterten Datenbank. Die cLogP-Werte sind linksschief verteilt. Das Maximum befindet sich bei Werten von 3 bis 4.

Entwicklung von Leitstrukturen zu Wirkstoffen erhöht sich der cLogP-Wert im Schnitt um 0.43 [45]. Lediglich 7.7 % wiesen einen cLogP-Wert auf, der größer als 4.5 war und damit bei Optimierungsschritten in den kritischen Bereich von 5 gelangen könnte.

– Auswertung

Mithilfe der ZINC-Datenbank und den erarbeiteten Filterkriterien (s. Tabelle 3.4 auf S. 72) wurde ein ausgewogener Datensatz erstellt. Die Ausgangsdatenbank bestand aus einer großen, schon weitestgehend den „Drug-Like“-Kriterien entsprechenden Substanzbibliothek. Durch weitergehende Filter, die vor allem kleinere Moleküle präferierten (insbesondere durch die Anzahl an drehbaren Bindungen und chiralen Zentren), konnte die Datenbank auf mittelgroße Substanzen mit sehr guten „Drug-Like“-Kriterien angereichert werden: Alle wichtigen Kriterien befinden sich auf einem guten Niveau, so dass auch weiter optimierte Hits, bei denen sich die physikochemischen Parameter verschlechtern könnten, noch den Ro5 entsprechen würden.

4 Virtual Screening

Das virtuelle Screening in dieser Arbeit erfolgte über ein Docking. Als virtuelles Docking, oder einfach Docking (engl. für Ankoppeln), wird die Generierung einer Pose einer Substanz (Ligand) in einer Zielstruktur (*Target*, Protein) bezeichnet. Die Pose beschreibt die genaue räumliche Platzierung und Orientierung, welche die Substanz in der Zielstruktur einnimmt. Ziel eines Dockings ist die korrekte, d. h. die in der Realität vorkommende, Vorhersage der Pose einer Verbindung [160, 161].

Handelt es sich um eine aktive Substanz (also eine Substanz, die in der Zielstruktur bindet), müssen während des Dockings zwei Probleme gelöst werden: Zunächst muss die korrekte Pose erzeugt werden (Docking-Problem). Danach muss die korrekte Pose zwischen den möglicherweise zahlreichen falschen Posen identifiziert werden (Scoring-Problem). Nur wenn beide Schritte gelingen, ist der gesamte Docking-Prozess erfolgreich.

Zur Lösung des Docking-Problems ist die Erstellung aller chemisch plausiblen Posen mit, je nach Anwendungsgebiet, möglichst wenig Rechenaufwand nötig. Dafür muss im optimalen Fall sowohl die Substanz an das Protein als auch das Protein an die Substanz angepasst werden. Für die flexible Anpassung der Substanzen an das Protein gibt es verschiedene Methoden. Docking-Methoden werden in systematischen, stochastischen und deterministischen Ansätzen eingeteilt [161, 162]. Der Unterschied besteht darin, wie sie die möglichen Freiheitsgrade der Substanzen untersuchen. Unter dem Begriff des Freiheitsgrads werden die Eigenschaften der Substanzen verstanden, die verändert werden können (wie Rotationen um Bindungen, Torsionswinkel). Bei der systematischen Vorgehensweise werden alle möglichen Freiheitsgrade einer zu dockenden Substanz untersucht. Dazu gehören die erschöpfenden Docking-Algorithmen (bspw. GLIDE [163]), inkrementell aufbauende Algorithmen (FlexX [164]) oder auf Konformerensemble-basierende Algorithmen (Fred [165]). Stochastische Methode versuchen, die Freiheitsgrade zufällig zu variieren und anhand von Wahrscheinlichkeitsberechnungen zu bewerten. Diese basieren bspw. auf Monte Carlo-Simulationen (ICM [166]) oder evolutionären Algorithmen (GOLD [167]). Unter deterministischen Algorithmen werden Energieminimierungen und MD-Simulationen verstanden. Wie schon aus dem Wort deterministisch abgeleitet, bestimmt bei diesen Methoden die Ausgangspose die Endpose. Häufig wird diese Methode mit anderen Algorithmen kombiniert, mit denen Ausgangsposen erstellt werden (z. B. bei GLIDE). Die Performance dieser verschiedenen Docking-Algorithmen wurde schon intensiv untersucht [56, 162, 168, 169]. Die Anpassung der Proteinstruktur an den Liganden wurde für Docking-Experimente lange Zeit vernachlässigt, weil die Vorhersage der Proteinflexibilität sehr komplex und rechenaufwendig ist. Inzwischen gibt es verschiedene Ansätze, die diese Flexibilität berücksichtigen [170–173]. Für ein HTVS eignen sie sich jedoch nicht, da sie sehr rechenaufwendig sind. In dem hier gewählten Ansatz wurden daher vier

starre Proteinstrukturen verwendet, die strukturell stark unterschiedliche, energiegunstige Konformationen der Proteinstruktur darstellten und damit den zur Verfügung stehenden konformativen Raum für das Protein möglichst umfassend beschreiben sollen (s. Abschnitt 2.1 auf S. 27).

Das Scoring-Problem wird an anderer Stelle diskutiert (s. Abschnitt 5 auf S. 100).

4.1 Material und Methoden

4.1.1 cRAISE

Unabhängig von der verwendeten Dockingmethode verarbeiten die meisten Algorithmen jede Substanz einzeln. Dies ist für sehr große Datensätze, wie sie bei diesem HTVS vorlagen, problematisch. Die Substanzbibliothek bestand aus ca. 3.4 Mio. Substanzen (s. Abschnitt 3.4 auf S. 73) und sollte in vier verschiedene Proteinstrukturen gedockt werden (s. Abschnitt 2.1 auf S. 27). Würde jede Substanz innerhalb einer Sekunde^{viii} erfolgreich gedockt und bewertet werden, dauerte das Docking über 150 Tage. Die Arbeitsgruppe um RAREY entwickelte eine nicht-sequentielle Methode zur Durchführung eines Screenings, die sich besonders für sehr große Datensätze eignet [175–177]. Anstatt jede Substanz einzeln zu docken, werden nur diejenigen Substanzen ausgewählt, die das Pharmakophor in der Zielstruktur erfüllen können. Dazu wird zuerst eine Datenbank mit den Struktureigenschaften der Substanzen erstellt. Beim anschließenden Docking-Vorgang selbst wird nicht jede einzelne Substanz verarbeitet, sondern nur die Substanzen, die anhand einer aus der Zielstruktur erstellten Abfrage ausgewählt wurden. Dadurch erfüllen sie schon wichtige Voraussetzung für eine gute Bindung in der Bindetasche. Im Detail verläuft ein cRAISE-Docking wie folgt:

Im ersten Schritt wird ein Substanz-Index der zu dockenden Moleküle erstellt. Dafür werden von allen Substanzen physikochemische Eigenschaften gespeichert und energiegunstige Konformere erzeugt. Letztere werden von CONFECT wissensbasiert generiert [178]. Für die weitere Verarbeitung stellt jedes einzelne Konformer eine eigene Substanz dar, die einzeln behandelt wird. Daher handelt es sich bei cRAISE auch um einen konformerensemble-basierenden Algorithmus: alle Konformere werden vor der eigentlichen Platzierung generiert. Für diese Substanzen werden daraufhin alle möglichen cRAISE-Deskriptoren erstellt. Ein cRAISE-Deskriptor stellt ein Interaktionsmuster aus drei Interaktionspunkten einer Substanz dar (bspw. Wasserstoffbrückenbindungsdonor, -akzeptor und hydrophobe Bereiche). Hierbei werden mithilfe eines *Valence State Combination Models* die wahrscheinlichsten Protonierungs- und Tautomerisierungszustände für jede Substanz generiert [179] und bei der Deskriptorberechnung berücksichtigt. Kann ein Atom anhand dieser Zustände beispielsweise sowohl Wasserstoffbrückenbindungsdonor als auch –akzeptor sein, werden hier alle möglichen Deskriptoren

^{viii} In einer vergleichenden Docking-Studie von 2006 benötigten verschiedene Docking-Tools zwischen 50 und > 300 s pro Liganden [174].

gebildet, die an der entsprechenden Position entweder eine Wasserstoffbrückendonator oder einen -akzeptor haben. Handelt es sich bei den Interaktionspunkten um gerichtete Interaktionen (z. B. Wasserstoffbrücken), wird deren Ausrichtung zusätzlich über einen Vektor beschrieben. Die drei Interaktionspunkte im Raum stellen ein Dreieck dar, das mögliche Interaktionen der Substanz beschreibt. Weiterhin werden auch noch sterische Parameter der Substanz gespeichert.

Im zweiten Schritt wird die Zielstruktur (Protein) vorbereitet. Ebenso wie bei den Substanzen werden cRAISE-Deskriptoren, also ebenfalls Interaktionsdreiecke, für die Bindetasche berechnet, sowie die Protonierungs- und Tautomerisierungszustände. Die Größe der Bindetasche wird mithilfe eines bekannten Liganden definiert. Alle Aminosäuren, die sich in einem Radius von 6.5 Å um ein Ligandatombefinden, werden dafür mit einbezogen.

Im dritten Schritt findet die Abfrage der cRAISE-Deskriptoren im Substanz-Index statt. Dafür werden die cRAISE-Deskriptoren der Zielstrukturen generiert und alle komplementären Deskriptoren (Wasserstoffbrückendonator zu -akzeptor) aus dem Substanz-Index angefragt. Erfüllt eine Substanz keinen cRAISE-Deskriptor der Abfrage, wird sie auch nicht für die Platzierung ausgewählt. Dies stellt einen großen Zeitgewinn bei dem Docking-Vorgang dar, da Substanzen ignoriert werden, wenn sie keine Möglichkeit haben, ein Interaktionsdreieck zu erfüllen. Es werden nur die Substanzen, genauer die dazugehörigen Konformere, Protonierungs- und Tautomerisierungszustände, weiter verarbeitet, die zumindest ein gutes Interaktionsmuster aufweisen könnten.

Im vierten Schritt werden die zuvor ausgewählten Substanzen in der Zielstruktur anhand der cRaisedeskriptor-Dreiecke platziert. Entstehen dabei grobe Überschneidungen der vdW-Radien, werden diese Verbindungen wiederum herausgefiltert. Anschließend findet eine Optimierung des Wasserstoffbrückennetzwerkes mit PROTOSS statt [180]. Dabei können sich Tautomerisierungs- und Protonierungszustände noch einmal ändern. Nach dem Scoring und Ranking der Posen mit einer internen empirischen Scoring-Funktion, wird eine Hit-Liste erstellt. Die Liste enthält von jeder ursprünglichen Substanz nur den bestbewerteten Zustand unter allen generierten Konformationen, Tautomeren und Protonierungszuständen.

Zusätzlich können Pharmakophore definiert werden, die von den Posen erfüllt werden müssen. Diese Beschränkungen werden sowohl bei der Anfrage der Deskriptoren in der Substanz-Datenbank benutzt also auch bei der finalen Optimierung mit PROTOSS. Auf diese Weise wird sichergestellt, dass nur Posen generiert werden, die das gewünschte Pharmakophor-Muster erfüllen.

Der Substanz-Index stellt eine wichtige Komponente für diese Methode dar. Dessen Erstellung ist rechenaufwendig. Dies ist insbesondere für Substanzen mit einer hohen Anzahl an drehbaren Bindungen problematisch, da die Menge an Konformeren stark zunimmt. Daher wurde die Zahl der rotierbaren Bindungen bei der Bibliothekserstellung der zu dockenden Substanzen eingeschränkt (s.

Abschnitt 3.3 auf S. 66). Der große Vorteil des Substanz-Indexes besteht darin, dass er, einmal erstellt, immer wieder zum Docking verwendet werden kann. Dies war für das hier gewählte Vorgehen sehr effektiv, da die Substanz-Bibliothek in vier verschiedene Zielstrukturen gedockt werden sollte.

Nach der Platzierung der Pose in die Bindetasche findet keine Optimierung oder Refinement in Bezug auf die Bindetasche statt. Daher sind hier zu einer gewissen Toleranz sterisch-negative Effekte (vdW-Clashes) möglich. Dies ist bis zu einem bestimmten Maße sogar ein gewünschter Effekt, da kleine Clashes die Flexibilität des Proteins repräsentieren können. Das Auftreten der vdW-Überschneidung kann reduziert werden, indem die Anzahl der Konformere erhöht wird. Dadurch nimmt jedoch der Rechenaufwand bei der Erstellung des Indexes und der Platzierung wieder zu. Da es sich bei cRAISE um ein HTVS-Tool handelt, ist der Fokus auf die Effizienz, zugunsten der Genauigkeit, beim Docking gesetzt. Dies ist insbesondere für das Postprocessing Prozess (s. Abschnitt 5 auf S. 100) relevant.

Standardmäßig werden alle Posen herausgeschrieben, die einen negativen Score-Wert besitzen, so dass je nach Ansatz entschieden werden kann, welche Docking-Ergebnisse weiter verarbeitet werden. Dadurch hat es zunächst den Anschein, dass cRAISE viele unspezifische Docking-Ergebnisse erzeugt. Eine definierte Score-Grenze für die Bewertung der Ergebnisse ist deshalb essentiell.

4.1.2 Evaluierungsmethoden

Die Bewertung der Qualität von Docking-Verfahren ist ein kontrovers diskutiertes Thema [168, 181–184]. Für die Optimierung des Dockings in diesem Abschnitt müssen sehr große Datenmengen, d. h. mehrere Millionen Docking-Ergebnisse, ausgewertet werden. Daher wird ein Fokus auf die Evaluierung des Rankings gelegt. Dabei wird ausschließlich beurteilt, ob eine aktive Substanz besser bewertet wird als eine inaktive Substanz. Es findet keine direkte Bewertung der Qualität der Pose statt, z. B. ob die richtige Pose generiert wird. Für die Bewertung des Rankings eignen sich Enrichment- und ROC-Plots.

– Enrichment-Plots

Bei Enrichment-Plots wird der Anteil der gescreenten Gesamtdatenbank gegen den Anteil der gefundenen Aktiven aufgetragen. Bei dieser Form der Auswertung besteht eine Abhängigkeit zwischen dem Anteil der aktiven Substanzen und der Güte der Anreicherung. Datenbanken mit sehr wenigen Aktiven zeigen eine deutlich bessere Anreicherung als mit vielen Aktiven (s. Abbildung 4.1). Dieser Effekt tritt insbesondere beim Vergleich von verschiedenen Datensätzen auf [185]. Während eines HTVS liegt der Hauptfokus auf dem frühen Enrichment, da aufgrund der großen Datenmenge nur ein kleiner Teil der gefundenen Posen weiterverarbeitet werden kann. Daher wird häufig eine logarithmische Auftragung gewählt, um die wichtigen Bereiche (bis ca. 10 %) optisch zu vergrößern. Zur Orientierung sind in den folgenden Abbildungen die beiden Grenzen, perfektes Enrichment (oben links) und zufällige Verteilung (unten rechts), grau dargestellt (s. Abbildung 4.2).

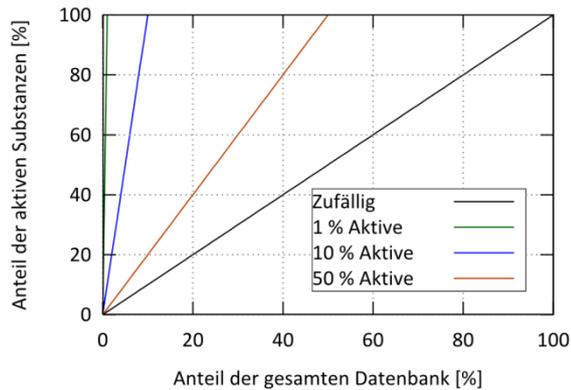


Abbildung 4.1 – Enrichment-Plots in Abhängigkeit von verschiedenen Anteilen an Aktiven im Vergleich zu einer zufälligen Verteilung (schwarz). Für jeden Anteil ist das perfekte Enrichment dargestellt (alle Aktive sind vor den Inaktiven platziert). Je weniger Aktive im Datensatz vorhanden sind, desto besser scheint das Enrichment, obwohl es in allen Fällen gleich ist.

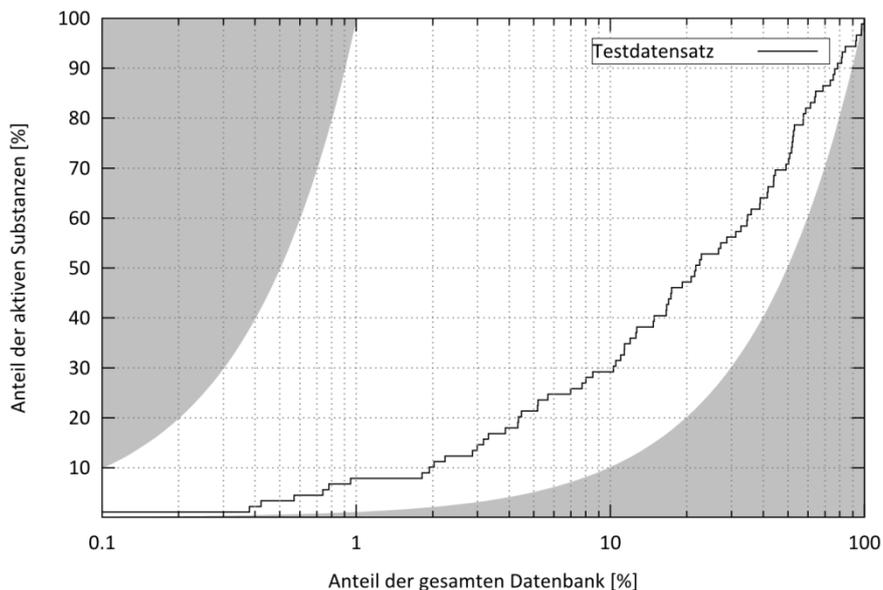


Abbildung 4.2 – Beispiel eines Enrichments. Die Grenzen, perfektes Enrichment und zufällige Auswahl, sind grau markiert. Durch die logarithmische Auftragung kann insbesondere die frühe Anreicherung genauer untersucht werden.

Aus dem Enrichment-Plot lässt sich auch der Enrichmentfaktor ($EF_{x\%}$) ableiten. Dieser beschreibt den Faktor, um den ein definierter Teil der Datenbank angereichert wird ($EF_{5\%} = 10$ bedeutet, dass in den TOP 5 % der gerankten Datenbank der Anteil Aktiver zehnmal höher ist als in der Ausgangsdaten-

bank). Berechnet wird der Enrichmentfaktor für $x\%$ mit der Anzahl aller Aktiven ($a_{100\%}$) und der gefundenen Aktiven ($a_{x\%}$) in diesem Bereich (s. Formel 4.1):

$$EF_{x\%} = \frac{a_{x\%}}{a_{100\%} \cdot x\%} \quad \text{Formel 4.1}$$

– ROC-Plots

Eine ähnliche Aussage trifft die „Receiver Operating Characteristic“-Kurve (ROC-Kurve). Bei dieser Kurve wird die richtig-positiv Rate (TPR, *sensitivity*) gegen die falsch-positiv Rate (FPR, *1-specificity*) aufgetragen. Die Auswertung hat den Vorteil, dass sie, anders als der Enrichment-Plot, unabhängig vom Verhältnis der Aktiven zu Inaktiven ist. Für den Datensatz der Abbildung 4.2 wurde in Abbildung 4.3 die ROC-Kurve erstellt. Je steiler die Kurve ist, d. h. je schneller eine richtig-positiv-Rate von 1 erreicht wird, desto besser ist die Anreicherung.

Anhand des ROC-Plots lässt sich auch ein Parameter für die Qualität des Rankings ableiten. Die Fläche unter der Kurve (*AUC*) beschreibt die Wahrscheinlichkeit dafür, dass ein zufällig ausgewählter Aktiver höher priorisiert wird als ein zufällig ausgewählter Inaktiver: Beträgt die Wahrscheinlichkeit eins, ist jeder Aktive höher bewertet als jeder Inaktive, bei 0.5 ist der Datensatz zufällig gemischt, bei null ist jeder Aktive schlechter bewertet als jeder Inaktive. Neben der Bestimmung des Integrals kann die *AUC* auch mit der Anzahl der Aktiven (a), der Anzahl der besser als die Aktiven gerankten Inaktiven (d_i) und der Gesamtanzahl der Substanzen (n) berechnet werden (s. Formel 4.2) [186].

$$AUC = \frac{1}{a} \cdot \sum_{i=1}^a \left(1 - \frac{d_i}{n}\right) \quad \text{Formel 4.2}$$

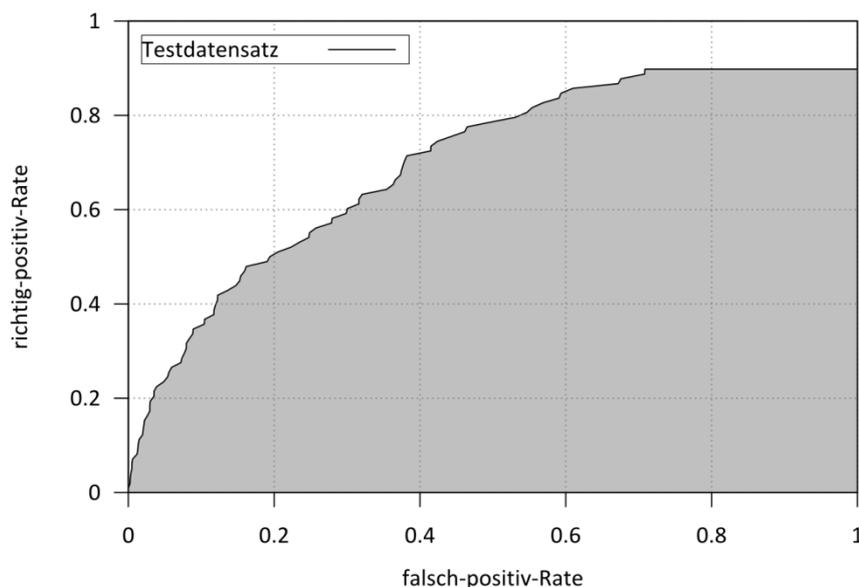


Abbildung 4.3 – Beispiel einer ROC-Kurve. Hierbei handelt es sich um den gleichen Datensatz wie in Abbildung 4.2. Die *AUC* (grau) beträgt 0.697.

Die größten Defizite dieser Bewertungsmethoden liegen darin, dass sie nicht in der Lage sind, Parameter für die Bewertung einer frühen Anreicherung zu liefern. Dies gilt insbesondere für Enrichment-Faktoren, die nur die Zahl der Aktiven innerhalb eines bestimmten Bereichs zählen und nicht die Platzierung selbst berücksichtigen. Dies ist für die Auswertung eines HTvS von großer Bedeutung, da nur eine geringe Zahl der Docking-Ergebnisse in den nächsten Schritten weiterverarbeitet werden kann und ein großer Teil der Ergebnisse verworfen wird. Insbesondere die Abhängigkeit der Enrichment-Plots von der Zusammensetzung des Datensatzes ist ein großer Nachteil [182]. Der große Vorteil der Enrichment-Plots liegt andererseits darin, dass sie intuitiv verstanden werden und, bei gleichen Datensätzen, auch visuell schnell interpretiert werden können.

Bei dem in dieser Arbeit verwendeten Vorgehen wurden die Docking-Experimente an einem vorher erstellten Testdatensatz optimiert (s. u.). Da dieser im Laufe der Versuche nicht verändert wurde, war die Analyse der gewählten Parameter anhand der Enrichment Plots möglich [181].

4.1.3 Directory of Useful Decoys – Enhanced

Das „*Directory of Useful Decoys – Enhanced*“ (DUD-E) ist eine Datenbank für das Benchmarking von Screening-Prozessen. Sie besteht aus aktiven und inaktiven Substanzen; die Inaktiven werden Decoys (engl. für Köder oder Falle) genannt. Dabei sind die Decoys so gewählt, dass sie ähnlich zu den Aktiven sind, aber dennoch so unterschiedlich, dass sie nicht an dem Enzym binden sollten [187]. Der DUD-E-Datensatz stellt eine Optimierung des DUD-Datensatz [188] dar, der noch einige Probleme enthielt (z. B. zu geringe Diversität [189], aktive Decoys [190]).

Für den DUD-E-Datensatz werden von den aktiven Liganden zunächst die möglichen Protonierungszustände in einem pH-Bereich von sechs bis acht berechnet. Anschließend wird in der ZINC-Datenbank für jeden Protonierungszustand nach ähnlichen Strukturen gesucht, die anhand der Molekülmasse, dem cLogP-Werten, Anzahl an drehbaren Bindungen, Wasserstoffbrückenakzeptoren, -donoren und der Gesamtladung bestimmt werden. Diese werden mittels Extended-Connectivity-Fingerprints (ECFP) gegenüber den aktiven Substanzen geclustert und die 25 % unähnlichsten Decoys behalten. Schließlich werden 50 Decoys für jeden Protonierungszustand zufällig ausgewählt [187, 188].

Dieses Protokoll wurde für eine Reihe von Targets mit entsprechenden aktiven Substanzen durchgeführt und die entsprechenden Decoy-Sets stehen für Benchmarkings zur Verfügung. Zusätzlich bietet die Arbeitsgruppe um SHOICHET eine webbasierte Plattform für die Generierung eines Decoy-Datensatzes mit diesem Protokoll aus einem eigenen Datensatz an^{ix}.

^{ix} <http://dude.docking.org/generate>; zuletzt aufgerufen: Januar 2015.

Für diese Optimierung wurde ein eigener Decoy-Datensatz mit 95 literaturbekannten aktiven Substanzen verwendet [35, 73, 74, 78, 85, 86, 88, 92–94, 191–194] (s. Anhang Tabelle 12.1 auf S. 231). Bei der Auswahl der aktiven Substanzen wurde auf die chemische Diversität und hohe Aktivität geachtet. Der so erhaltene DUD-E-Datensatz bestand aus 6,295 Decoys.

Die aktiven Substanzen wurden zusammen mit dem DUD-E-Datensatz zum DUD-E Docking-Testdatensatz vereint (6,390 Strukturen).

4.2 Docking mit cRAISE

Die Docking-Experimente und Variationen der Parameter für diese wurden in Kooperation mit Angela Henzler^x und Therese Inhester^x durchgeführt.

Zunächst wurde ein cRAISE -Docking mit dem DUD-E Docking-Testdatensatz und, nach Optimierung der Parameter, der gefilterten ZINC-Datenbank durchgeführt. cRAISE wurde bisher noch nicht mit einem Datensatz dieser Größe und physikochemischer Eigenschaften (s. Abschnitt 3.4 auf S. 73) getestet. Daher war nicht bekannt, wie die Docking-Ergebnisse ausfallen würden und wie stark einzelne Parameter die Ergebnisse beeinflussen (z. B. Qualität oder Anzahl der Posen). Entsprechend der Evaluation der Dockings sollten Kriterien erarbeitet werden, um den Datensatz weiter zu reduzieren. Anschließend sollte mit diesem reduzierten Datensatz ein zweites Docking mit cRAISE durchgeführt werden, bei dem durch die Verwendung von deutlich mehr Konformeren pro Substanz (optimierter Substanz-Index) Posen besserer Qualität erhalten werden.

4.2.1 Erstellung der Pharmakophore

cRAISE ist, wie oben beschrieben, ein pharmakophor-gelenktes Docking-Tool. Die Interaktionen, die cRAISE zum Docken verwendet, können auch manuell angepasst und unterschiedlich gewichtet werden. Dadurch favorisiert cRAISE die entsprechenden Interaktionen und kann Posen, die nicht den Anforderungen entsprechen, direkt herausfiltern. Für das Docking wurden essentielle Interaktionen definiert, die ein mögliche Pose mindestens erfüllen muss: die Wechselwirkung mit Asp73 als Wasserstoffbrückenbindungsakzeptor und dem Wat A als Wasserstoffbrückenbindungsdonor (s. Abschnitt 2.3.2 auf S. 50).

Die Erstellung des Pharmakophors läuft unabhängig vom Docking ab. Um das Pharmakophor bestmöglich zu beschreiben, wurden verschiedene Einstellungen für die Größe der Sphären und deren Orientierung getestet. Eine Sphäre beschreibt den Bereich, in dem sich das entsprechende Atom für die definierte Interaktion befinden darf. Der Einfluss der Größe der Sphären war bei den verschiedenen Dockings nur gering: Weder größere noch kleinere Sphären veränderten die Anreicherungen entscheidend. Einerseits schränken aber zu kleine Sphären den gesamten Docking-Vorgang ein, andererseits führen zu große Sphären wiederum zu unspezifischen Ergebnissen. Daher wurde ein mittlerer Sphärenradius von 1.0 Å gewählt.

Beim Docking in die 1E11-Proteinstruktur wurde der tiefe P-Bereich sehr häufig angesprochen. Da bekannte Inhibitoren vornehmlich im Bereich des Adeninkörpers banden und der P-Bereich von verschiedenen Resten flexibel besetzt werden konnte (s. Abschnitt 2.1 auf S. 27), wurde dieser Bereich für das Docking eingeschränkt. Dafür wurden zwei *Exclusion Volumes* in den P-Bereich platziert. Das

^x Arbeitskreis von Matthias Rarey Zentrum für Bioinformatik Universität Hamburg.

sind Sphären, in deren Bereich während des Dockings kein schweres Atom platziert werden darf (s. Abbildung 4.4, schwarze Sphären). Als weitere Definitionen wurden Asp73 und Wat A verwendet, die in ihrer Funktion als Wasserstoffbrückenbindungsakzeptor bzw. -donor zum essentiellen Pharmakophor gehörten. Aus Sicht des Liganden wurde die Sphäre des Wasserstoffbrückenbindungsakzeptors auf das N1 des Adenins und die Sphäre des Wasserstoffbrückenbindungsdonors auf das primäre Amin des Adenins platziert.

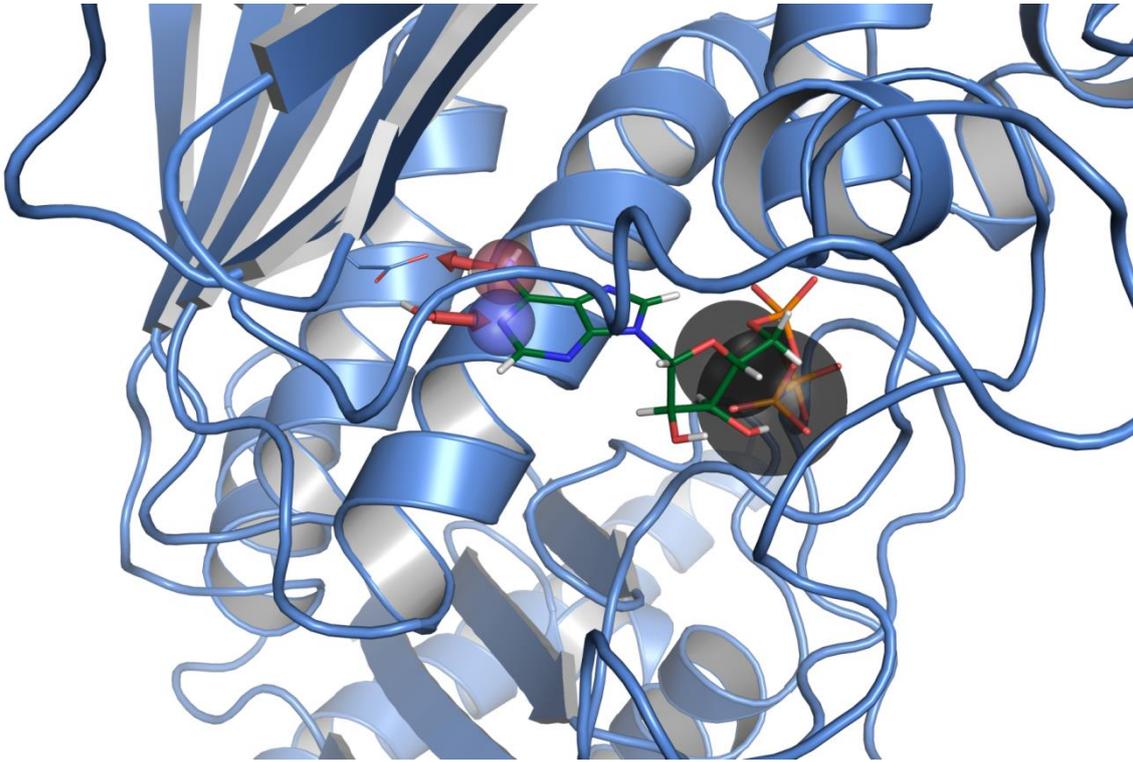


Abbildung 4.4 – cRAISE-Pharmakophor für 1E11. Die im Docking gesuchten Wasserstoffbrückenbindungsdonor-Komponenten des Liganden sind rot, Wasserstoffbrückenbindungsakzeptor-Komponenten rot und *Exclusion Volume* schwarz dargestellt. Die Pfeile (rot) geben die Richtung des jeweiligen Wasserstoffbrückenbindungsdonors an.

Für die Definition des Pharmakophors für 1AJ6 wurde ein modifizierter Ligand und nicht Novobiocin selbst verwendet. Novobiocin bindet in einem weiten Bereich der Bindetasche. Die Interaktion innerhalb der P-Tasche ist nicht sehr ausgeprägt, dafür zeigt es eine Bindung bis zum Loop A (s. Abschnitt 2.1 auf S. 27). Bei dem cRAISE-Docking konnte keine Posen generiert werden, wenn die Bindetasche nur durch Novobiocin definiert wurde und die essentiellen Interaktionen mit Asp73 und Wat A vorgegeben wurden. Das Problem bestand darin, Posen zu erzeugen, die auch tief in der Tasche banden. cRAISE erzeugt die Bindetasche mithilfe eines 6.5 Å-Radius (s. Abschnitt 4.1.1 auf S. 80). Dies ist aufgrund der Bindeeigenschaften für diesen Liganden nicht ausreichend. Daher wurde Novobiocin modifiziert und formal über das Amin mit dem C3-Atom des Pyrazols des Liganden von 3G7E verknüpft (s. Abbildung 4.5). Dies hatte keine direkte Auswirkung auf die erstellten Pharmakophore und den Ablauf des Dockings selbst, es erweiterte jedoch die räumliche Ausdehnung bei der Definition der Bindetasche in Richtung der tiefen Bindetasche.

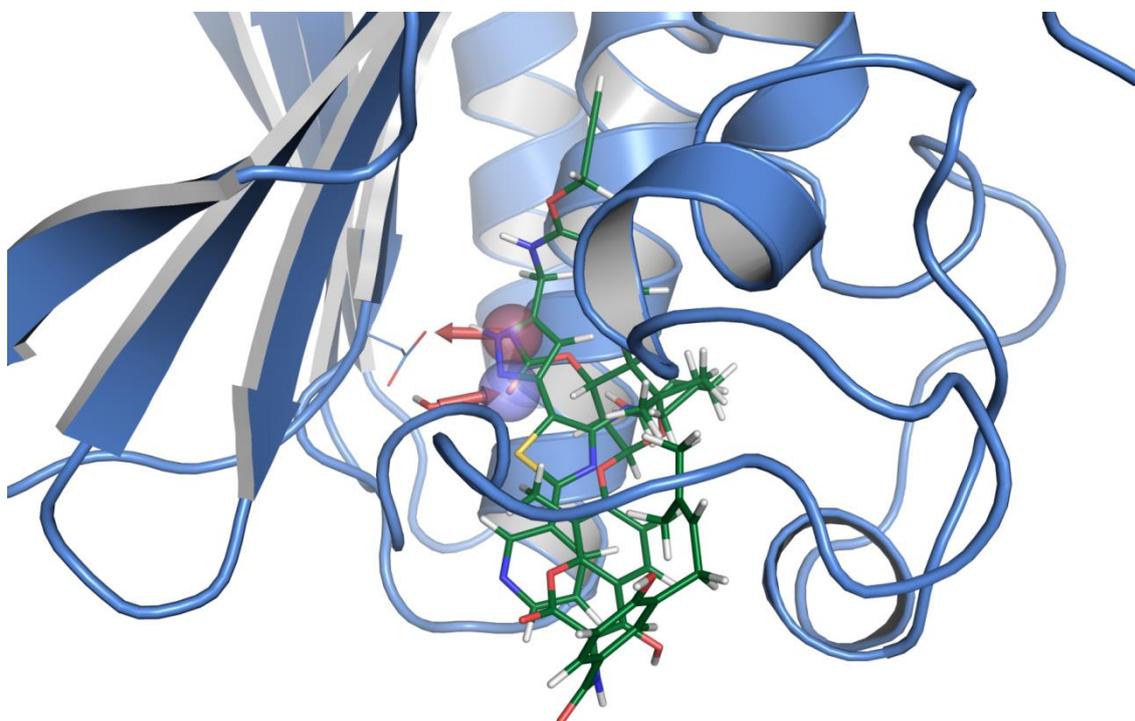


Abbildung 4.5 – cRAISE-Pharmakophor für 1AJ6. Der Ligand besteht aus dem formal miteinander verknüpften Novobiocin und Pyrazolderivat von 3G7E. Die im Docking gesuchten Wasserstoffbrückenbindungsdonor-Komponenten des Liganden sind rot und Wasserstoffbrückenbindungsakzeptor-Komponenten blau. Die Pfeile (rot) geben die Richtung des jeweiligen Wasserstoffbrückenbindungsdonors an.

Bei den Strukturen von 3G7E (3G7E_1 und 3G7E_2) bestand die Problematik bei der Definition der Bindetasche darin, dass Wasserstoffbrückenbindungsdonor und -akzeptor des essentiellen Pharmakophors direkt nebeneinander positioniert waren (die beiden Stickstoffatome des Pyrazolringes). Die Sphären besaßen einen Radius von ca. 1.0 Å und die Entfernung zwischen den bei-

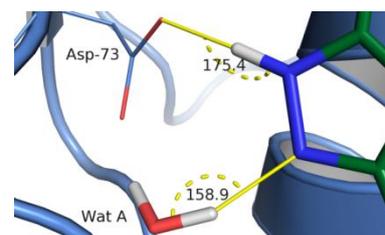


Abbildung 4.6 – Geometrie des essentiellen Pharmakophors von 3G7E.

den Atomen betrug ungefähr 1.4 Å. Daher überlagerten sich die beiden Pharmakophordefinitionen, wenn sie auf den beiden Stickstoffatomen platziert wurden. In Abbildung 4.6 sind die Winkel für Donor–H···Akzeptor dargestellt. Der optimale Winkel beträgt 180 ° [99], für die Interaktion vom Asp73 mit dem Liganden ist er hier 175.4 °, Wat A bildet einen Winkel von 158.9 mit dem Liganden aus. Die Geometrie war für die Interaktion mit dem Asp73 also deutlich besser. Weiterhin befand sich dort auch der C=O···H-Winkel mit 121.7 ° nahe am Optimum von 120 ° [99]. Daher wurde die Akzeptorsphäre um 0.8 Å in der Verlängerung der Bindung zwischen den beiden Stickstoffatomen verschoben, so dass Donor–H···Akzeptor-Winkel 165.2 ° zum Zentrum der Sphäre betrug. Der Abstand zwischen den beiden Sphären war 2.2 Å, damit zwischen den beiden Interaktionssphären noch ein kleiner Abstand von 0.2 Å bestand und sich das Stickstoff des Pyrazols, das als Akzeptor fungierte, trotzdem in der Sphäre befand. Für das Docking in 3G7E_1 bildeten diese beiden Sphären die essentiellen Interaktionen (s. Abbildung 4.7).

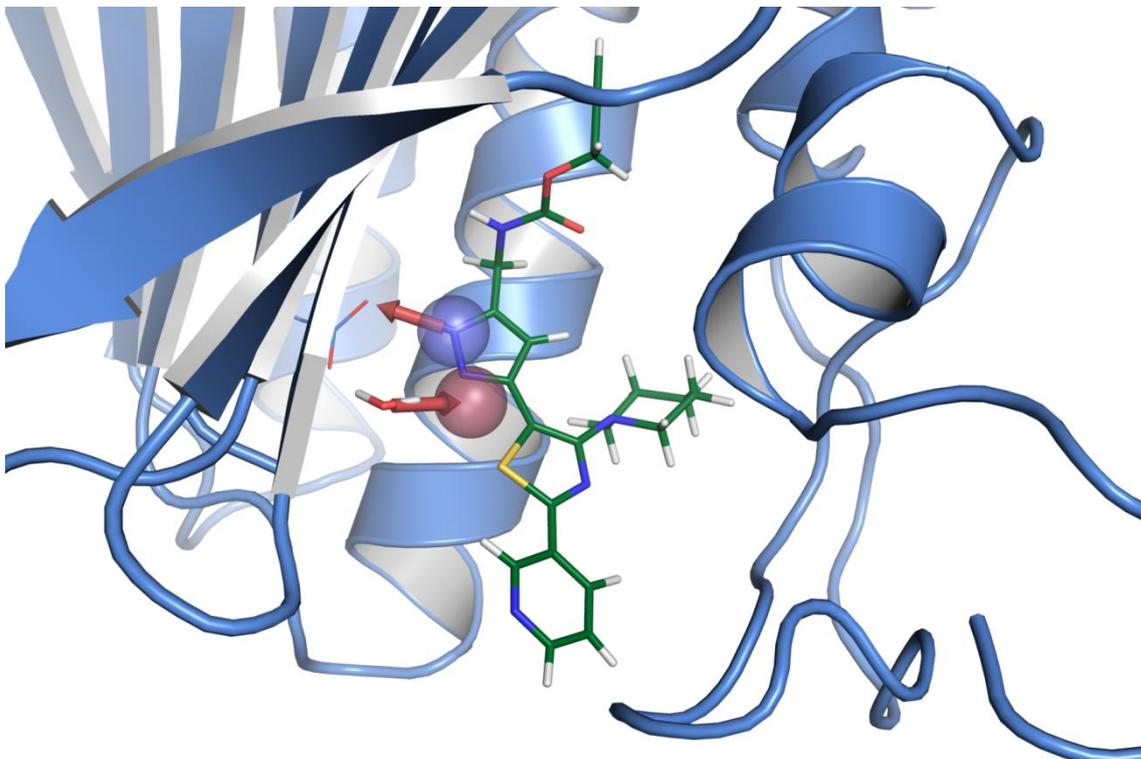


Abbildung 4.7 – cRAISE-Pharmakophor für 3G7E_1. Akzeptor- und Donorsphären überlagerten sich, wenn sie auf die Zentren der entsprechenden schweren Atome des Pyrazols gelegt wurden. Die Akzeptorsphäre wurde daher, auch um bessere Geometrien zu ermöglichen, um 0.8 Å verschoben. Die im Docking gesuchten Wasserstoffbrückenbindungsdonor-Komponenten des Liganden sind rot und Wasserstoffbrückenbindungsakzeptor-Komponenten blau. Die Pfeile (rot) geben die Richtung des jeweiligen Wasserstoffbrückenbindungsdonors an. Loop B wurde aus Übersichtsgründen ausgeblendet.

Für das Docking mit 3G7E_2 wurden die beiden Interaktionssphären von 3G7E_1 übernommen, da die Strukturen bis auf das Wat B identisch sind. Zusätzlich wurde jedoch noch eine weitere Sphäre hinzugefügt, welche die Interaktion mit dem Wat B beschreibt. Dafür wurde eine weitere Wasserstoffbrückenbindungsakzeptor-Sphäre mit der Position des Stickstoffs des Carbaminsäureesters defi-

niert. Diese musste beim Docking zusätzlich erfüllt werden (s. Abbildung 4.8). Diese Sphäre sollte hauptsächlich eine Einschränkung des Dockings darstellen. Die Konservierung des Wassers war unsicher (s. Abschnitt 2.3.2 auf S. 50). Würde ein Ligand jedoch in 3G7E_2 ohne Beteiligung des Wat B binden, wäre die Pose auch in 3G7E_1 wiederzufinden. Damit diene die Definition der dritten Interaktion hauptsächlich dem Zweck, neue Posen zu generieren, die in dem Docking mit 3G7E_1 nicht erzeugt werden konnten. Daher werden bei diesem Docking auch deutlich weniger erfolgreich gedockte Liganden und Posen erwartet. Das Docking mit 3G7E_2 ist als Erweiterung des Dockings mit 3G7E_1 zu verstehen.

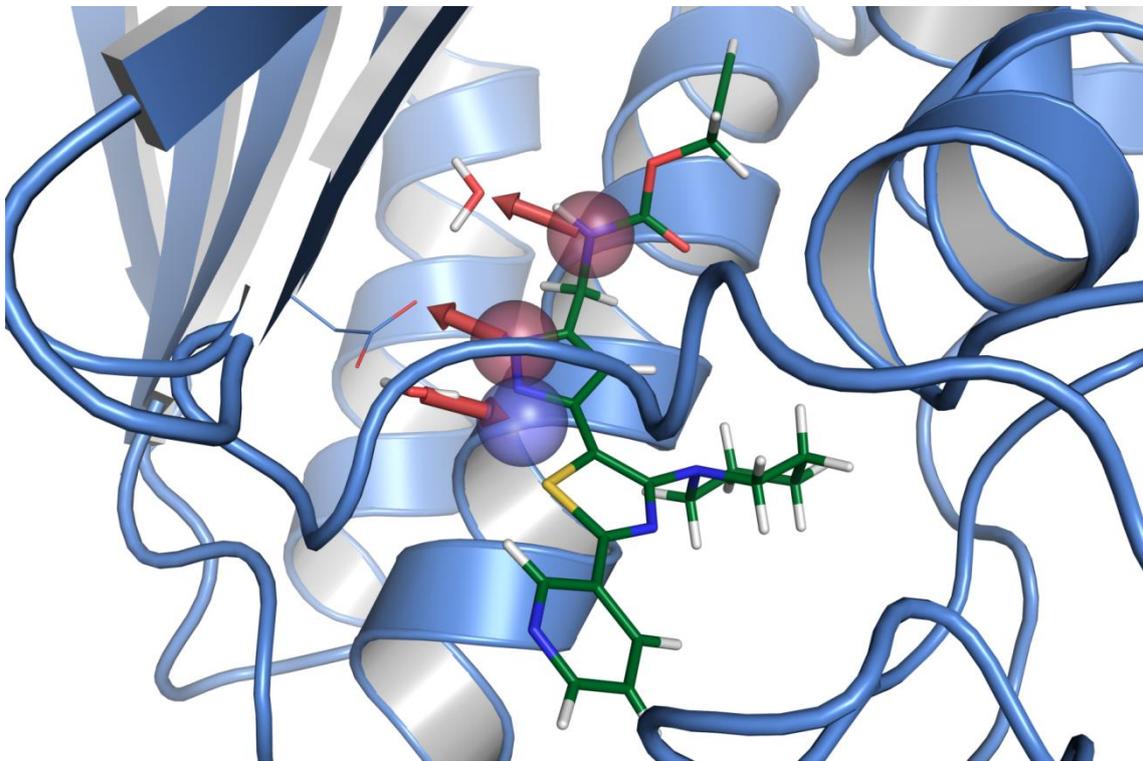


Abbildung 4.8 – cRAISE-Pharmakophor für 3G7E_2. Die im Docking gesuchten Wasserstoffbrückenbindungsdonor-Komponenten des Liganden sind rot und Wasserstoffbrückenbindungsakzeptor-Komponenten blau. Die Pfeile (rot) geben die Richtung des jeweiligen Wasserstoffbrückenbindungsdonors an.

4.2.2 Erstes cRAISE-Docking

– cRAISE Docking des DUD-E Docking-Testdatensatzes

In Abbildung 4.9 sind das Enrichment und der ROC-Plot vom cRAISE-Docking in 1E11 mit dem DUD-E Docking-Testdatensatz dargestellt. Die erste Anreicherung war besser als eine zufällige Verteilung. Der ROC-Plot zeigte eine AUC von 0.676. Eine visuelle Analyse der Posen ergab, dass trotz der *Exclusion Volumes* viele Posen generiert wurden, welche den hydrophilen Abschnitt in der P-Tasche bevorzugten.

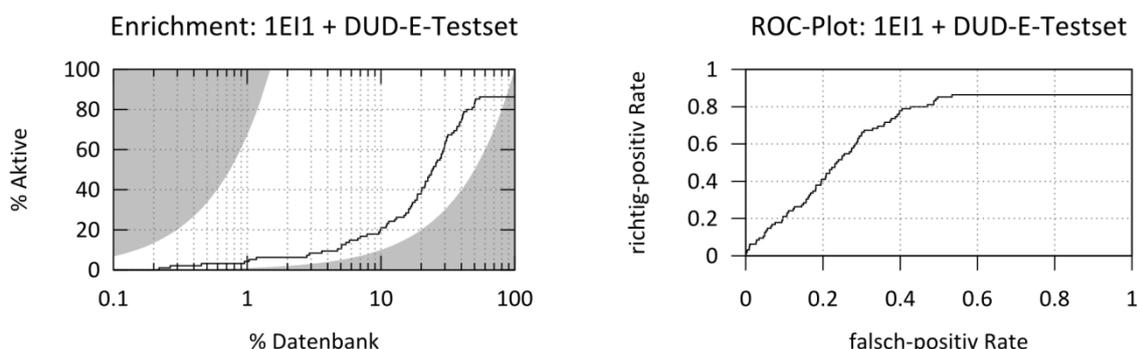


Abbildung 4.9 – Links: Enrichment des cRAISE-Dockings in 1E11 mit dem DUD-E Docking-Testdatensatz. Rechts: ROC-Plot des gleichen Dockings. Die AUC beträgt 0.676.

Das Docking des DUD-E Docking-Testdatensatz von 1AJ6 wies zu Beginn, bis ca. 1 %, ein schlechtes Enrichment auf. Erst im weiteren Verlauf nahm das Enrichment deutlich zu. 98 % der Aktiven konnten insgesamt wieder in die Substanz gedockt werden. Der ROC-Plot zeigte die gleichen Tendenzen. Insgesamt war die AUC mit 0.698 auf einem vergleichbaren Niveau wie bei 1E11.

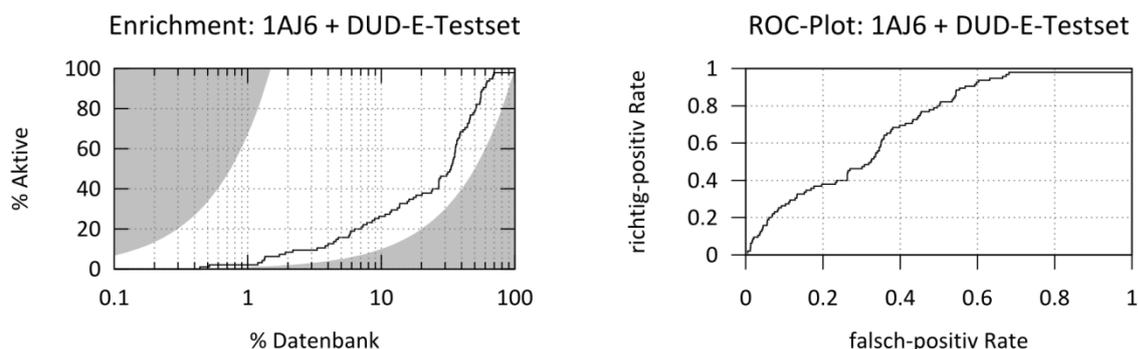


Abbildung 4.10 – Links: Enrichment des cRAISE-Dockings in 1AJ6 mit dem DUD-E Docking-Testdatensatz. Rechts: ROC-Plot des gleichen Dockings. Die AUC beträgt 0.698.

3G7E_1 (s. Abbildung 4.11) zeigte, ähnlich wie die anderen Dockings, zu Beginn nur eine geringe Anreicherung. Diese stieg aber im weiteren Verlauf deutlich an. Bei 10 % der Datenbank wurden ca. 36 % der Aktiven gefunden. Dies zeigte ein deutlich bessere Anreicherung als es bei den bisher vorgestellten Dockings der Fall war.

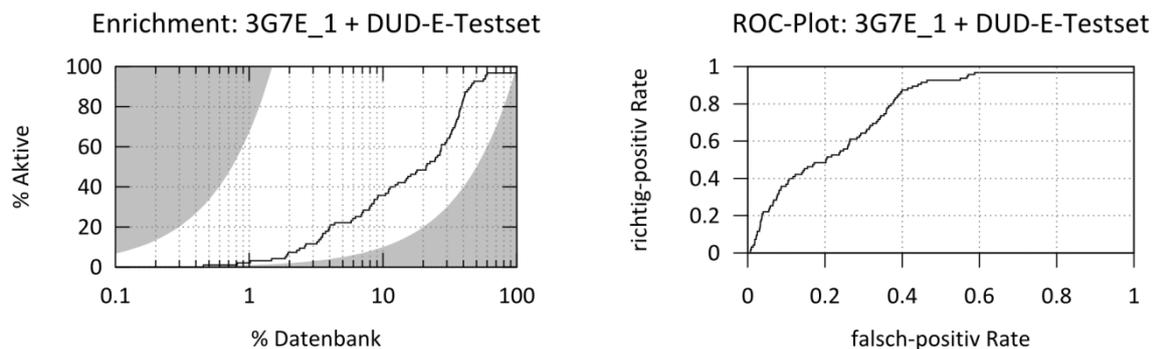


Abbildung 4.11 – Links: Enrichment des cRAISE-Dockings in 3G7E_1 mit dem DUD-E Docking-Testdatensatz. **Rechts:** ROC-Plot des gleichen Dockings. Die AUC beträgt 0.768.

Beim Docking des DUD-E Docking-Testdatensatz in 3G7E_2 unterschieden sich die Ergebnisse deutlich zu den bisher vorgestellten (s. Abbildung 4.12). Das frühe Enrichment war besser als bei allen anderen Dockings. Dennoch war die Gesamtanreicherung deutlich schlechter als bei den bisherigen Dockings. Dies wurde auch aus der AUC des ROC-Plots ersichtlich, die mit 0.416 erstmalig einen Wert unter 0.5 annahm. Dies entspräche einer Anreicherung, die schlechter als eine zufällige Verteilung wäre. Bei diesem Docking handelte es sich jedoch um einen besonderen Fall: Die vorgegebenen Interaktionen, die zum Docking verwendet wurden, können nicht von allen Aktiven erfüllt werden (s. o.). Daher war der Ausschluss von vielen Aktiven auch korrekt. Das Ergebnis ist also trotzdem als positiv zu bewerten: Die Aktiven, die dieses Pharmakophor erfüllten, konnten erfolgreich und früh angereichert werden. Das Docking mit 3G7E_2 war (wie oben schon erwähnt) von vornerein als Erweiterung des 3G7E_1-Dockings ausgelegt und wird dieser Anforderungen gerecht.

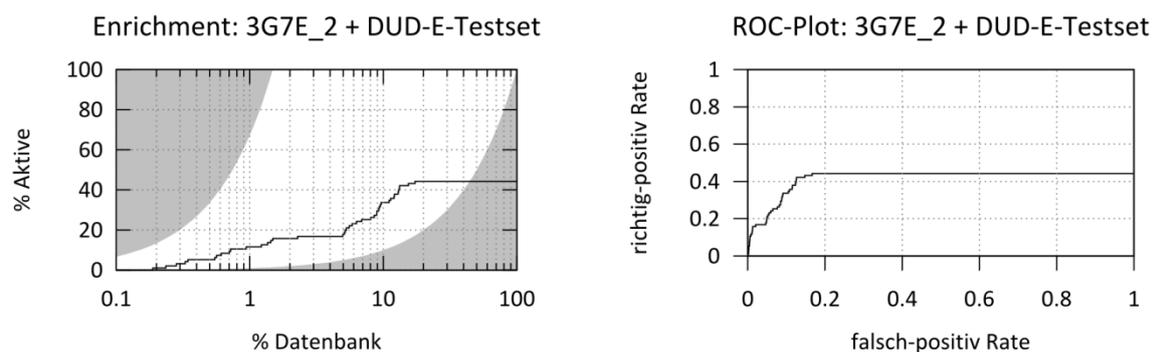


Abbildung 4.12 – Links: Enrichment des cRAISE-Dockings in 3G7E_2 mit dem DUD-E Docking-Testdatensatz. **Rechts:** ROC-Plot des gleichen Dockings. Die AUC beträgt 0.416.

Insgesamt konnte der DUD-E Docking-Testdatensatz damit in allen Enzymstrukturen mit akzeptablen Anreicherungen gedockt werden. Im Folgenden wurden die Dockings mit der gefilterten ZINC-Datenbank (s. Abschnitt 3.4 auf S. 73) durchgeführt. Während des Screenings wurden die aktiven Substanzen, die zur Erstellung des DUD-E-Datensets verwendet wurden, mitverwendet, um Anhaltspunkte über die Qualität des Screenings zu gewinnen.

– cRAISE-Docking der ZINC-Datenbank

Die gefilterte ZINC-Datenbank umfasste ca. 3.4 Millionen Substanzen. Insgesamt wurden 4.8 Millionen Posen erzeugt, die einen Score kleiner null besaßen (s. Tabelle 4.1). Jede Substanz konnte auch in mehr als in einer

Tabelle 4.1 – Ergebnisse des ersten cRAISE-Dockings.

	Anzahl
Anzahl an Posen	4.8 Millionen
davon 1E11	1.0 Millionen
davon 1AJ6	1.8 Millionen
davon 3G7E_1	1.5 Millionen
davon 3G7E_2	0.5 Millionen
Anzahl gedockter Substanzen	2.0 Millionen

Struktur gedockt werden, daher war die Anzahl der Posen höher als die Anzahl der Substanzen in der gefilterten ZINC-Datenbank. Die Verteilung gab schon Aufschluss über die Ergebnisse. 1AJ6 und 3G7E_1 lieferten die meisten Docking-Ergebnisse mit jeweils 1.5 Mio. Posen oder mehr. Bei 1E11 nahm die Anzahl der Posen schon sehr deutlich auf knapp 1 Mio. Posen ab, bei 3G7E_2 gab es nur noch 0.5 Mio. Posen. Insgesamt wurden 2.0 Mio. verschiedene Substanzen gedockt, statistisch betrachtet wurde damit jede erfolgreich gedockte Substanz in 2.4 verschiedene Proteine gedockt. Die starke Reduktion der Posen bei 3G7E_2 war nachvollziehbar, da das viel strengere Pharmakophor weniger Ergebnisse zuließ. Die geringere Anzahl an Dockingergebnissen bei 1E11 war hingegen nicht vorhersehbar. Die Anforderungen und Ähnlichkeit zu den beiden Enzymstrukturen 3G7E und 1AJ6 ließen auch ähnliche Ergebnisse erwarten. Zwar waren in 1E11 *Exclusion Volumes* definiert, die das Docking einschränken sollten, diese befanden sich jedoch in einem Bereich, der bei den anderen Bindetaschen nicht ausgebildet war (s. Abschnitt 2.1 auf S. 27). Die Tendenz, dass sich die Posen bei 1E11 anders als bei den anderen beiden ähnlichen Strukturen verhielten, konnte auch anhand der Score-Werte nachvollzogen werden. In Abbildung 4.13 ist der jeweilige Rank der Substanz gegen ihren cRAISE-Score aufgetragen. Dabei konnte sehr deutlich zwischen den Dockings von 1AJ6 und 3G7E_1 gegenüber 1E11 und 3G7E_2 unterschieden werden. Die Werte der Scores der erstgenannten Gruppe waren in ihrem gesamten Verlauf niedriger, deutlich länger auf einem niedrigen Niveau und stiegen nur langsam an. Der Verlauf der Kurven von 3G7E_1 und 1AJ6 war nahezu parallel. Die Differenz von 1AJ6 zu 3G7E_1 war während des Anfangs konstant bei ca. 1.6. Erst im späteren Verlauf nahm die Differenz langsam zu (bspw. betrug sie bei Rank 800,000 2.1). Bei 1E11 und 3G7E_2 waren die Scores im gesamten Verlauf deutlich höher. Dabei waren die Scores auf den niedrigen Rängen gleichauf, im weiteren Verlauf nahmen sie aber bei 3G7E_2 schneller zu.

Damit gab es eine generelle Tendenz: Je mehr Posen in einer Proteinstruktur erzeugt werden konnten, desto waren auch die Scores. Im DUD-E-Testdatensatz war die Anreicherung bei 1E11 am schlechtesten (s. Abbildung 4.9 S. 92). Bei dem Docking der Gesamtdatenbank waren die Scores von 1E11 zu Beginn sogar schlechter als bei 3G7E_2, bei dem es deutlich weniger Posen gab. Erst im weiteren Verlauf wurden die Scores von 1E11 besser als bei 3G7E_2.

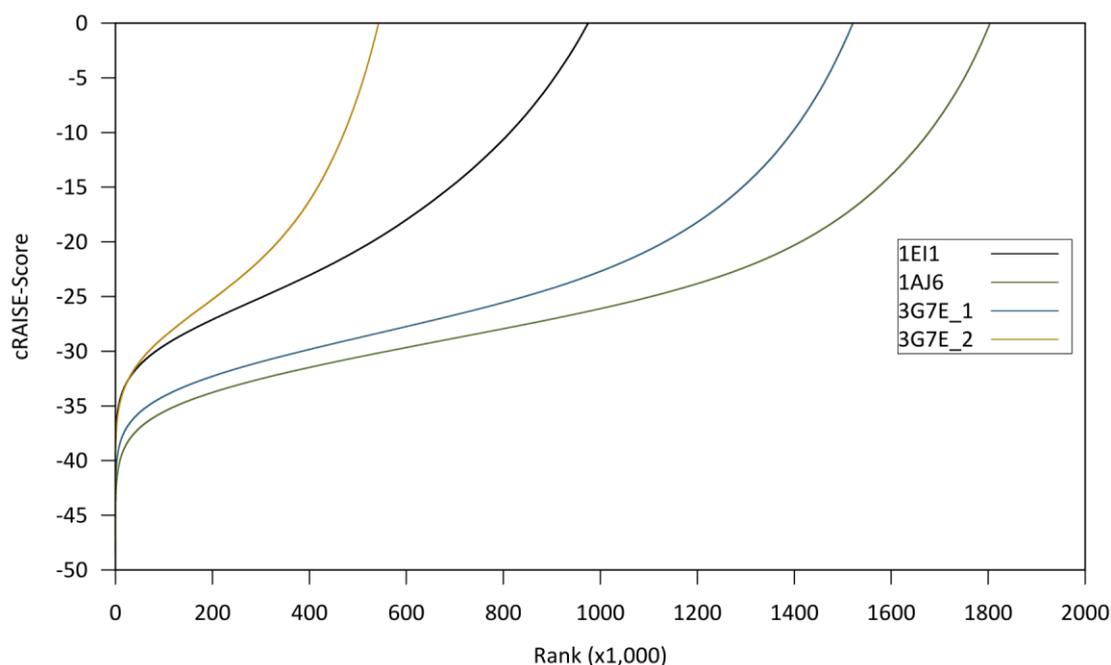


Abbildung 4.13 – Vergleich der verschiedenen cRAISE-Dockings. Für jede Pose ist der Rank gegen seinen jeweiligen Score aufgetragen.

Dem gefilterten ZINC-Datensatz wurden auch die 95 Aktiven, die zur Erstellung des DUD-E-Datensatzes verwendet worden sind, hinzugefügt. Aufgrund der großen Anzahl von Verbindungen in diesem Datensatz hatten die Aktiven nur einen Anteil von unter 0.03 %. Die Enrichments können daher nicht mit denen des DUD-E-Testdatensatz verglichen werden. Zusätzlich wurde für das Docking ein Datensatz verwendet, der auch potentiell aktive Substanzen enthält. Eine Substanz wird möglicherweise auch deswegen besser bewertet, weil sie aktiv ist.

In Abbildung 4.14 ist das Enrichment der Dockings dargestellt. Abbildung 4.15 zeigt die entsprechenden ROC-Plots. 1AJ6 und 3G7E_1 besaßen eine gute Anreicherung. Bei den ROC-Plots wurde die gleiche Tendenz deutlich. Bei 3G7E_2 wurde gerade zu Beginn eine gute Anreicherung erreicht, die bei 1 % der Datenbank sogar die beste war. Lediglich 1E11 zeigte im gesamten Bereich nur eine schwache Anreicherung.

Zusammengefasst konnten damit folgende Schlüsse gezogen werden: Konnten viele Substanzen in die jeweilige Zielstruktur gedockt werden, wurden im Vergleich besser bewertete Posen generiert und es konnten auch mehr Aktive in höheren Rängen wiedergefunden werden. 3G7E_2 nahm eine Sonderstellung ein, da die Struktur aufgrund des deutlich strengeren Pharmakophors weniger Posen generieren konnte und auch viele Aktive des Testdatensatzes diese nicht erfüllen konnten.

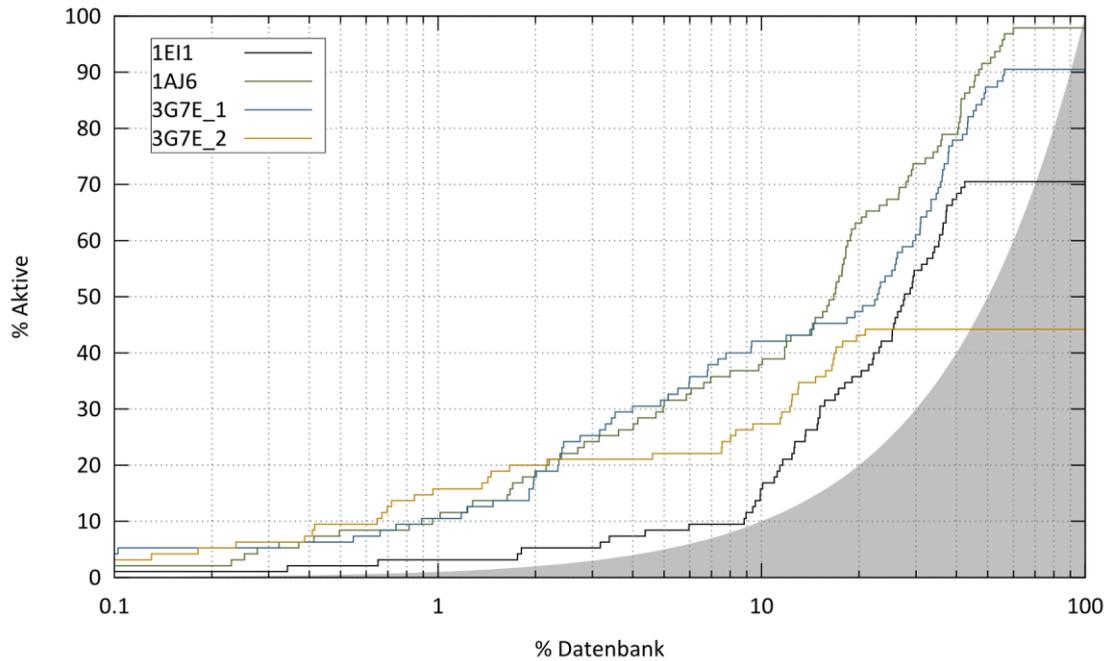


Abbildung 4.14 – Enrichments der Dockings mit dem Datensatz der gefilterten ZINC-Datenbank.

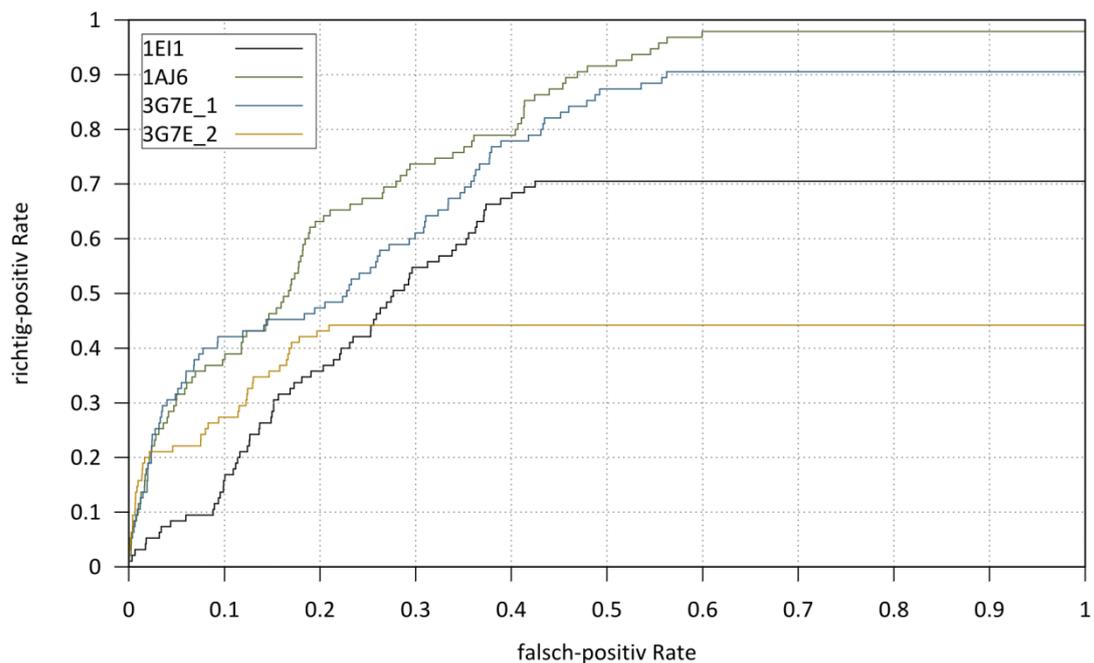


Abbildung 4.15 – ROC-Plot der Dockings mit dem Datensatz der gefilterten ZINC-Datenbank. Die AUC betrug für 1E11 0.563, für 1AJ6 0.793, für 3G7E_1 0.730 und für 3G7E_2 0.410.

Für die Durchführung des zweiten Dockings mussten die Docking-Ergebnisse noch stark gefiltert werden, da für den zweiten Docking-Prozess eine deutlich rechenaufwendigere Methode verwendet wurde. Dafür wurden die hier aufgeführten Filterkriterien verwendet, die im Folgenden diskutiert werden:

1. cRAISE-Score-Cutoff von -20
2. Top 15 % der verbleibenden Posen.

Der erste Schritt führte dazu, dass Posen bevorzugt wurden, die gute Scores und damit eine höhere Qualität besaßen. Das zweite Kriterium bevorteilte die Dockings, die viele Ergebnisse erzeugten und bessere Anreicherungen gezeigt hatten. Diese zwei Schritte favorisierten damit die Dockings mit vielen und guten Ergebnissen gegenüber den Dockings mit weniger Ergebnissen und schlechteren Scores. Diese Reduktion

besaß Vorteile gegenüber reinen Score-Cutoffs – dort wären die Dockings von 3G7E_1 und 1AJ6 stark überrepräsentiert gewesen – oder reinen prozentualen Filtern – bei denen anteilmäßig viele Posen von den schlechteren Docking-Ergebnissen weiter mit eingeflossen wären. Mit der gewählten Kombination findet eine gewichtete Reduktion aller Ergebnisse statt. Bspw. wurde die Zahl der Docking-ergebnisse in 3G7E_1 auf 11.1 % seiner ursprünglichen Posen reduziert, bei 1EI1 betrug der Anteil 8.1 %. Durch diese Filterkriterien wurde der Datensatz von den ursprünglich 3.4 Millionen Substanzen der gefilterten ZINC-Datenbank auf 360,000 Substanzen reduziert (entspricht 10.6 %). Der Vorteil durch den Ansatz, verschiedene Zielstrukturen zu benutzen, wird anhand der Anreicherung deutlich. Durch Kombination^{xi} der vier Dockings wurde eine deutlich bessere Anreicherung erreicht als bei jedem Docking einzeln. In Tabelle 4.2 ist der *Enrichment Factor* für den letztendlich gewählten Anteil von 10.6 % der Substanzen gezeigt. Diese Anreicherungen mussten, wie oben beschrieben, vorsichtig beurteilt werden. Sie zeigten jedoch, dass durch die verschiedenen Zielstrukturen unterschiedliche Aktive gefunden werden konnten, die in einer kombinierten Auswertung eine höhere Anreicherung brachten. In Abbildung 4.16 ist dieser Zusammenhang grafisch dargestellt.

Durch das hier eingesetzte Auswahlverfahren konnten bei den verschiedenen Dockings die Anteile der einzelnen Dockings weiterverwendet werden, die Vorteile zeigten. Bspw. wurde nur ein kleiner Anteil von 3G7E_2 verwendet, der aber in diesem Bereich eine deutliche Überlegenheit gegenüber den anderen Dockings aufwies.

Tabelle 4.2 – Anreicherung der Einzeldockings verglichen mit einer kombinierten Auswertung.

	EF_{10.6 %}
1EI1	1.6
1AJ6	3.7
3G7E_1	4.0
3G7E_2	2.6
kombinierte Auswertung	5.9

^{xi} Für die kombinierte Auswertung wurden alle vier Dockings entsprechend der gewählten Kriterien gefiltert und „Rank by Rank“ vereint. Dabei wurden alle Duplikate entfernt und nur der jeweils bessere Rank behalten.

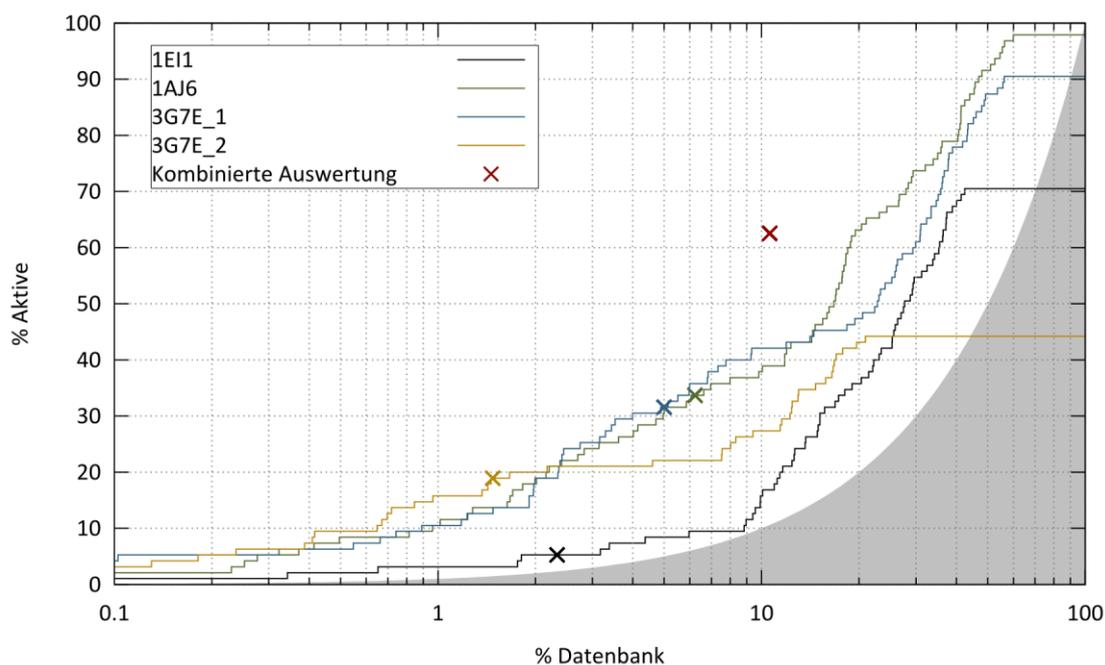


Abbildung 4.16 – Enrichment-Plot des ersten cRAISE-Screenings. Für jedes Docking sind die Enrichments bei einem Scorecutoff von -20 und der TOP 15 % markiert (Kreuze). Das rote Kreuz markiert das durch die kombinierte Auswertung insgesamt erreichte Enrichment. Dafür wurden die verschiedenen Dockings „Rank by Rank“ vereint. Kam eine Substanz mehrmals vor, wurde der bessere Rank gewählt.

4.2.3 Zweites cRAISE-Docking

Im zweiten Docking wurden die ca. 360,000 Substanzen des ersten Dockings (darin sind noch 59 Aktiven Substanzen enthalten) erneut mit cRAISE gedockt, um die Qualität der Posen zu verbessern (s. Abschnitt 4.1.1 auf S. 80). Jede dieser gefilterten Substanzen wurde in diesem Schritt erneut in alle verwendeten Proteinstrukturen gedockt, unabhängig davon, in welcher Proteinstruktur die Substanz im ersten Docking entdeckt wurde. Die deutlich kleinere Menge an Substanzen erlaubte es bei der erneuten Generierung des Substanz-Indexes, mehr Konformationen pro Substanz zu erstellen als im ersten Docking. Dadurch steigt zwar der Rechenaufwand für die Erstellung des Indexes und des eigentlichen cRAISE-Dockings stark an, die Qualität der Posen nimmt aber erheblich zu: Durch die größere Anzahl an Konformeren nimmt bspw. die Anzahl der Clashes der vdW-Radien erheblich ab. Dies wurde schon aus den Scores ersichtlich (s. Abbildung 4.17), die deutlich geringer, also besser, waren als im ersten cRAISE-Docking. Gab es beispielsweise im ersten Docking bei 1AJ6 207 Substanzen mit einem guten Score von unter -45, waren es im zweiten Docking über 2,700. Die Gruppierungen des ersten Dockings waren auch hier wieder ersichtlich: 3G7E_2 und 1E11 hatten leicht höhere Scores als 1AJ6 und 3G7E_1. Die Unterschiede waren jedoch kleiner und 1E11 näherte sich der anderen Gruppe ein wenig an. Die Ergebnisse von 3G7E_1 und 1AJ6 waren nahezu identisch in ihrem Verlauf. In 1AJ6 konnten lediglich einige Substanzen mehr gedockt werden. Auffällig war auch, dass deutlich mehr

Posen in 1E11 generiert werden konnten. Daher war der Ansatz, alle Substanzen erneut in alle Zielstrukturen zu docken, sinnvoll.

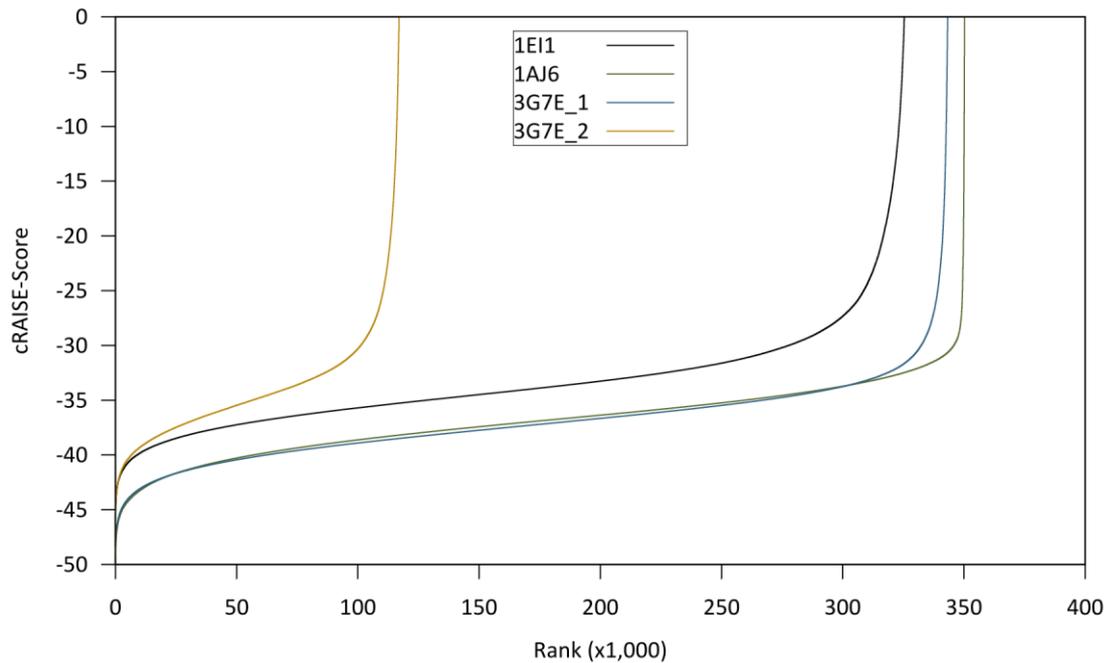


Abbildung 4.17 – Vergleich der verschiedenen cRAISE-Dockings. Für jede Pose ist der Rank gegen seinen jeweiligen Score aufgetragen.

Der Vorteil der intensiveren Dockingsuche wird auch in dem Enrichment deutlich (s. Abbildung 4.18). 1E11 befand sich nach diesem Docking auf einem ähnlichen Niveau wie 1AJ6 und 3G7E_1. Bei den beiden letztgenannten Strukturen zeigten sich keine Verbesserungen der Anreicherungen mehr, die AUC des ROC-Plots nahm sogar leicht ab (s. Abbildung 4.19). Dennoch war die Durchführung eines erweiterten Dockings für jede Zielstruktur sinnvoll, um Posen mit höherer Qualität zu generieren, die später im Postprocessing weiterverarbeitet werden konnten (s. Abschnitt 5 auf S. 100).

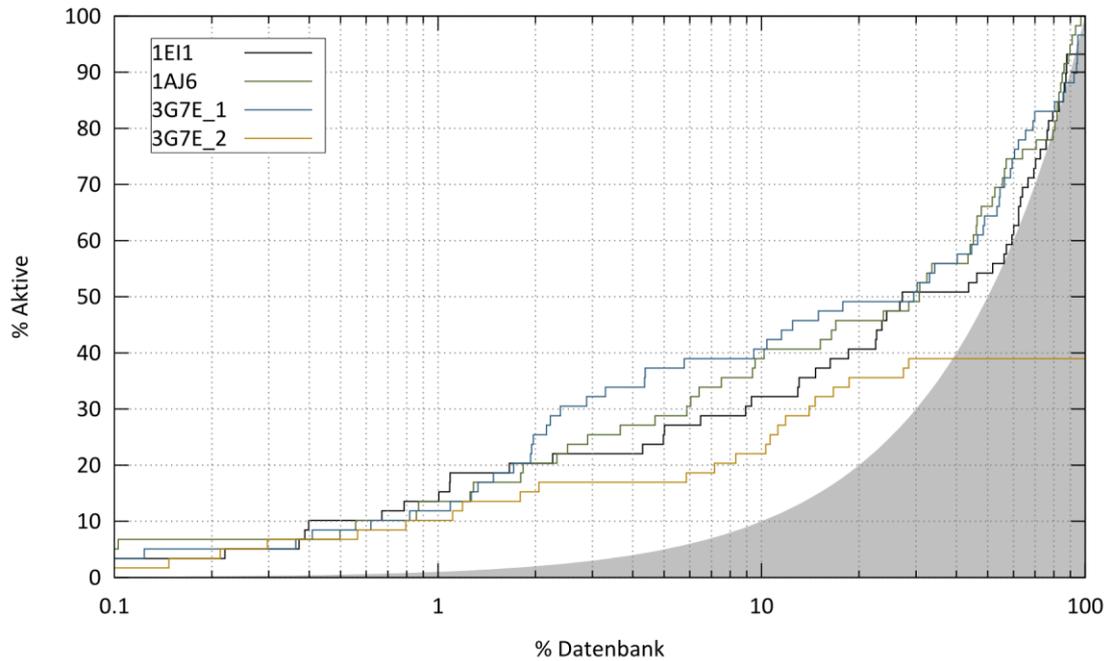


Abbildung 4.18 – Enrichments des zweiten cRAISE-Dockings.

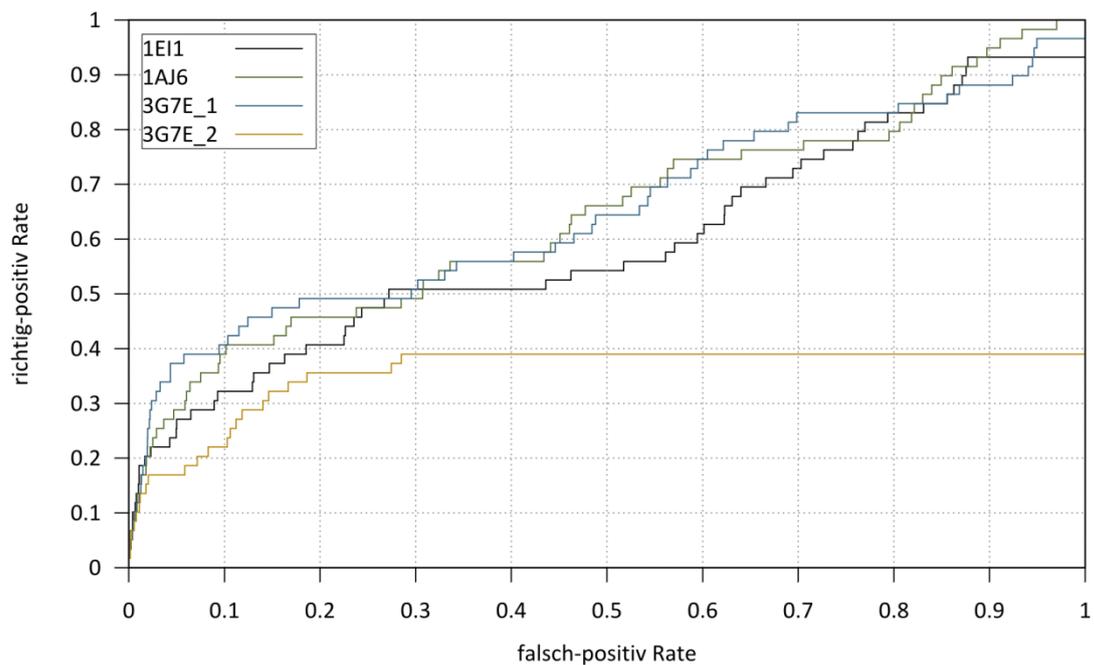


Abbildung 4.19 – ROC-Plots des zweiten cRAISE-Dockings. Die AUC betrug für 1EI1 0.593, für 1AJ6 0.642, für 3G7E_1 0.651 und für 3G7E_2 0.357.

Damit konnten in diesem HTVS 3.4 Mio. Substanzen in vier verschiedenen Proteinstrukturen gescreent werden. Insgesamt wurde die Substanzbibliothek auf 360,000 verschiedene Substanzen reduziert. Dies stellte schon eine deutlich kleinere Anzahl an Substanzen dar, war jedoch noch zu groß, um sie einzeln zu untersuchen. Die nächste Aufgabe bestand in der schrittweisen Reduktion und Anreicherung der Substanzen in der Nachprozessierung.

5 Postprocessing

Das Postprocessing stellt einen weiteren elementaren Schritt bei einer strukturbasiertem Leitstruktursuche dar. Es dient dazu, die durch das erste Docking erhaltenen Strukturen neu zu bewerten, zu sortieren sowie die Anzahl der Substanzen zu reduzieren (s. Abschnitt 1.2 auf S. 9). Dieser Prozess ist notwendig, da die beim Docking verwendeten Scoring-Funktionen für eine genaue Differenzierung unzureichend sind. Dies ist bei dem hier gewählten umso wichtiger, da das Docking mit einem HTVS-Tools durchgeführt wurde, bei dem der Fokus auf eine schnelle Platzierung und Bewertung zugunsten der Qualität der Pose ausgelegt ist (s. Abschnitt 4.1.1 auf S. 80).

Die einzelnen Postprocessing-Schritte können sehr heterogen sein: Von weiteren Docking-Schritten [195, 196], über einzelne Scoring-Funktionen [197], Consensus-Scorings [198, 199], verschiedenen Clusteringmethoden [200], pharmakophorbasierten Ansätze [201], MD-Simulationen mit GB/PBSA-Berechnungen [202] bis zur visueller Auswahl [203] und der jeweiligen Kombination der einzelnen Schritte [204] gibt es viele Möglichkeiten. Im Gegensatz zu der Vorbereitung der Substanzbibliothek (s. Abschnitt 3 auf S. 61) können hier im weiteren Verlauf des Workflows, aufgrund der durch das Docking schon erfolgten Reduktion der Anzahl an Substanzen, rechenintensivere Schritte durchgeführt werden. Im Folgenden werden die hier relevanten Methoden kurz vorgestellt.

– Scoring-Funktion

Scoring-Funktionen sollen das Scoring-Problem des Dockings lösen (s. Abschnitt 4 auf S. 79): Sie sollen richtig platzierte Posen von falsch platzierten unterscheiden und Posen entsprechend ihrer Aktivität sortieren. Da es sich beim Postprocessing vornehmlich um die Bewertung von bereits vorhandenen Posen handelt, bezieht sich die Diskussion hier primär auf das Rescoring bereits gedockter Posen und nicht auf den Einfluss des Scoring auf den eigentlichen Dockingprozess.

Klassischerweise wird zwischen drei Ansätzen unterschieden: kraftfeldbasiert, empirisch und wissensbasiert [205–207].

1. Bei den kraftfeldbasierten Scoring-Funktionen wird die Bindungsenergie des Protein-Ligand-Komplexes über ein Kraftfeld (s. auch Abschnitt 2.2.1.3 auf S. 41) berechnet, das auf atomare Interaktionen zurückzuführen ist. Dabei gehen in der Regel enthalpische Effekte wie van-der-Waals-, elektrostatische Interaktionen und Bindungskräfte (Dehnung, Torsion usw.) in die Berechnung mit ein. Problematisch sind insbesondere entropische Effekte, die in diesen Modellen nur durch sehr rechenintensive Arbeitsschritte oder ungefähre Approximationen berücksichtigt werden können.

2. Bei empirischen Funktionen wird die Bindungsenergie in einfachere, skalierbare Fragmente (z. B. Wasserstoffbrückendonoren, hydrophobe Interaktionen) zerlegt, die zu einem Gesamtscore aufaddiert werden. Jedes Fragment enthält einen Skalierungsfaktor, der durch eine Regressionsanalyse mit empirischen Daten bestimmt wird. Der Vorteil besteht darin, dass auch komplexe Interaktionen effizient berechnet und bewertet werden können (z. B. entropische Effekte). Dafür ist die Qualität der Regression von den vorhandenen empirischen Daten abhängig: zum einen von der Richtigkeit, zum anderen von den Eigenschaften des Datensatzes. Interaktionsmuster, die nur unzureichend von dem zur Erstellung der Funktion verwendeten Datensatz beschrieben werden, können von dieser Funktion nicht richtig bewertet werden.
3. Die wissensbasierten Funktionen werden mithilfe von statistischen Analysen von Kristallstrukturen erstellt. Diese versuchen die Bindungsenergie mittels Potentialen, die auf der Boltzmann-Verteilung basieren, zu bestimmen. Mit der inversen Boltzmann-Verteilung (s. Formel 5.1) lässt sich die Bindungsenergie ΔE über die Boltzmannkonstante k_B , die Temperatur T und die paarweisen Potentiale, die anhand der Dichtefunktion vom Zustand $\rho(i)$ und eines Referenzzustandes ρ_{ref} bestimmt werden, berechnen. Bei dieser Gleichung wird ebenso wie bei den empirischen Funktionen auf experimentelle Daten zurückgegriffen. Allerdings werden die Kristallstrukturen direkt ausgewertet und in einer Häufigkeitsverteilung verarbeitet (z. B. wie häufig tritt ein Alkohol in einer bestimmten Entfernung zu einem Carbonyl auf). Dazu gibt es mehr und verlässlichere Daten als bei den freien Bindungsenthalpien, die bei den empirischen Scoring-Funktionen verwendet werden. Wissensbasierte Scoring-Funktionen sind ähnlich schnell wie empirische. Dennoch bereiten hier unterrepräsentierte Interaktionen (z. B. Metallbindungen) und die Definition des Referenzzustandes ρ_{ref} Probleme.

$$\Delta E = -k_B T \ln \left(\frac{\rho(i)}{\rho_{\text{ref}}} \right) \quad \text{Formel 5.1}$$

Diese Einteilung stellt nur eine grobe Kategorisierung dar. Manche Scoring-Funktionen bestehen auch aus der Kombination verschiedener Methoden (z. B. kombiniert die AutoDock Scoring-Funktion den empirischen und kraftfeldbasierten Ansatz [208]).

Die Evaluierung von Scoring-Funktionen ist ein in der Literatur intensiv untersuchtes Thema [205, 206, 209–211]. Zusammengefasst gibt es keine Scoring-Funktion, die allen anderen überlegen ist. Grundsätzlich erlauben Scoring-Funktionen eine gute Priorisierung von Bindungsposes. Lediglich bei Entropie-, Solvatationseffekte und selten auftretenden Interaktionen (z. B. Metallkoordination) sind die Ergebnisse unbefriedigend. Ebenso sind Korrelationen zu wahren Bindungsaffinitäten noch nicht verlässlich möglich. Die Performance von Scoring-Funktionen ist außerdem noch sehr häufig vom

Target abhängig. STAHL *et al.* konnten 2001 zeigen, dass der PLP-Score (s. 5.1.1 auf S. 106) für die ATPase-Untereinheit der Gyrase eine deutlich bessere Performance zeigt als drei andere Scores [209], während der PLP Score bei anderen Targets diesen Scoring Funktionen unterlegen war.

– Fingerprints und Clustering

Die automatische Charakterisierung und Analyse struktureller Informationen von Molekülen ist für Computer schwierig: Idealerweise sollten die gesamten strukturellen zwei- oder dreidimensionalen Eigenschaften eines Moleküls eindeutig und vergleichbar in eine einfache Zahlenfolge übersetzt werden. Nur diese kann dann von einem PC analysiert und verarbeitet werden. Das Problem besteht also darin, mehrdimensionale Eigenschaften in einem eindimensionalen Raum darzustellen, ohne dabei Informationen zu verlieren.

Fingerprints stellen genau diese Übersetzungen dar. Dabei werden strukturelle Eigenschaften von Molekülen in binäre Zahlenfolgen umgewandelt. Durch die Reduktion auf den eindimensionalen Raum gehen jedoch immer Informationen verloren. Daher gibt es verschiedene Ansätze, auf welchen Eigenschaften die Fingerprints basieren. Die zuerst entwickelten Fingerprints bestehen aus einfach definierten Strukturelementen, denen jeweils eine Stelle des Binärcodes zugeordnet ist. Hat die Struktur ein bestimmtes Strukturelement (z. B. Ladung, 6-Ring), wird an der entsprechenden Stelle des Binärcodes eine 1, ansonsten eine 0, gesetzt. Die Güte dieses Fingerprints ist sehr stark von den vorher definierten Strukturelementen abhängig und kann aufgrund ihrer begrenzten Anzahl und der mangelnden Informationen über die Konnektivität große Datensätze nur schlecht differenzieren. Zudem kann er nur auf vorher festgelegten Strukturmerkmalen basieren (bspw. MACCS von MDL). Weiterführende Fingerprints ordnen bestimmte Atome, Atomgruppen (ECFP [212]), Verknüpfungen (Daylight Fingerprint [213]) oder auch Pharmakophoren (FCFP [212]) eindeutigen Werten zu, die anschließend „*hashed*“ werden (s. auch 5.1.6 auf S. 109). Dadurch sind diese Fingerprints zwar abstrakter zu interpretieren, da die einzelnen Bits nicht einer bestimmten Struktur oder Merkmal zuzuordnen sind, dafür können sie aber für jede Art von Molekülen verwendet werden und ihre Länge kann entsprechend der Anforderung angepasst werden.

Nach dem Erstellen der Fingerprints können diese entsprechend ihrer Übereinstimmung verglichen werden. Der bei weitem am häufigsten genutzte Parameter dafür ist der Tanimoto-Koeffizient, auch wenn es eine Vielzahl von anderen Parameter gibt (z. B. Cosine, Pearson, Simpson) [214]. Dabei berechnet sich der Tanimoto-Koeffizient T_K aus den Bits von zwei Fingerprints (a und b) und der Summe der übereinstimmenden Bits c (s. Formel 5.2):

$$T_K = \frac{c}{a + b - c}$$

Formel 5.2

Sind beide Fingerprints identisch, beträgt der Koeffizient definitionsgemäß 1. Gibt es keine Übereinstimmung in den beiden Fingerprints nimmt der Tanimoto-Koeffizient einen Wert von 0 an. Da der Koeffizient maximal 1 beträgt, kann der Wert auch umgekehrt werden (s. Formel 5.3). Dann beschreibt der Koeffizient t nicht die Ähnlichkeit, sondern wie stark sich die Substanzen unterscheiden (0 kein Unterschied; 1 Fingerprints komplett verschieden).

$$t = 1 - T_K \quad \text{Formel 5.3}$$

Im Allgemeinen werden Substanzen mit einem Tanimoto-Koeffizienten, der größer als 0.7 (entspricht $t \leq 0.3$) ist, als sehr ähnlich angesehen [215].

5.1 Material und Methoden

5.1.1 Piecewise Linear Potential Score

Bei dem „*Piecewise Linear Potential*“ (PLP)-Score handelt es sich um einen empirischen Score, der schon 1995 von GEHLHAAR *et al.* beschrieben wurde [217]. Er berechnet das Interaktionspotential zwischen den schweren Atomen eines Liganden und den sich innerhalb von 6 Å befindlichen schweren Atomen des Proteins und summiert diese auf. Dabei handelt es sich bei dem Algorithmus, der die Energie der einzelnen Interaktion beschreibt, um eine stückweise lineare Funktion, deren Verlauf Ähnlichkeit mit dem Lennard-Jones-Potential hat. Allerdings wird die Abstoßungsenergie nicht unendlich groß, wenn der Abstand zwischen den schweren Atomen null beträgt. Das führt dazu, dass der Score gegenüber kleinen Überlappungen der vdW-Radien toleranter ist und die Pose nicht verwirft, wenn andere starke Interaktionen überwiegen [216].

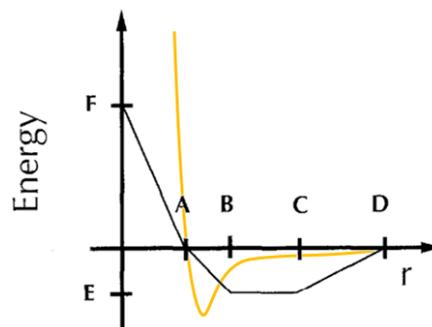


Abbildung 5.1 – Verlauf der PLP-Funktion (schwarz) und des Lennard-Jones-Potential (gelb). Jeder Abschnitt der PLP-Funktion ist linear. Der wichtigste Unterschied zum Lennard-Jones-Potential stellt der endliche Wert F dar: Bei Überschneidung der vdW-Radien strebt das Lennard-Jones-Potential gegen unendlich, während sich die PLP-Funktion einem definierten Wert nähert. (mod. nach [216])

Zur Bestimmung des PLP-Scores wurde die Implementierung von LeadIt (BioSolveIT [218]) verwendet.

5.1.2 Glide

Glide (*Grid-based Ligand Docking with Energetics*) [163, 219] ist ein systematisch erschöpfendes Docking-Tool (s. Abschnitt 4 auf S. 79). Das Docking verläuft in vier Schritten: (1) Zunächst werden die Form und Interaktionsmöglichkeiten von dem Rezeptor mit verschiedenen Feldern auf einem *Grid*

(engl. Gitter) dargestellt. (2) Daraufhin werden energetisch günstige Konformationen des Liganden generiert und innerhalb des *Grids* platziert. (3) Erfolgsversprechende Posen werden im dritten Schritt mithilfe des OPLS-AA Kraftfeld [108] in Verbindung mit einem distanzabhängigen Dielektrizitätsmodell minimiert. (4) Abschließend werden mithilfe von Monte Carlo-Simulationen Minima für die Torsionswinkel des Liganden gesucht. Für die finale Bewertung und das Ranking der Posen wird der GlideScore (s. Abschnitt 5.1.3) verwendet.

Die Dockings wurden mit Glide 6.2 (Schrödinger [220]) durchgeführt.

5.1.3 GlideScore

Der GlideScore ist eine empirische Scoring-Funktion, die bei der Bewertung von Dockingergebnissen von Glide (s. Abschnitt 5.1.2) verwendet wird. Dabei können zwei unterschiedliche Methoden, *Single Precision (SP)* oder *Extra Precision (XP)*, mit unterschiedlicher Parametrisierungen verwendet werden. Der GlideScore SP hat deutlich weichere Grenzen, so dass auch Posen, die kleinere Fehler enthalten, nicht verworfen werden [163]. Der GlideScore wurde aus dem ChemScore [221] entwickelt und erweitert. Der in dieser Arbeit verwendete GlideScore 2.5 SP wird wie folgt berechnet (s. Formel 5.4):

$$\begin{aligned}
 \Delta G_{\text{bind}} = & C_{\text{lipo-lipo}} \sum f(r_r) + C_{\text{hbond-neut-neut}} \sum g(\Delta r) h(\Delta \alpha) \\
 & + C_{\text{hbond-neut-charged}} \sum g(\Delta r) h(\Delta \alpha) \\
 & + C_{\text{hbond-charged-charged}} \sum g(\Delta r) h(\Delta \alpha) \\
 & + C_{\text{max-metal-ion}} \sum f(r_{\text{im}}) + C_{\text{rotb}} H_{\text{rotb}} \\
 & + C_{\text{polar-phob}} V_{\text{polar-phob}} + C_{\text{coul}} E_{\text{coul}} + C_{\text{vdW}} E_{\text{vdW}} \\
 & + \text{solvation term}
 \end{aligned}
 \tag{Formel 5.4}$$

Die *C*-Koeffizienten sind empirisch ermittelte Werte für die einzelnen Terme. *f*, *g* und *h* sind Funktionen, die über Abstände *r* oder Winkel *α* die einzelnen Koeffizienten gewichten. Die Funktionen nehmen Werte zwischen 0 und 1 an. *H_{rotb}* beschreibt den Anteil eingeschränkt drehbarer Bindungen und der gesamte Term ist damit ein Maß für die Entropie. Der *C_{polar-phob}V_{polar-phob}*-Term beschreibt den Einfluss polarer, nicht-wasserstoffbrücken-bildender Atome in hydrophoben Bereichen. Als weitere Parameter gehen noch die coulombschen (*E_{coul}*) und van-der-Waals-Energien (*E_{vdW}*) ein, die mithilfe einer modifizierten Version des OPLS-Kraftfelds bestimmt werden. Der *solvation term* bewertet den Einfluss von expliziten Wassermolekülen auf die Pose. Dies schließt sowohl Solvatisierungseffekte von sehr hydrophilen Bereichen des Liganden oder des Proteins als auch in hydrophoben Bereichen eingeschlossene Wassermoleküle mit ein. Der GlideScore wurde mit Glide (s. 5.1.2) bestimmt.

5.1.4 DSX Score

Der DSX-Score ist ein wissensbasierter Score. Es handelt sich dabei um eine Weiterentwicklung des DrugScore [222] und zeigt insbesondere für die Bewertung von Dockingposen eine gute Performance [223]. Der Score berechnet sich aus dem distanzabhängigen Paarpotential ($\text{score}_{\text{pair}}$), Torsionswinkelpotential ($\text{score}_{\text{tors}}$) und SAS-abhängigen Potential (score_{SR}) (s. Formel 5.5):

$$\begin{aligned} \text{score}_{\text{pair}} &= \sum_{a_i \in P} \sum_{a_j \in L} \text{score}_{\text{pair}}^{\text{DSX}}(c(a_i, a_j), r(a_i, a_j)) \\ \text{score}_{\text{tors}} &= \sum_b \sum_{T \in b} \frac{\text{score}_{\text{tors}}^{\text{DSX}}(t(T), \phi(t))}{n_T} \\ \text{score}_{\text{SR}} &= \sum_{a \in P} \text{score}_{\text{SR}}^{\text{DSX}}(c(a), \text{SR}(a)) + \sum_{a \in L} \text{score}_{\text{SR}}^{\text{DSX}}(c(a), \text{SR}(a)) \end{aligned} \quad \text{Formel 5.5}$$

Dabei ist a ein Protein- P oder ein Ligand L -Atom. c ist ein Clustertyp mit r als Abstand für die gegebenen Atome. b stellt eine zentrale Bindung eines Torsionswinkel T vom Typ t mit der Anzahl n_T an möglichen Torsionswinkeln dar. SR beschreibt schließlich das Verhältnis der SAS von gebundenen und ungebundenen Atomen.

Diese einzelnen Terme können schließlich mit einem Gewichtungsfaktor w , der zwischen 0 und 1 liegt zum Endscore ($\text{score}_{\text{total}}$) aufaddiert werden (s. Formel 5.6).

$$\text{score}_{\text{total}} = w_p \text{score}_{\text{pair}} + w_t \text{score}_{\text{tors}} + w_s \text{score}_{\text{SR}} \quad \text{Formel 5.6}$$

Der Gewichtungsfaktor kann dabei für jedes Scoring variabel angepasst werden. Abhängig vom Target und den Liganden kann der Score daher weiter optimiert werden. In dieser Arbeit wurde der Score in seiner Standardeinstellung verwendet.

5.1.5 Prime MM-GBSA-Score

Der Prime MM-GBSA (Molecular Mechanics - Generalized Born Surface Area)-Score ist ein kraftfeldbasierte Scoring-Funktion, die die Freie Bindungsenthalpie mithilfe eines impliziten Lösungsmodells berechnet. Sie basiert auf dem in Formel 5.7 angegebenen Zusammenhang:

$$\Delta G_{\text{Bindung}} = E_{\text{Komplex}} - E_{\text{Protein}} - E_{\text{Ligand}} \quad \text{Formel 5.7}$$

Demnach wird die Freie Bindungsenthalpie $\Delta G_{\text{Bindung}}$ über die verschiedenen Energien E des Komplex, des isolierten Proteins und des isolierten Liganden berechnet. Diese enthalten verschiedene Terme, wie Protein-Ligand-vdW-Kontakte, elektrostatische Interaktionen, Desolvatisierungsenergie vom

Liganden und Deformationsenergien von Liganden und Proteinen. Der entropische Einfluss auf die freie Bindungsenthalpie wird dabei ignoriert. Zur Berechnung wird das OPLS2005-Kraftfeld [103] mit dem impliziten Lösungsmodell VSGB 2.0 [224] verwendet.

Für die Bestimmung der einzelnen Energieterme wird jeder der Strukturen (Komplex, Protein und Ligand) zunächst minimiert. Für die Minimierung kann gewählt werden, ob das Protein komplett starr gehalten wird oder innerhalb eines bestimmten Radius um den Liganden frei minimiert werden kann. Je größer der Radius um den Liganden ist, desto rechenaufwendiger werden die Minimierungen.

5.1.6 Daylight-Fingerprint

Der Daylight-Fingerprint [213] berechnet für jedes Atom ein *Pattern*^{xii}, ein *Pattern* für jedes Atom mit seinem unmittelbaren Nachbarn (mit den entsprechenden Bindungen) und sequentiell weitere *Patterns* für jedes Atom mit steigender Anzahl an Bindungen (bis sieben). Diese *Patterns* werden schließlich *hashed*, d.h. die Information wird reduziert, so dass für jedes *Pattern* vier bis fünf Bits erzeugt werden, die in die Erstellung des Fingerprints eingehen [225, 226]. Daher wird der Daylight-Fingerprint auch häufig als ein pfadbeschreibender Fingerprint bezeichnet, da bei jedem Schritt der *Pattern* durch die Verlängerung des Pfades um eine Bindung erstellt wird und der *Pattern* somit die Pfade im Molekül beschreibt [212].

5.1.7 Clustering

Das Clustering erfolgte nach dem Leader-Algorithmus [227, 228]. Dabei wird zum ersten Fingerprint einer Liste der Tanimoto-Koeffizient (T_K) zu allen anderen Strukturen berechnet. Liegt der Koeffizient eines Fingerprints unterhalb eines festgelegten Grenzwertes, wird die Substanz dem Leader zugeordnet und aus der Liste entfernt. Im nächsten Schritt wird die nächste Substanz aus der Liste verwendet, die nicht dem ersten Fingerprint zugeordnet wurde und erneut nach dem gleichen Schema verfahren. Dies wird solange wiederholt, bis kein Ligand mehr in der Liste vorhanden ist [227].

Der Vorteil dieser Methode liegt darin, dass die Reihenfolge der verwendeten Liste bestimmt, welche Struktur zum Cluster-Leader wird. Bei einer bereits priorisierten Liste an Substanzen, z. B. durch ein Docking oder Rescoring, wird dadurch die am besten bewertete Struktur des jeweiligen Clusters zum Leader. Weiterhin werden alle Strukturen aufgrund ihrer Ähnlichkeit zum Leader bestimmt, die erfolgreich und gut gedockte Posen darstellen.

^{xii}Ein Pattern (engl. für Muster) stellt die Übersetzung der Struktur zu einer Zahl dar. Dies können z. B. Integers oder binäre Zahlen sein (daher auch der Name; die Abfolge von Einsen und Nullen bei binären Zahlen stellen ein Muster dar).

5.2 Planung

Die Optimierung des Postprocessing erfolgte mit einer Rationalen. Aufgrund der vielfältigen Möglichkeiten für die Durchführung, lag der Fokus beim Postprocessing auf einer frühen und hohen Anreicherung von aktiven Substanzen, die aber auch Freiheiten für spätere Hit-Erweiterungen lassen sollte. Das Postprocessing sollte folgende Schritte enthalten:

1. Eine Verbesserung der Anreicherung aktiver Verbindungen aus den cRAISE-Dockings durch ein geeignetes *Rescoring*. Bei den Ergebnissen des cRAISE-Dockings konnten zwar die meisten Aktiven gedockt werden, doch insbesondere die frühe Anreicherung war unzureichend (s. Abschnitt 4.2.3 auf S. 98).
 - Diese Optimierung wurde mithilfe des gedockten DUD-E Docking-Testdatensatz durchgeführt, da dieser Datensatz für die Testung der Differenzierung von aktiven und inaktiven Substanzen optimiert ist (s. Abschnitt 4.1.3 auf S. 85).
2. Ein Leader-Clustering der Neubewerteten Posen. Mit dem Clustering sollte die Anzahl der weiter untersuchten Strukturen deutlich reduziert werden, ohne die bisher gewonnen Informationen zu verlieren: Bei jedem weiteren Schritt des Clusterleaders kann nachträglich auch der Cluster untersucht werden (bspw. potentielle Kandidaten für die Hit-Erweiterung).
 - Die Optimierung für das Clustering konnte an dem gesamten Screening-Datensatz durchgeführt werden.
3. Redocking mit Glide. Der vorletzte Schritt sollte ein Redocking der Clusterleader beinhalten, das verlässlichere Posen als das HTVS-Tool cRAISE generiert.
 - Für die Evaluierung des Dockings wurde ebenfalls der DUD-E Docking-Testdatensatz (s. o.) verwendet.
4. Visuelle Auswahl. Schließlich erfolgte die finale Auswahl der zu kaufenden Substanzen anhand verschiedener Parameter, wie Güte der Pose oder auch chemische Variabilität.

5.3 Rescoring

Die Ausgangsdaten für das Postprocessing stellen die im letzten Kapitel durchgeführten Dockingergebnisse dar (s. Abschnitt 4.2.3 auf S. 98). Wiederum wurden die Parameter zunächst mithilfe des DUD-E Docking-Testdatensatz (s. Abschnitt 4.1.3 auf S. 85) optimiert. Dafür wurde der DUD-E Docking-Testdatensatz mit cRAISE unter den gleichen Bedingungen wie der ZINC-Datensatz gedockt (s. Abschnitt 4.1.3 auf S. 85). Abbildung 5.2 zeigt die Ausgangssituation für diese Optimierung: Die aktiven Substanzen konnten in 3G7E_1 fast alle erfolgreich gedockt werden (93.7 %). Dennoch ist die Anreicherung gerade am Anfang, nur unzureichend. Damit war ein gutes, zum Docking-Prozess der ZINC-Datenbank vergleichbares Ausgangsszenario vorhanden, um eine Optimierung durchzuführen und diese zu analysieren.

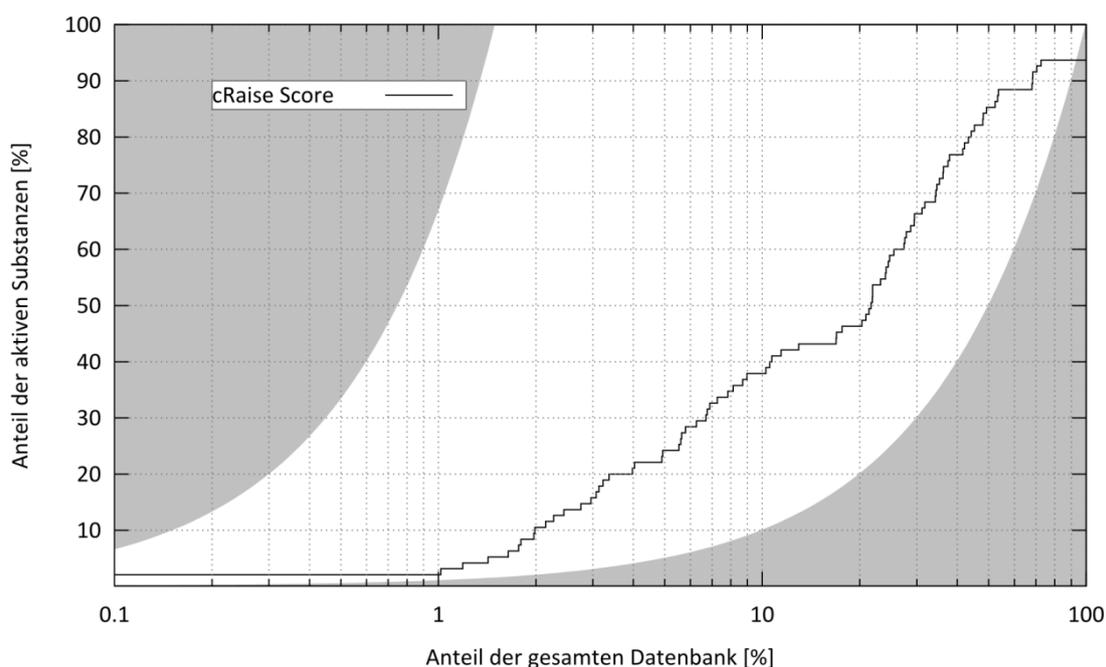


Abbildung 5.2 – Enrichment des cRAISE-Scores nach Docking des DUD-E Docking-Testdatensatz in 3G7E_1. Auch wenn fast alle aktiven Substanzen erfolgreich gedockt wurden, war die Anreicherung zu Beginn nur unzureichend. Erst ab ca. 1 % der gesamten Datenbank findet eine effektive Anreicherung der Aktiven statt.

– GlideScore

Ein komplettes Rescoring mit dem GlideScore war nicht möglich. Nur 136 Substanzen der 4,738 erfolgreich gedockten Substanzen konnten überhaupt bewertet werden. Die meisten Posen wurden aufgrund von Überschneidung des vdW-Radius von Ligand und Protein nicht bewertet. Dabei besaßen nur 59 Posen einen Score, der kleiner war als -5 kcal/mol. Wie in Abbildung 5.3 zu sehen, wurden 29.5 % der Aktiven erfolgreich gescort. Dennoch ist das beginnende Enrichment sehr gut. Unter den ersten 20 Posen sind nur drei Decoys.

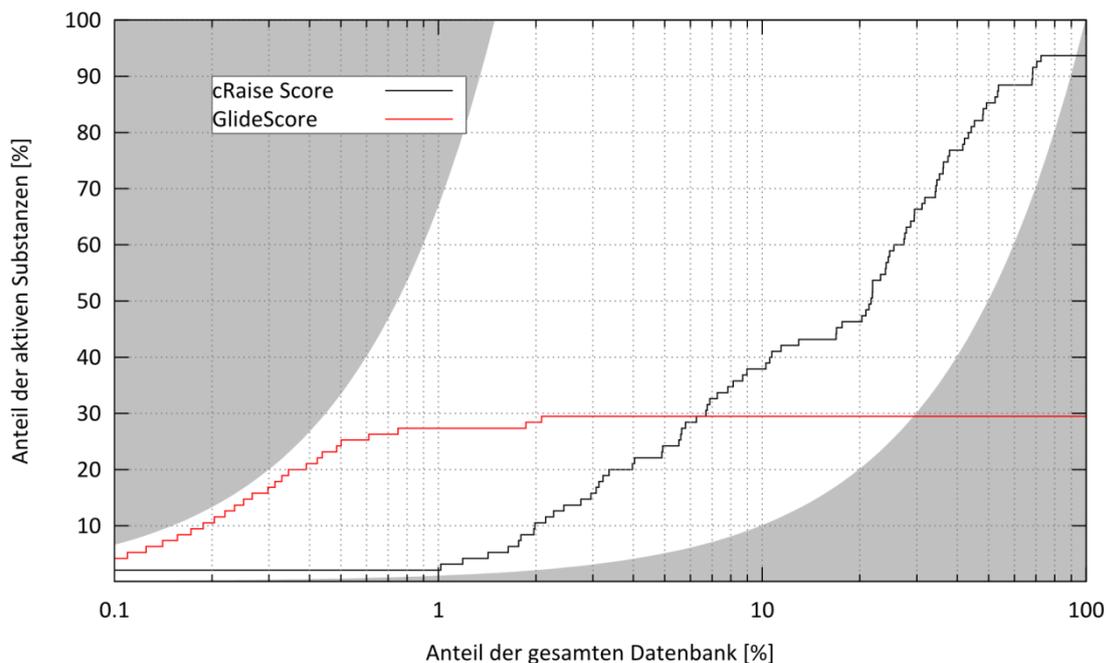


Abbildung 5.3 – Enrichment des GlideScores. Der GlideScore (rot) zeigte ein deutlich besseres frühes Enrichment als der cRAISE-Score (schwarz). Dennoch konnte nur ein kleiner Anteil der gedockten Posen überhaupt bewertet werden.

– DSX-Score

Ein DSX-Rescoring zeigte eine leichte Verbesserung des cRAISE-Scores sowohl im Anfangs- als auch im mittleren Bereich (s. Abbildung 5.4). Dieser Vorteil verringerte sich jedoch ab einem Anteil von 6 % der gesamten Datenbank maßgeblich und war ab 10 % schwach. Insgesamt wurden aber fast alle Aktiven erneut gescort.

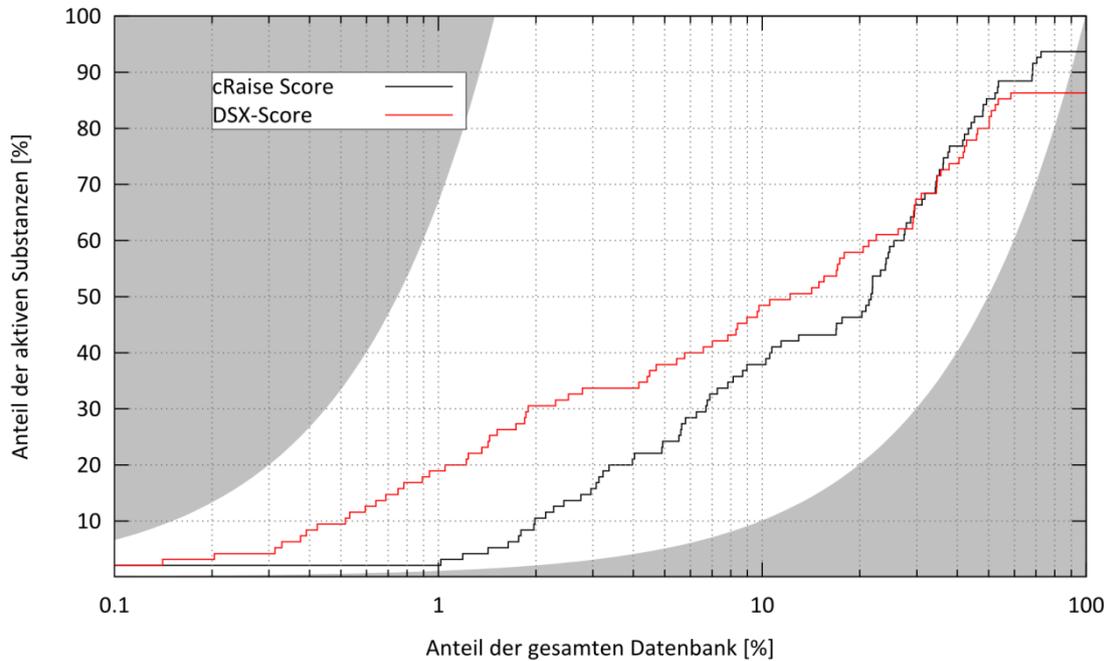


Abbildung 5.4 – Enrichment des DSX-Scores. Der DSX-Score (rot) zeigt eine moderate Verbesserung des frühen Enrichments im Vergleich zum cRAISE-Score (schwarz). Dieser Vorteil geht aber im weiteren Verlauf verloren.

– GBSA-Score

Als kraftfeldbasierte Scoring-Methode wurde der Prime MM-GBSA-Score verwendet, GBSA freezed (ohne Minimierung des Proteins) und GBSA 8 Å (mit Minimierung des Proteins; s. Experimenteller Teil 11.3.11 auf S. 226). Die statischere Methode GBSA freezed ohne Minimierung des Proteins zeigte eine deutlich bessere Anreicherung als die flexiblere mit minimierten Seitenketten (s. Abbildung 5.5). Das Enrichment für GBSA 8 Å ist nicht besser als der cRAISE-Score. Demgegenüber zeigt der GBSA freezed-Methode gerade zu Beginn eine deutlich bessere Anreicherung.

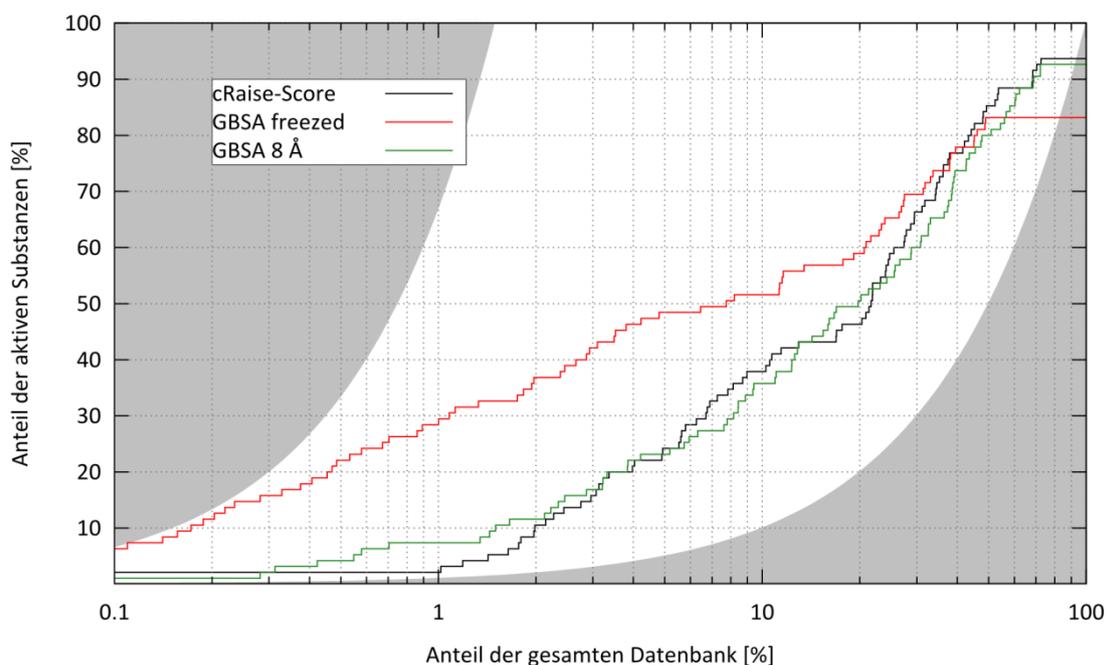


Abbildung 5.5 – Enrichment der kraftfeldbasierten Scoring-Funktionen. GBSA freezed (rot) erreichte von Anfang an eine sehr gute Anreicherung und bewertete auch den Großteil der aktiven Substanzen. GBSA-8 Å zeigt keine Verbesserung gegenüber den cRAISE-Score.

– PLP-Score

Der PLP-Score zeigte zum Beginn des Enrichments (s. Abbildung 5.6) eine perfekte Anreicherung. Die ersten 18 Substanzen sind ausschließlich Aktive. Dabei wurden unterschiedliche chemische Strukturen aus verschiedenen Literaturquellen erfolgreich gerankt.

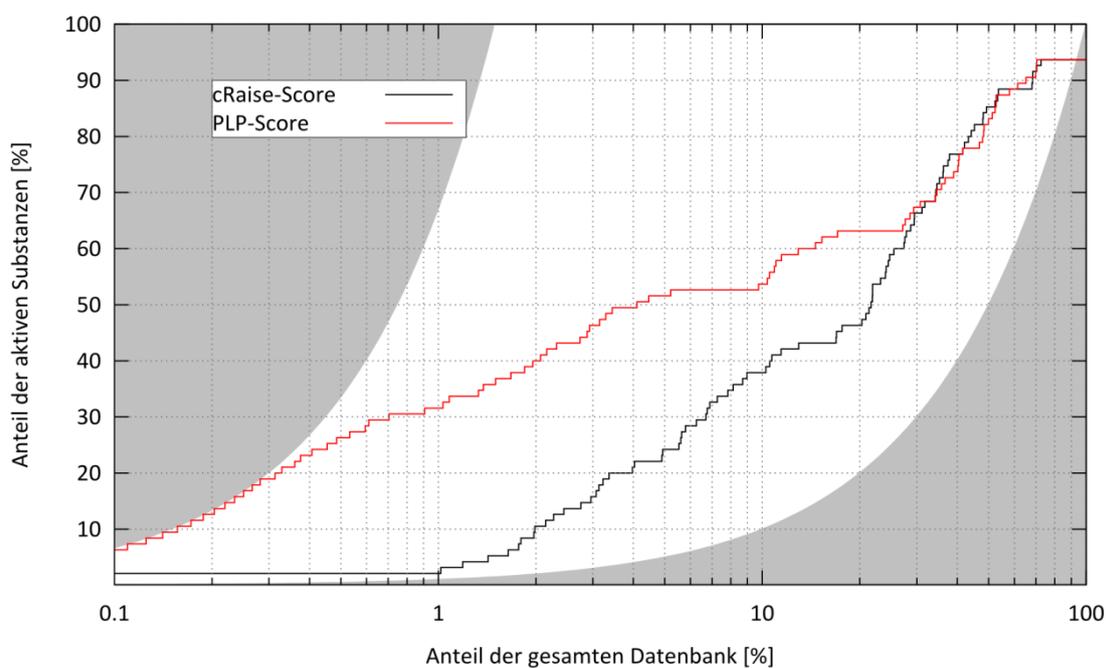


Abbildung 5.6 – Enrichment des PLP-Scores. Der PLP-Score (rot) verbesserte das Enrichment erheblich. Zu Beginn ist die Anreicherung perfekt und die Anreicherung ist im gesamten Verlauf auf einem hohen Niveau. Es konnten alle Aktiven erneut bewertet werden.

In Abbildung 5.7 sind alle Enrichments zusammengefasst. Die GBSA 8 Å-Methode und der DSX-Score lieferten nur eine schwache bis keine Verbesserung der Anreicherung für die Posen. Der GlideScore, GBSA freezed und der PLP-Score zeigten alle eine sehr gute frühe Anreicherung, waren aber im weiteren Verlauf und ihrer Aussagekraft unterschiedlich zu bewerten.

Die GBSA-8 Å-Methode war in allen Bereichen unterlegen. Weder war die Anreicherung signifikant besser als im cRAISE-Score noch war der Zeitaufwand für die Berechnung in einem annähernd akzeptablen Bereich (über 3 Tage für 4,738 Posen). Die Gründe für das schlechte Abschneiden könnten in der Güte der Posen liegen: Die Geometrien der Ausgangsposten von cRAISE sind für die GBSA-8 Å-Methode unzureichend.

Der Vorteil des DSX-Scores lag in seiner hervorragenden Effizienz: Für die Bewertung der 4,738 Posen benötigte die Scoring-Funktion nur vier Sekunden. Zwar war die Anreicherung besser als im cRAISE-Score, sie war aber noch deutlich schlechter als in den anderen Scoring-Funktionen (s. u.).

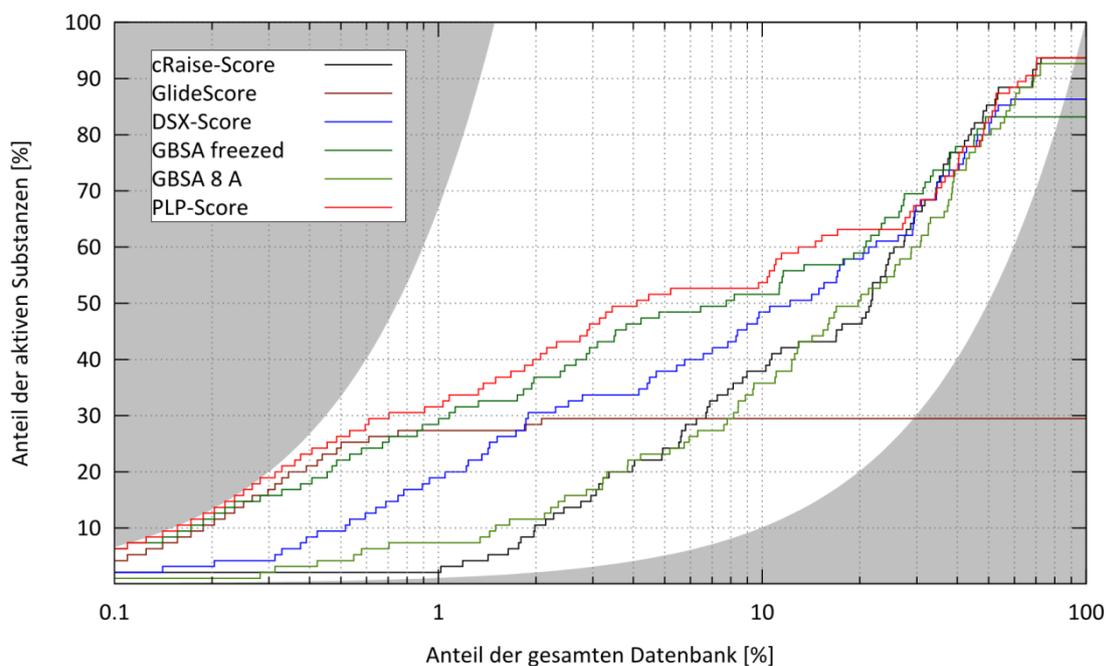


Abbildung 5.7 – Übersicht der Enrichments aller verwendeten Scoring-Funktionen. Der GBSA-8 Å-Score und der DSX-Score waren nur wenig bis gar nicht besser als der cRAISE-Score und brachten keine großen Vorteile für ein erneutes Ranking. Der GlideScore, GBSA freezed und der PLP-Score zeigten alle eine gute frühe Anreicherung.

Der GlideScore besaß ein gutes frühes Enrichment, erreichte aber kein hohes Gesamtenrichment, weil die meisten Posen, auch von aktiven Substanzen, nicht bewertet worden. Die Ursache dafür lag ebenfalls in den von cRAISE generierten Posen. Es traten die gleichen Probleme, wie beim GBSA 8 Å-Rescoring auf. Der GlideScore ist für ein Docking optimiert, das solche Clashes verhindert. Daher wurden die meisten Posen nicht bewertet und die bewerteten Posen besaßen vergleichsweise schlechte Score-Werte. Nichtsdestotrotz konnten die aktiven Substanzen sehr gut angereichert werden. 26 der 28 erfolgreich gescorten Posen von Aktiven waren in den TOP50. Damit war der Gli-

deScore dafür geeignet, aktive Substanzen zu finden, sofern die Posen für die Scoring-Funktion bewertbar waren. Für den Datensatz von 4,738 Posen wurden 7 min benötigt.

Die GBSA freezed-Methode zeigte eine ähnlich gute Anreicherung und konnte im Gegensatz zum GlideScore fast alle Posen bewerten. Dafür waren die Berechnungen für diese Funktion deutlich aufwendiger: Für die 4,738 Posen wurde eine Rechenzeit von 8:50 h benötigt. Im Vergleich zu der GBSA 8 Å-Methode waren die Score-Werte jedoch deutlich schlechter (freie Bindungsenthalpie von unter -100 kcal/mol bei GBSA 8 Å: 197, bei GBSA freezed: 8). Folglich konnte die freie Komplexminimierung der GBSA 8 Å-Methode die Posen in absoluten Werten erheblich verbessern, führte aber zu einem schlechteren relativen Ranking. Möglicherweise wiesen die Decoys größere Clashes als die Aktiven auf. Die GBSA freezed-Methode konnte die Decoys mit einem großen Clash von den Liganden mit einem kleineren Clash differenzieren. Bei der ausführlicheren Minimierung der GBSA 8 Å-Methode konnten die Clashes bei beiden reduziert und aneinander angeglichen werden und führten zu einheitlicheren Score-Werten.

Über den PLP-Score gelang durchgehend die beste Anreicherung. Zusätzlich konnte er die meisten Posen des cRAISE-Dockings erneut gut bewerten. Dafür gab es verschiedene Gründe: Schon bei der Entwicklung des Scores wurde darauf geachtet, ihn toleranter gegenüber Überlappungen der vdW-Radien zu machen (s. Abschnitt 5.1.1 auf S. 106). Dies geschah unter der Annahme, dass sich das Protein noch an den Liganden anpassen kann [216]. Damit verfolgt die Scoring-Funktion eine ähnliche Rationale wie cRAISE (s. Abschnitt 4.1.1 auf S. 80). Daher konnte der Score auch die hohe Anzahl an Posen wieder bewerten, da genau diese Clashes für die anderen Scoring-Funktionen (z. B. GlideScore) häufig problematisch waren. Dabei ermittelte der PLP-Score auch sehr niedrige Score-Werte und bewertete die Posen damit gut. Dies war auch ein Vorteil gegenüber den GlideScore oder den GBSA-freezed-Score, bei welchen grundsätzlich eine gute Anreicherung stattfand, die Score-Werte aber in einem schlechten Bereich waren. Aufgrund dieser Scores wurden die Posen also als nicht sehr vertrauenswürdig bewertet. Zusätzlich konnten schon STAHL *et al.* zeigen, dass der PLP-Score gute Ergebnisse für dieses Target lieferte [209].

Damit war der PLP-Score die optimale Wahl für die Neubewertung der Posen. Die Scoring-Funktion war konzeptionell sowohl für den Docking-Algorithmus als auch für das Target die beste Wahl. Die Zeit für die Bewertung der Posen war mit 4:50 min auch ausreichend effizient, um sehr große Datensätze verarbeiten zu können.

5.4 Leader-Clustering

Nach dem erfolgreichen Rescoring und Ranking durch den PLP-Score wurde der gedockte ZINC-Datensatz im nächsten Schritt für jede Proteinstruktur mithilfe von Daylight-Fingerprints geclustert. Dafür wurde ein Clustering mittels des Leader-Algorithmus mit verschiedenen Cutoffs durchgeführt. Die Cutoffs beschreiben den Grenzwert, ab welchem Unterschied t ($1 - \text{Tanimotokoeffizient}$) der gerade verarbeitete Fingerprint nicht mehr dem Leader zugeordnet wird.

Die Problemstellung bestand darin, ein ausgewogenes Verhältnis zwischen der Anzahl der Cluster und deren Größe zu finden: Einerseits sollten möglichst wenige Einzelcluster, also ungeclusterte Substanzen, und andererseits nicht zu große Cluster, die zu heterogen wären, auftreten. Da die Erstellung des Fingerprints und das Clustering nicht anhand eines Pharmakophors stattfinden, besteht ein Problem darin, dass auch als ähnlich eingestufte Strukturen verschiedene Bindungsmodi besitzen könnten. Daher war eine strenge Auslegung von Ähnlichkeit vorteilhaft. In der Literatur werden Strukturen mit einem Cutoff von 0.3 häufig als ähnlich klassifiziert (s. Abschnitt 5.1.7 auf S. 109). Daher wurde im Folgenden der Cutoff-Bereich von 0.20 bis 0.35 untersucht, um die Einflüsse zu analysieren. In Tabelle 5.1 sind charakteristische Merkmale für die unterschiedlichen Cluster aufgelistet.

Erwartungsgemäß nahm die Anzahl der Cluster mit zunehmenden Cutoff ab. Es gab jedoch schon bei einem sehr niedrigen Cutoff von 0.20 bei allen Zielstrukturen sehr große Cluster (> 100 Substanzen). Diese nahmen im weiteren Verlauf noch drastisch zu. Bei einer Erweiterung des Cutoffs auf 0.25 entstanden bei den drei größeren Dockings (1E11, 3G7E_1 und 1AJ6) bereits Cluster, die mehr als 500 Substanzen beinhalteten. Dies bedeutete im Umkehrschluss, dass eine Substanz über 500 Substanzen repräsentieren würde und deren Bindungseigenschaften beschreiben sollte. Beim Clustering von 3G7E_1 gab es bei einem Cutoff von 0.35 einen Cluster mit 1,765 Substanzen. In diesem einen Cluster wären einer Substanz über 0.5 % aller Dockingergebnisse zugeordnet.

5.4 Postprocessing

Tabelle 5.1 – Eigenschaften der Leader-Cluster aus dem PLP-Score entstandenem Ranking der gedockten ZINC-Datenbank. Für jedes Protein (1AJ6, 1E11, 3G7E_1, 3G7E_2) sind die charakteristischen Clustermerkmale für unterschiedliche Cutoffs ($t = 0.20$ bis 0.35) dargestellt. #-Posen: Anzahl der Posen. „Größter Cluster“: Anzahl der Substanzen aus dem der größte Cluster bestand. „Verteilung der Clustergrößen“: Gibt die Anzahl der Cluster mit entsprechenden Clustergrößen wieder.

1AJ6 (#-Posen: 349,263)	0.20	0.25	0.30	0.35
Anzahl der Cluster	137,664	100,672	70,664	46,740
Größter Cluster	424	688	826	948
Verteilung der Clustergrößen				
≥ 500	0	2	3	12
499 – 100	36	97	242	409
99 – 50	168	373	539	702
49 – 10	4,701	6,023	6,473	6,256
9 – 2	53,524	43,750	32,801	22,013
1	79,235	50,427	30,606	17,348

1E11 (#-Posen: 316,526)	0.20	0.25	0.30	0.35
Anzahl der Cluster	124,448	91,228	64,581	42,986
Größter Cluster	297	531	763	957
Verteilung der Clustergrößen				
≥ 500	0	1	3	9
499 – 100	33	94	218	356
99 – 50	166	355	514	656
49 – 10	4,283	5,335	5,674	5,439
9 – 2	47,209	38,836	29,401	20,085
1	72,757	46,607	28,771	16,441

3G7E_1 (#-Posen: 338,412)	0.20	0.25	0.30	0.35
Anzahl der Cluster	132,182	96,891	68,156	45,309
Größter Cluster	360	706	1,165	1,765
Verteilung der Clustergrößen				
≥ 500	0	2	3	12
499 – 100	32	102	242	374
99 – 50	191	371	559	752
49 – 10	4,692	5,778	6,136	5,853
9 – 2	50,151	41,204	30,861	20,981
1	77,116	49,434	30,355	17,337

3G7E_2 (#-Posen: 111,397)	0.20	0.25	0.30	0.35
Anzahl der Cluster	50,930	39,367	29,219	20,750
Größter Cluster	181	288	447	531
Verteilung der Clustergrößen				
≥ 500	0	0	0	1
499 – 100	6	18	48	89
99 – 50	34	83	130	191
49 – 10	1,262	1,704	2,006	2,128
9 – 2	17,214	15,005	12,202	9,167
1	32,414	22,557	14,833	9,174

Neben der Anzahl und Größe der Cluster ist auch die Verteilung der Cluster wichtig. Da es sich um das Leader-Clustering von Dockingergebnissen handelte, waren die einzelnen Strukturen schon vorselektiert und sortiert. Das Ranking des PLP-Rescoring gab die Reihenfolge der Cluster vor: Der erste Cluster repräsentierte den bestbewerteten Liganden vom jeweiligen Docking, der zweite Cluster den nächstbesten bewerteten Liganden usw. Dies führte dazu, dass insbesondere im Anfangsbereich größere Cluster auftraten: Substanzen, die spezifische Strukturmerkmale zeigten, um im Enzym erfolgreich gedockt zu werden, wurden stark angereichert. Diese waren mit höherer Wahrscheinlichkeit auch ähnlicher zueinander. Schlecht bindende Strukturmerkmale können in einer bestimmten Konformation möglicherweise ebenfalls eine nachvollziehbare Dockingpose ergeben. Die Wahrscheinlichkeit, dass so ein Strukturmerkmal angereichert wird, ist dann jedoch deutlich kleiner. Zusätzlich würde die Pose vermutlich schlechter bewertet werden. Dies hieß aber im Umkehrschluss nicht, dass Einzelcluster grundsätzlich schlechtere Substanzen waren. Es konnte sich auch um einzigartige Strukturvariationen handeln, die nicht häufig vorkamen.

Dies wird am Beispiel vom Clustering der 3G7E_1-Ergebnisse durch Tabelle 5.2 verdeutlicht, bei der die nach dem PLP-Score 1,000 bestbewerteten Leader enthalten sind. Wieder sind bei verschiedenen Cutoffs charakteristische Kenngrößen für die Cluster angegeben. Zusätzlich ist die Anzahl der Substanzen angegeben, die aus der Summe der Clusterleader und den geclusterten Substanzen besteht, und die Anzahl an den Aktiven, die dem Datensatz zum Beginn des Screenings hinzugefügt worden waren. Im Vergleich zu den Eigenschaften aller geclusterten Substanzen (s. Tabelle 5.1) hatte sich bei einem Cutoff von 0.20 der Anteil der Einzelcluster deutlich gesenkt. Er machte nun nur noch knapp 20 % aus, statt eines Anteils von über 50 %. Durch die Anreicherung der großen Cluster bei guten Score-Werten, waren die riesigen Cluster (> 500 Substanzen) wieder bei größeren Cutoff-Werten vertreten.

Tabelle 5.2 – Verteilung des TOP1000 Clusters von 3G7E_1. Anzahl an Substanzen: Anzahl an Clusterleadern und repräsentierte Substanzen. Weitere Parameter wie bei Tabelle 5.1.

3G7E_1	0.20	0.25	0.30	0.35
Anzahl der Cluster	1,000	1,000	1,000	1,000
Anzahl an Substanzen	13,311	22,711	37,414	60,765
Anzahl an Aktiven	24	24	22	25
Größter Cluster	360	706	1,165	1,765
Verteilung der Clustergrößen				
≥ 500	0	2	3	10
499 – 100	12	40	100	154
99 – 50	45	90	99	165
49 – 10	259	338	410	418
9 – 2	496	425	340	227
1	188	105	48	26

Im gleichen Maße wie die Anzahl der Einzelcluster abnahm, nahm die Anzahl der zu großen Cluster zu. Die Anzahl der Substanzen, die insgesamt repräsentiert waren, stieg zwar stark an, dies hatte jedoch kaum Auswirkungen auf den Anteil der Aktiven.

Zusammenfassend wies die Wahl eines Cutoffs von $t = 0.2$ deutliche Vorteile auf. Die Cluster besaßen eine höhere Übereinstimmung mit dem möglichen Pharmakophor. Gleichzeitig wurde durch das Leader-Clustering von priorisierten Docking-Ergebnissen eine homogenere Datenbank untersucht, so dass ein niedriger Cutoff für das Ziel diverser Substanzklassen ausreichte. In der Literatur werden Substanzen mit einem kleineren Unterschied als 0.2 die gleiche biologische Aktivität zugeordnet [229]. Im Folgenden wurden die Cluster mit einem Cutoff von $t = 0.2$ weiterverwendet.

5.5 Redocking mit Glide

In dem finalen Docking sollten nur die Cluster-Leader neu gedockt werden, um Posen mit höherer Qualität zu erzeugen. Die verschiedenen Rescoring-Ergebnisse (s. Abschnitt 0 auf S. 111) verdeutlichten die unzureichende Qualität der Posen für eine abschließende Bewertung. Der GlideScore konnte bspw. ein Großteil der Posen aufgrund von Clashes nicht bewerten.

– DUD-E Docking-Testdatensatz

Zur Identifizierung des Redocking-Effektes wurden die 1,000 bestbewerteten Substanzen des PLP-Rescorings des DUD-E-Datensatz aus dem 1. Schritt (s. Abschnitt 5.3 auf S. 111) verwendet. Damit konnte gewährleistet werden, dass das Testscenario auch mit durch cRAISE erfolgreich gedockten und vom PLP-Score bewerteten Substanzen durchgeführt wurde.

Als Target wurde wieder 3G7E_1 genutzt. Zunächst mussten die Substanzen für das Docking vorbereitet werden. Dafür wurden die Strukturen mithilfe von LigPrep von Schrödinger [230] aufgearbeitet. LigPrep bereitet die Substanzen für das Docking in Glide vor, für welches es Tautomere, Protonierungen und Ringkonformationen bestimmt. cRAISE passt während des Dockings die Protonierung und Tautomere (im Folgenden Zustände genannt) an (s. Abschnitt 4.1.1 auf S. 80). Daher gab es von einigen Strukturen Zustände, die nach den Bewertungskriterien von Ligprep energetisch ungünstig gewesen und zugunsten günstigerer Zustände verworfen worden wären. Damit wären diese interessanten und schon im Docking bewährten Zustände verloren. Daher wurden zwar neue Zustände erzeugt, die von cRAISE erstellt wurden aber beibehalten. Weitere Zustände wurden mit Epik bei einem pH-Wert von sieben berechnet. Für jeden Ring wurden bis zu fünf Konformere erstellt. Im Testdatensatz befanden sich im Gegensatz zum späteren Screening-Datensatz noch chirale Substanzen. Die chiralen Zentren wurden nicht mehr verändert. Die abschließend erhaltenen Substanzen wurden automatisch mit dem OPLS_2005-Kraftfeld [103] minimiert (s. Experimenteller Teil 11.3.14 auf S. 227).

Für das SP-Docking mit Glide wurde ein Grid für 3G7E_1 erstellt, das als Pharmakophor das Asp73 als Wasserstoffbrückenakzeptor und Wat A als Wasserstoffbrückendonor enthielt, analog zu dem Vorgehen bei cRAISE (s. Abschnitt 4.2.1 auf S. 87). Während des Docking-Vorgangs mussten alle Restriants erfüllt werden. Es wurde das SP-Docking unter Standardbedingungen ohne Epic-Penalties durchgeführt [220]. Die fünf besten Posen wurden minimiert und die beste Pose gespeichert (s. Experimenteller Teil 11.3.14 auf S. 227).

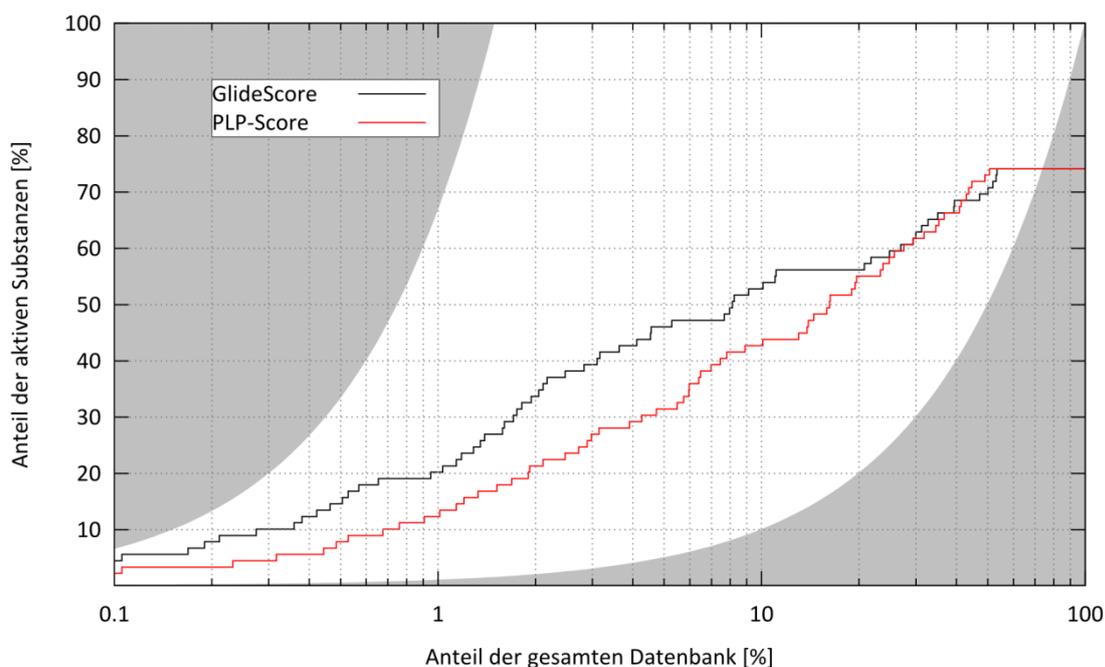


Abbildung 5.8 – Enrichment des Glide-Redocking des Testdatensatzes. Das Enrichment des GlideScores (schwarz) war auf einem hohen Niveau, auch wenn es schlechter war als bei der Anreicherung im ersten Schritt. Ein Rescoring der Glide-Posen mit dem PLP-Score brachte keine Verbesserung.

Das Enrichment nach Glide Redocking mit GlideScore ist in Abbildung 5.8 dargestellt: Es war auf einem hohen Niveau, wenn auch schlechter als beim PLP-Rescoring der cRAISE-Dockingergebnisse (s. Abschnitt 0 auf S. 114). Ein PLP-Rescoring (rot dargestellt) verbesserte die Anreicherung nicht. Dafür besaßen die Posen aufgrund des anspruchsvolleren Dockings eine deutlich höhere Qualität als nach dem Docking mit cRAISE (z. B. geringere Überlappung der vdW-Radien). Die Verschlechterung der Anreicherung war in diesem Fall akzeptabel: Beim ersten Schritt des Postprocessings der ZINC-Datenbank wurden 100,000de Substanzen neu bewertet und gerankt. Dort war die optimale Anreicherung ausschlaggebend, da im weiteren Verlauf nur noch unter 1 % der Substanzen weiter verarbeitet wurden. Zwar war die Anreicherung im letzten Docking-Schritt nicht zu vernachlässigen, der Fokus lag aber auf der Generierung optimaler Posen für die abschließende visuelle Bewertung der Posen. Dies ist beispielhaft mit einer Substanz des Decoy-Datensatzes dargestellt (s. Abbildung 5.9). In der cRAISE-Pose (links dargestellt) gibt es eine deutliche Überlappung der vdW-Radien (rote Kreise). Teilweise lagen sogar die Zentren der schweren Atome innerhalb des vdW-Radius des Proteins. Da-

her war die Substanz mit der Pose vom cRAISE-Docking nicht mit dem Glide-Score bewertbar. Nach dem Glide-Docking wurde eine deutlich bessere Pose gefunden, die bewertet werden konnte.

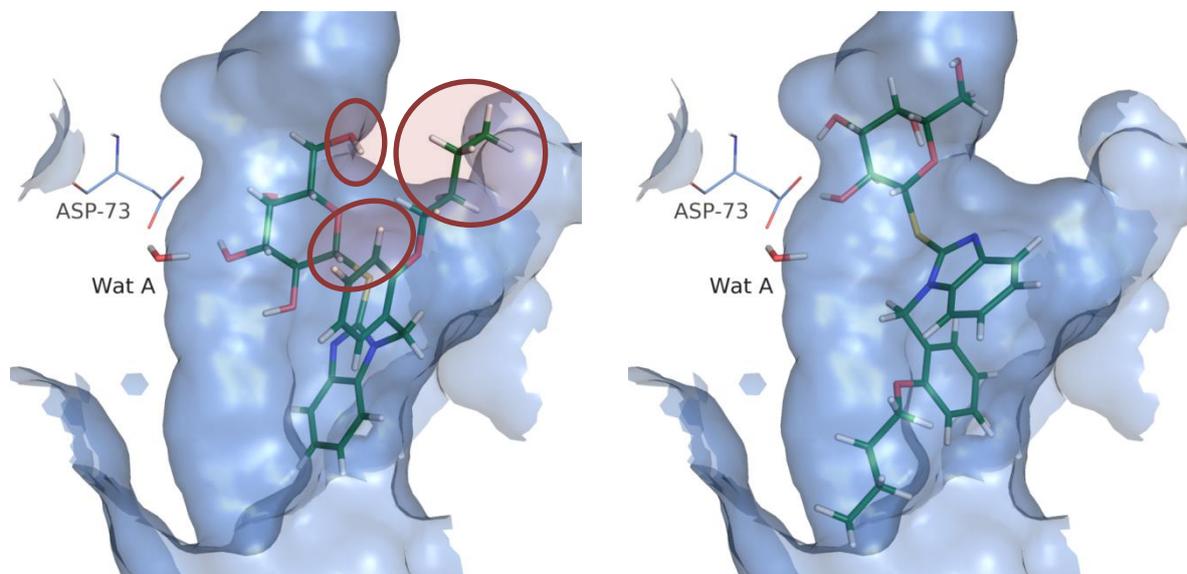


Abbildung 5.9 – cRAISE-Pose (links) und Glide-Pose (rechts) von einem Decoy. Bei der cRAISE-Pose ist die Überlappung der vdW-Radien sehr deutlich zu erkennen (rote Kreise).

Im Glide-Docking konnten von den 4,738 Substanzen des Testdatensatzes 796 (16.8 %) erfolgreich gedockt werden. Die niedrige Anzahl an Dockingergebnissen war nachvollziehbar: Bei dem DUD-E-Testdatensatz handelte es sich explizit um Substanzen, die nicht im Enzym binden sollten. Dies führte bei einer empfindlicheren Bewertung der Posen dazu, dass eine größere Anzahl an Substanzen nicht gedockt wurde. Und obwohl nur 16.8 % des Testdatensatzes gedockt wurden, konnten über 70 % der Aktiven wiedergefunden werden.

Mit diesen Docking-Experimenten konnte gezeigt werden, dass Glide in der Lage ist, Aktive verlässlich zu docken und die Substanzen in einem akzeptablen Bereich anzureichern. Das primäre Ziel, verlässliche Posen zu generieren, konnte mit diesem Vorgehen sichergestellt werden.

– Screening-Datensatz

Für das abschließende Docking des Screening-Datensatzes wurden für jede Struktur die 1,000 Cluster-Leader gedockt, die vom PLP-Score am besten bewertet wurden. Das prinzipielle Vorgehen war identisch zum Testdatensatz: Die Inputzustände wurden beibehalten, gleichzeitig wurden aber auch alle möglichen Zustände mit Epik generiert. Für jeden Ring wurden bis zu fünf Konformere erstellt. Die abschließende Minimierung fand mit dem OPLS_2005-Kraftfeld statt. Für die SP-Dockings wurden für 3G7E_1, 3G7E_2, 1AJ6 und 1EI1 die Grids mit Asp73 als Wasserstoffbrückenbindungsakzeptor und dem Wat A als Wasserstoffbrückenbindungsdonor definiert. 3G7E_2 hatte ein weiteres Pharmakophor mit Wat B als Akzeptor, wieder analog zum Vorgehen bei cRAISE (s. Abschnitt 4.2.1 auf S. 87). Bei jedem Docking mussten alle Pharmakophore erfüllt werden. Die besten fünf Posen jedes Liganden wurden minimiert und die dabei bestbewertete am Ende abgespeichert. Erwartungsgemäß

konnten deutlich mehr Substanzen gedockt werden als beim Testdatensatz (s. Tabelle 5.3). Bei 3G7E_1 und 1AJ6 wurden ca. 94 % erfolgreich

Tabelle 5.3 – Übersicht der Dockingergebnisse der Cluster-Leader.

	3G7E_1	1E11	1AJ6	3G7E_2
Anzahl an Substanzen	1,000	1,000	1,000	1,000
Gedockte Substanzen	943	875	937	520

gedockt, bei 1E11 konnte Glide bei fast 88 % eine plausible Pose generieren. Lediglich bei 3G7E_2 gab es weniger Dockingergebnisse. Dies wurde durch die größeren Einschränkungen des anspruchsvolleren Pharmakophors, das erfüllt werden musste, verursacht.

5.6 Visuelle Auswahl

Bei der visuellen Auswahl wurden die von Glide erstellten Posen analysiert. Dabei wurden verschiedene Parameter verwendet: Neben dem Glide-Score, dem PLP-Score und der „Qualität“ der Pose, waren die chemische Diversität der ausgewählten Substanzen untereinander und zu bereits bekannten Strukturen entscheidend.

Neben der Pose selbst wurden die Protonierungs- und Tautomerisierungszustände der vorliegenden Verbindungen bewertet. Die verschiedenen Zustände wurden zum einen während des cRAISE-Dockings (s. Abschnitt 4.1.1 auf S. 80) und zum anderen bei der Vorbereitung des Glide-Dockings durch Ligprep (s. o.) erstellt. Insbesondere die von cRAISE erzeugten Zustände waren zum Teil kritisch zu bewerten. Dabei traten Tautomere auf, durch die aromatische Systeme kaputt gingen (z. B. Substanz **43** und **44**). Ebenso kamen energetisch ungünstige Keto-/Enolautomerien in Ringsystemen vor.

Auch bei den Protonierungen gab es fragwürdige Zustände, die beurteilt werden mussten. Dabei traten sowohl Deprotonierungen von schwachen Säuren (z. B. Amide (**45**), Sulfonamide) als auch Protonierungen in heterozyklischen, aromatischen Ringsystemen (z. B. **46**) auf. Erschienen diese Zustände nicht plausibel und waren sie für die Interaktionen wichtig, wurden diese Substanzen nicht weiterverfolgt.

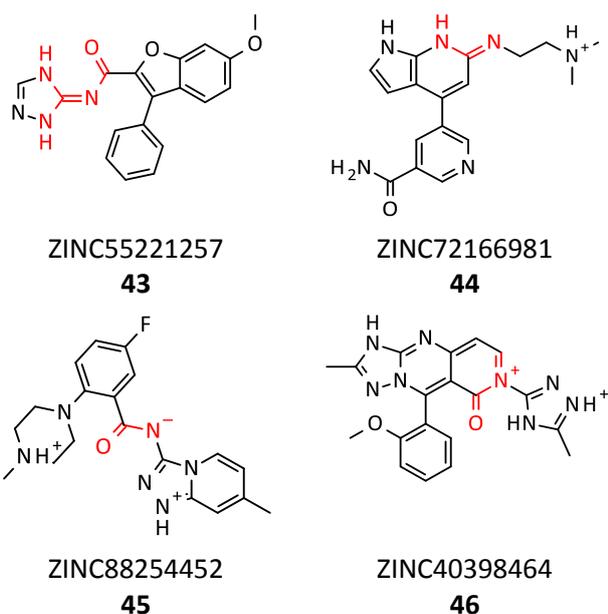


Abbildung 5.10 – Beispiele für kritische Protonierungs- und Tautomerisierungszustände. Bei den Verbindungen sind die zur Diskussion stehenden Zustände rot markiert.

5.6.1 3G7E_1

In Abbildung 5.11 ist die Pose von **47** dargestellt. Dabei handelte es sich um die am besten bewertete Pose des Glide-Dockings für dieses Enzym (GlideScore -11.288 kcal/mol).

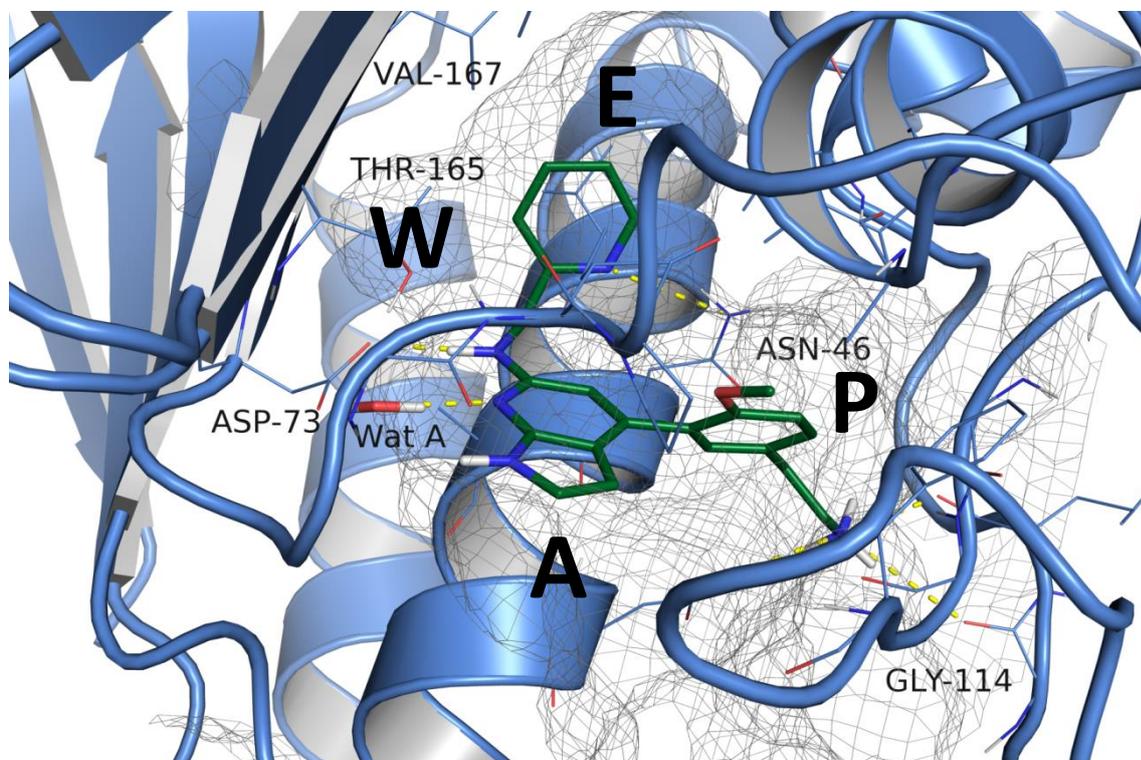


Abbildung 5.11 – Glide-Pose von ZINC72432021 (**47**). Die zentralen Bereiche sind in der Abbildung gekennzeichnet.

Die Abbildung zeigt die Bindetasche und deren Merkmale. Zur Orientierung wurden die Bereiche der Bindetasche definiert: Der Ausgang der Tasche wird als A bezeichnet. P stellt den Bereich dar, in dem sich bei 1E1I die Ribose und der Triphosphat-Rest des ATPs befinden. E ist der Bindungsbereich bei dem sich in 3G7E ein Ethinyl-Rest befindet. W ist der Bereich, in dem Wat B vorkommt. Bei allen Dockingergebnissen stellten Asp73 und Wat A das Hauptpharmakophor (HP) dar (s. auch Abschnitt 2.1 auf S. 39).

Es konnte eine Vielzahl von verschiedenen Strukturelementen gefunden werden, die mit dem Hauptpharmakophor interagierten. Häufig fanden sich Heterozyklen, Aminoheterozyklen oder Strukturelementen, die Carbonyle und in enger Nachbarschaft ein Stickstoffatom Wasserstoffbrücken-Akzeptor und eine Donorgruppe in enger Nachbarschaft zueinander enthielten (s. Abbildung 5.12). Zwischen dem Akzeptor- und Donoratom des Liganden waren damit zwischen null (bspw. Pyrazol-Derivate) und drei Atome Abstand (Pyrrol-2-yl-amide) zu finden.

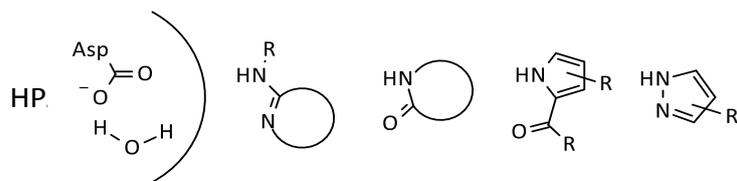


Abbildung 5.12 – Häufig aufgetretene Interaktionspartner für das Hauptpharmakophor (HP) beim Glide-Docking in 3G7E_1.

Hydroxyl-Funktionen agierten nur sehr selten als Donorgruppe. Und wenn sie vorkamen, war die Geometrie für diese Interaktion unzureichend, so dass diese Substanzen nicht in die weitere Auswahl kamen. Dies wird anhand der Pose von **48** deutlich, die mit einem GlideScore von -9.519 die bestbewertete Pose mit einem Hydroxyl-Pharmakophor darstellte (s. Abbildung 5.13). Die Winkel der Akzeptor- und Donorgruppen für das Hauptpharmakophor entsprachen nicht den Erwartungen, auch wenn die Abstände zwischen den schweren Atomen (2.7 Å und 3.3 Å) akzeptabel waren. Dies ist insofern beachtenswert, als bei dem einzig zuvor durchgeführten virtuellen Screening, gezielt nach Phenolderivaten gesucht wurde (s. Abschnitt 1.3 auf S. 22).

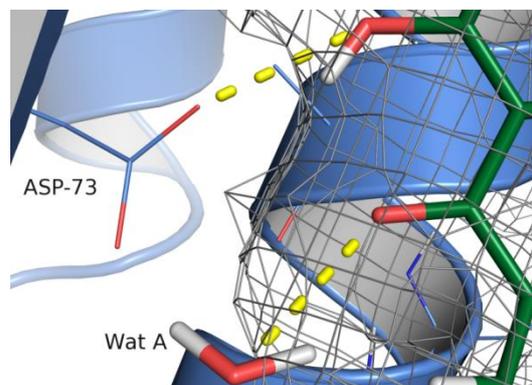


Abbildung 5.13 – Glide-Pose von ZINC06651910 (**48**). Bei dieser Struktur stellte eine Hydroxyl-Gruppe den Donor für Asp73 dar.

Eine wichtige Besonderheit in der 3G7E Struktur stellte die E-Tasche dar. Diese ist bei 3G7E besonders stark ausgeprägt (s. Abschnitt 2.1 auf S. 37). Bei vielen Docking-Ergebnissen war diese Tasche durch den Liganden besetzt. Dabei gab es deutliche Unterschiede zu den bisher bekannten GyrB-Inhibitoren. Die ausführlichsten Untersuchungen wurden dazu von RONKIN *et al.* durchgeführt [88]. Dabei wurden aber hauptsächlich kleine oder lineare Fragmente verwendet (Methyl-, Ethinyl- oder Propinylreste), um diese Seitentasche zu besetzen. Andere Arbeitsgruppen führten noch kleinere Fragmente ein (s. Abschnitt 1.3 auf S. 14). Die E-Tasche war bei dem Glide-Docking häufig mit Ringsystemen besetzt. Ein häufiges Fragment war ein Pyridyl-Rest, z. B. in Abbildung 5.11 auf S. 124 gezeigt. Die Tasche wurde auch von anspruchsvolleren Resten belegt (s. Abbildung 5.14). Dies ermöglichte weitere Interaktionen mit dem Protein. In Abbildung 5.14 gab es eine gute Wasserstoffbrücke

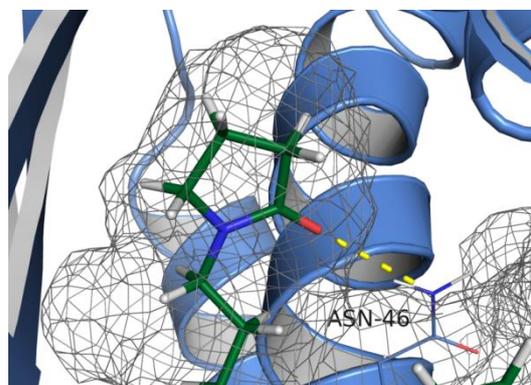


Abbildung 5.14 – Glide-Pose von ZINC12232705 (**49**). Die E-Tasche ist mit einem ungesättigten γ -Lactam besetzt.

zwischen dem Sauerstoff des Lactamrings von **49** und Asn46 (Entfernung 3.1 Å). Bisher gibt es keine Inhibitoren mit derart ausgeprägten Ringsystemen in diesem Bereich.

In der W-Tasche gab es nur wenige Variationen. Häufig war diese Position nicht besetzt. In einigen Fällen befanden sich dort kleine hydrophobe Teilstrukturen, wie z. B. Methyl- oder Ethyl-Reste.

Die Teilstrukturen, die in die P-Tasche zeigten, waren sehr variabel. Dies deckte sich mit den literaturbekannten Strukturen. Auch dort sind viele Teilstrukturen bekannt, die in diesem Bereich binden. Generell wurde dieser Bereich, da hier die meiste Flexibilität in der Bindetasche besteht, häufiger zur Optimierung der Struktur verwendet, bspw. für die Verbesserung der Pharmakokinetik [83]. Bei den Docking-Ergebnissen von 3G7E_1 wurde die P-Tasche teilweise vor der A-Tasche, also dem Ausgang des Enzyms, adressiert. Substanz **50** besaß bspw. sowohl in der E-

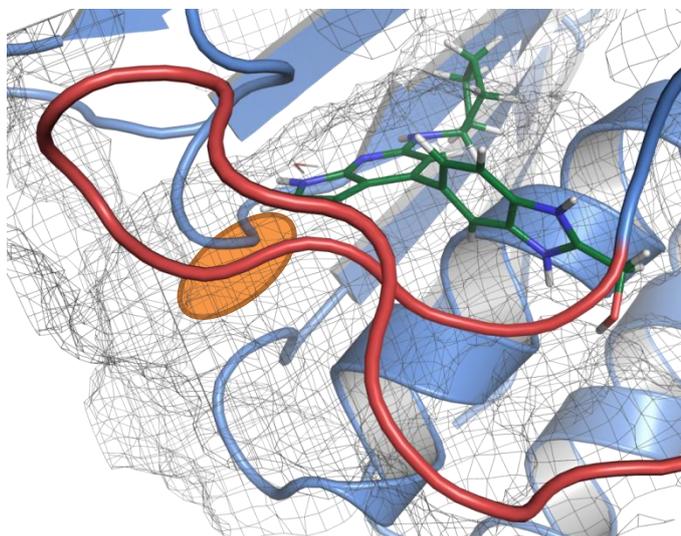


Abbildung 5.15 – Glide-Pose von ZINC72480381 (**50**). Bei dieser Pose lag ein großer Rest in der P-Tasche, während der Ausgang (orange markiert) nicht belegt war. Der blockierende Loop B ist rot dargestellt.

als auch in der P-Tasche einen großen Rest, während der A-Bereich unbesetzt war (s. Abbildung 5.15). Wenn diese Substanz in dieser Pose binden soll, blockiert Loop B den Zugang für den Liganden. Dies setzt folglich einen starken und schwer vorhersagbaren *Induced Fit* voraus.

Der A-Bereich ist durch seine geladenen Aminosäuren gekennzeichnet (Arg76, Arg136, Glu50, Asp106, s. Abschnitt 2.1 auf S. 27). Bei 3G7E trat die Besonderheit auf, dass dieser Bereich aufgrund des umgeklappten B-Loops sterisch eingeschränkt war. Aus diesem Grund war die A-Tasche schwer adressierbar. Dies führte dazu, dass die meisten Liganden beim Docking eher die P-Tasche besetzten und nicht die A-Tasche (s. o.). In Abbildung 5.16 ist die Pose von **51** dargestellt, welches im Ausgangsbereich einen Triazol-Ring besitzt. Dieser ist im Vergleich zum Pyridin-Ring, der beim Liganden der Kristallstruktur vorkommt, kleiner. Dennoch passt der Triazol-Ring nur knapp hinein. Der Bereich wurde insbesondere durch den Backbone-Carbonylsauerstoff von Phe104 und der Seitenkette von Arg76 eingeschränkt (orange gestrichelten Linie).

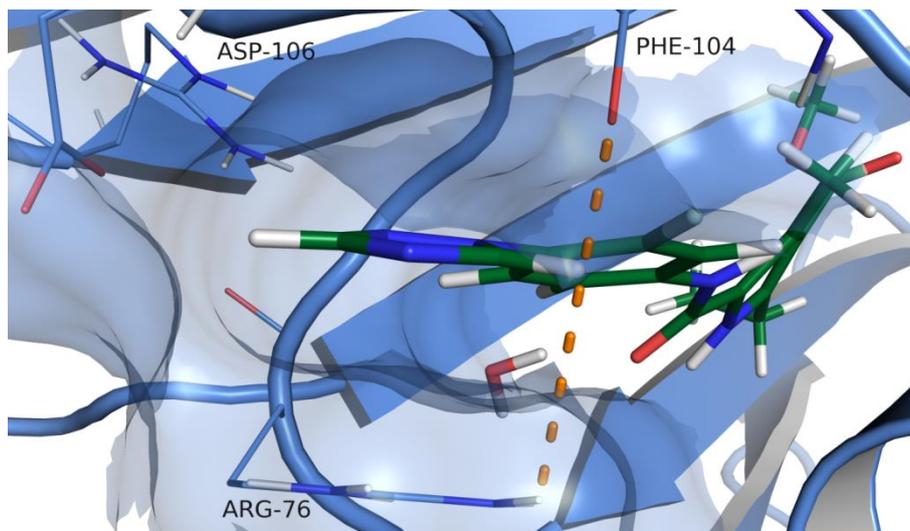


Abbildung 5.16 – Glide-Pose von ZINC58175024 (**51**). Der A-Bereich der Tasche ist sterisch gehindert. Auch der vergleichbar kleine Triazolring der Substanz passt knapp in die Tasche. Die gestrichelte Linie (orange) zeigt die stärkste Verengung in diesem Bereich.

Abbildung 5.17 zeigt die erste Auswahl von zwölf Substanzen, die aufgrund der beschriebenen Beobachtungen selektiert wurden. Die Auswahl dient vor allem dazu, einen allgemeinen Überblick über die Liganden und deren Posen zu geben. Bei **52** – **56** handelt es sich um Pyrrolamide, einer Gruppe die schon von EAKIN *et al.* beschrieben wurde [77], allerdings ohne die Carbonylfunktion an Position 5 des Pyrrols. Zwar handelt es sich nicht um ein komplett neues Hauptfragment, es bestätigt aber grundsätzlich das Vorgehen, wenn auch bekannte Fragmente angereichert und gut bewertet werden. Zudem waren die weiteren Substitutionen sehr variabel und unterschieden sich deutlich von den bekannten Substanzen. Ähnliches gilt für die gefundenen Pyrazole (**59** und **60**), die in einem ähnlichen Interaktionsmuster von RONKIN *et al.* beschrieben wurden [88]. Es konnten aber auch vielversprechende bisher noch nicht beschriebene Strukturen gefunden werden, die mit dem Hauptpharmakophor interagieren. Dabei sind als Vertreter einer neuen Klasse insbesondere die Xanthinderivate **57** und **58** zu nennen, aber auch verschiedene Substanzen mit neuen Interaktionsmustern (**61** – **63**).

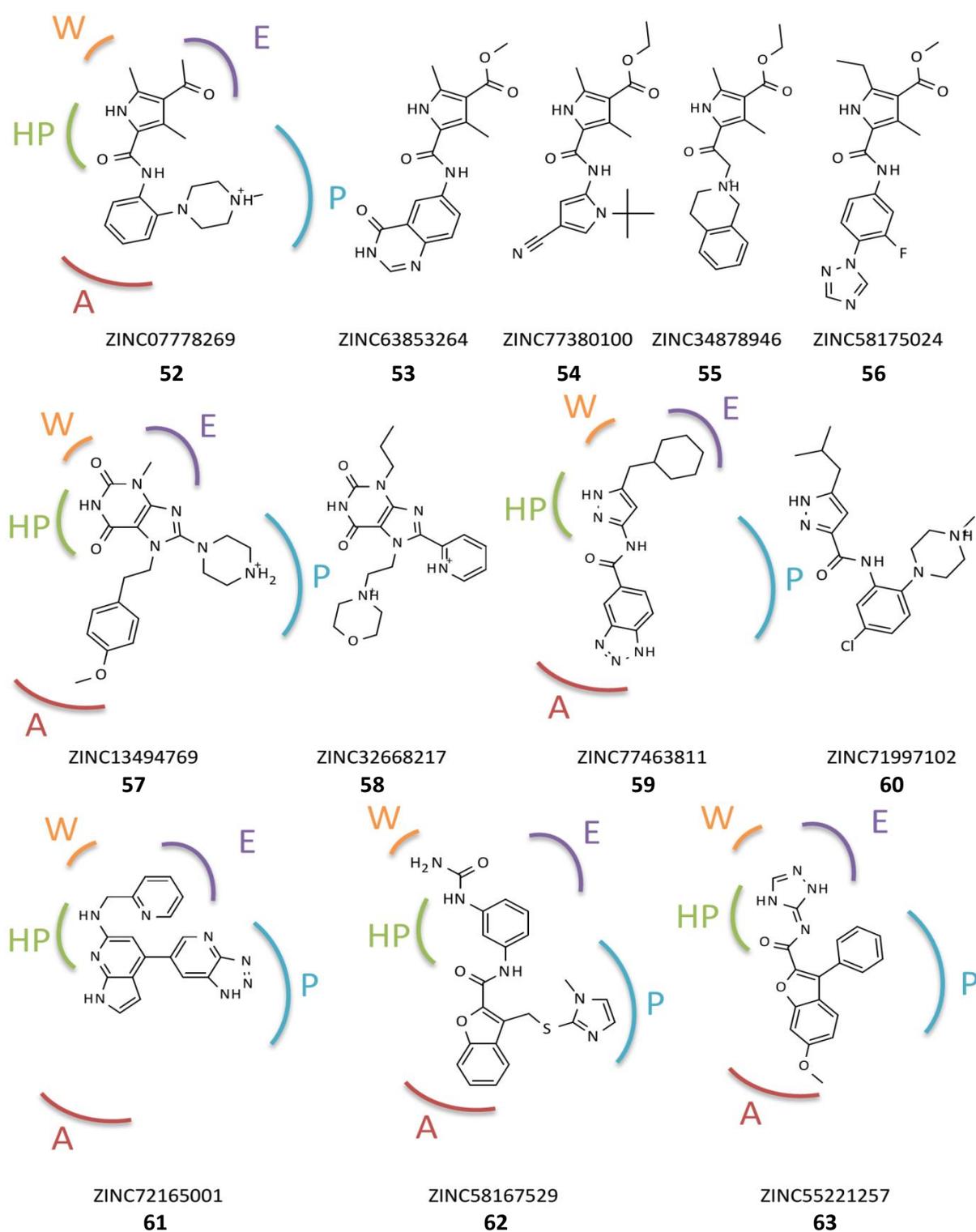


Abbildung 5.17 – Erste Auswahl interessanter Vertreter vom 3G7E_1-Docking mit Glide. Die Substanzen sind entsprechend ihrer Docking-Pose ausgerichtet. Zur Orientierung sind die Bindungsbereiche der Bindetasche dargestellt.

5.6.2 3G7E_2

3G7E_2 war in seiner Struktur nahezu identisch zu 3G7E_1. Der einzige Unterschied bestand im Wat B, das hier zusätzlich vorhanden war. Beim Docking musste der Ligand dabei, wie oben beschrieben, zusätzlich als Wasserstoffbrückendonator für dieses Wasser fungieren.

Von den Bindetascheneigenschaften und den damit gefundenen Substanzen war der Unterschied zu 3G7E_1 nur gering. Die meisten dort diskutierten Beobachtungen trafen auch hier zu.

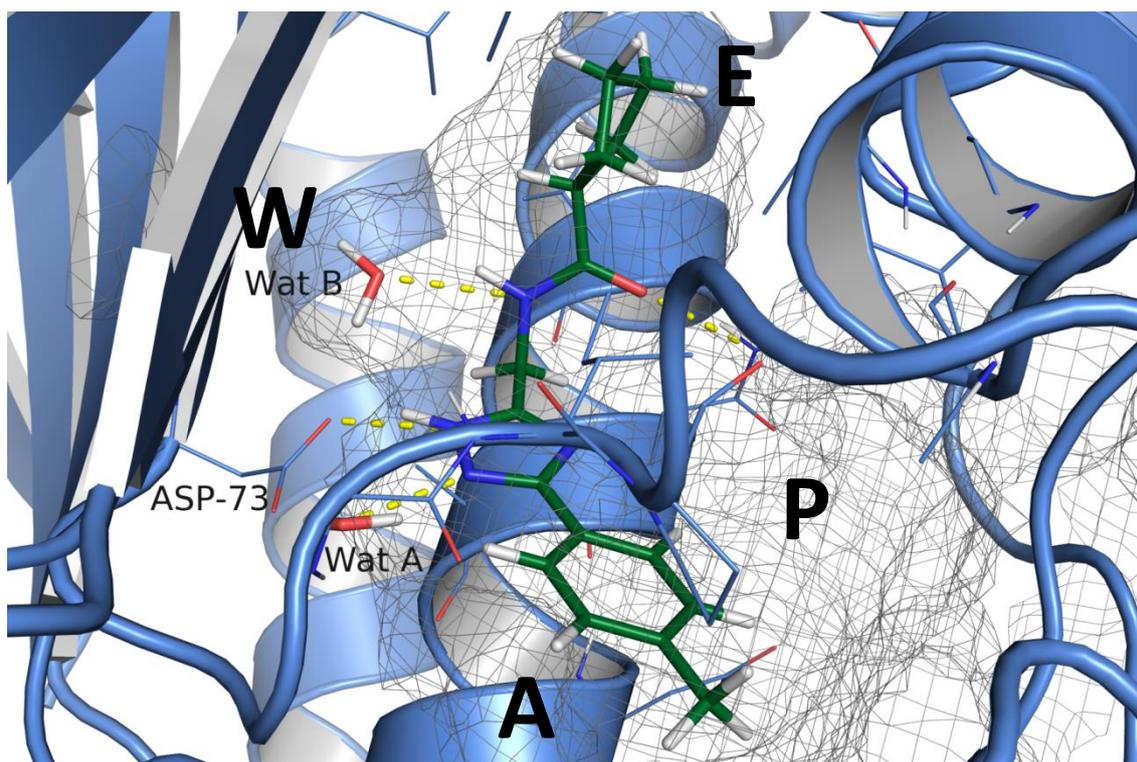


Abbildung 5.18 – Glide-Pose von ZINC13565291 (64) in 3G7E_2. Die zentralen Bereiche sind in der Abbildung gekennzeichnet.

Abbildung 5.18 zeigt eine der bestbewerteten Posen (GlideScore: -10.206 kcal/mol) und die zugehörigen Bereiche.

Die Hauptpharmakophore waren ebenfalls vergleichbar zu denen von 3G7E_1. Insbesondere die Pyrazol- und Aminopyridinderivate dominierten bei den Docking-Ergebnissen (s. Abbildung 5.19).

Dabei nahm die Komplexität der Moleküle zu, weil das zu erfüllende Pharmakophor anspruchsvoller war. Im Vergleich zum 3G7E_1-Docking traten mehr Posen auf, welche die E-Tasche weiter ausfüllten. Dies ließ sich dadurch erklären, dass durch die zusätzliche Interaktion Substanzen präferiert wurden, die tiefer in der Tasche binden.

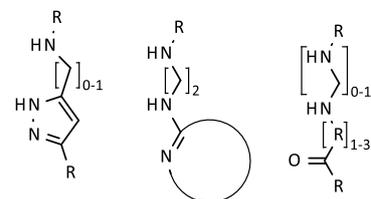


Abbildung 5.19 – Häufig aufgetretene Interaktionspartner für das Hauptpharmakophor beim Glide-Docking in 3G7E_2.

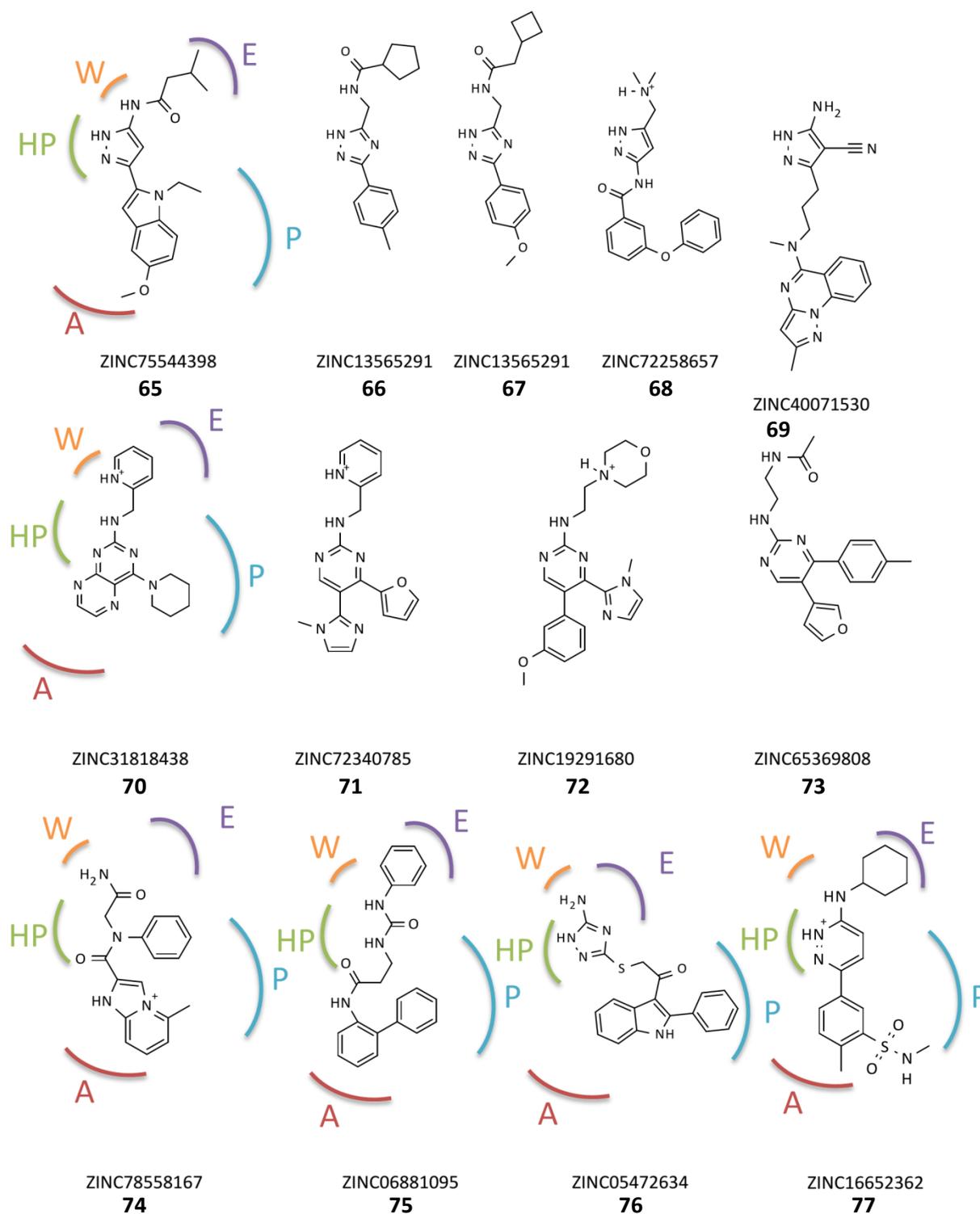


Abbildung 5.20 – Erste Auswahl interessanter Vertreter vom 3G7E_2-Docking mit Glide. Die Substanzen sind entsprechend ihrer Docking-Pose ausgerichtet. Zur Orientierung sind die Bindungsbereiche der Bindetasche dargestellt.

In Abbildung 5.20 ist eine Auswahl von interessanten Vertretern des Glide-Dockings in 3G7E_2 aufgeführt. Aufgrund des anspruchsvolleren Pharmakophors wiesen nahezu alle Strukturen neue, noch nicht in der Literatur beschriebene Eigenschaften auf. Viele Substanzen hatten jedoch sehr große Reste in der E-Tasche (z. B. Substanz **77**), die sehr viel Platz einnahmen. Wie bei 3G7E_1 treten einige interessante Pyrazole (**65 – 69**) auf. Ebenfalls häufig konnten 2-Aminopyrimidine (**70 – 73**) dieses strenge Pharmakophor erfüllen, neben weiteren sehr unterschiedlichen Strukturen (**74 – 77**).

5.6.3 1E11

Abbildung 5.21 zeigt die bestbewertete Pose des Glide-Dockings in 1E11 (s. Abbildung 5.21, GlideScore -11.229 kcal/mol). Die wichtigsten Unterschiede liegen in der E- und P-Tasche. Die E-Tasche war sehr schwach ausgeprägt. Die P-Tasche war dafür größer und deutlich besser für hydrophile Interaktionen präformiert (s. Abschnitt 2.1 auf S. 29). Die Posen erhielten deutlich höhere Score-Werte als bei allen anderen Dockings. Die 1,000ste Pose hatte noch immer einen Glide-Score von -8.356 kcal/mol. Beim Docking von 3G7E_2 besaß bspw. die 134ste Pose einen vergleichbaren Score mit -8.351 kcal/mol. Dies lag darin begründet, dass die neuen hydrophilen Interaktionen innerhalb der P-Tasche einen starken Beitrag zum Score beisteuerten.

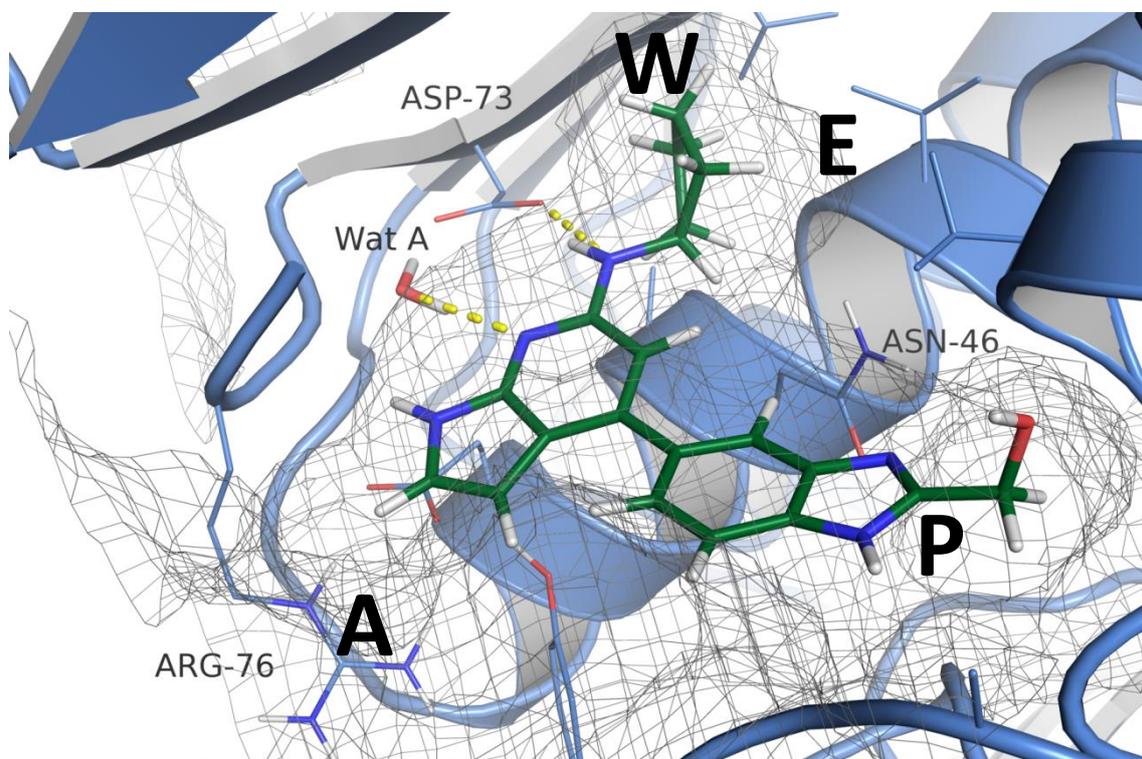


Abbildung 5.21 – Glide-Pose von ZINC72480381 (**78**) in 1E11. Die zentralen Bereiche sind in der gekennzeichnet. Loop A ist zur besseren Übersicht ausgeblendet.

Im Vergleich zu den bisher aufgetretenen Interaktionspartnern für die Hauptpharmakophore gab es einige Gemeinsamkeiten mit den Ergebnissen von 3G7E. Dies war nicht überraschend, da dieser Bereich bei den

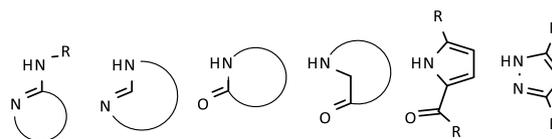


Abbildung 5.22 – Häufig aufgetretene Interaktionspartner für das Hauptpharmakophor beim Glide-Docking in 1E11.

verschiedenen Strukturen sehr stark konserviert war. Bei den ähnlichen Strukturmerkmalen waren insbesondere die Pyrrolamide, Pyrazole und Aminopyrimidine vertreten. Dennoch gab es auch Unterschiede. Die vorkommenden Pharmakophore waren (s. Abbildung 5.22) deutlich variabler und wechselten sich häufiger als bei den anderen Dockings ab. Weiterhin kamen auch neue Pharmakophore vor, wie α -Aminocarbonylderivate. Auch traten häufiger Alkoholfunktionen auf, die das Pharmakophor erfüllten. Die Geometrien dieser Verbindungen waren jedoch wie bei 3G7E_1 (s. Abbildung 5.13 auf S. 125) zu schlecht, um in eine weitere Auswahl zu gelangen.

Die W- und E-Tasche wurden nur noch von einem gemeinsamen Rest belegt, da insbesondere die E-Tasche nur noch schwach geprägt war. Sofern diese Bereiche der Bindetasche besetzt waren, platzierten sich ausschließlich hydrophobe Reste, zumeist kurze Alkylketten, dort. In Abbildung 5.23 ist die Glide-Pose von **79** gezeigt. Dabei war ein Cyclopropyl-Rest in diesem Bereich. Neben aliphatischen Resten gab es auch aromatische Systeme, wenn bspw. ein protoniertes Pyridiniumkation mit dem Asp73 interagierte.

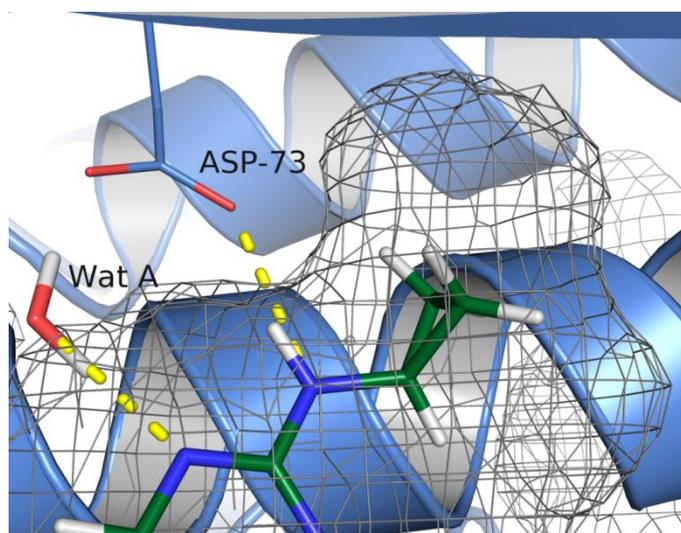


Abbildung 5.23 – Glide-Pose von ZINC19582299 (**79**). Der W- und E-Bereich der Tasche war mit einem Cyclopropyl-Rest belegt.

Die P-Tasche wurde noch deutlich stärker als bei den Dockings mit 3G7E betont. Wie in Abschnitt 2.1 auf S. 31 beschrieben, war diese auch deutlich stärker ausgeprägt. Folglich war die Bindetasche für viele gute hydrophile Interaktionen präformiert. In Abbildung 5.24 ist die Glide-Pose von **80** dargestellt. In dieser Struktur sind der Furan-Ring und das Amid sehr tief in der P-Tasche verankert. Dies ermöglichte insbesondere hydrophile Interaktionen mit den polaren Seitenketten. Mit einem Glide-Score von -11.040 kcal/mol wurde dies auch sehr positiv bewertet.

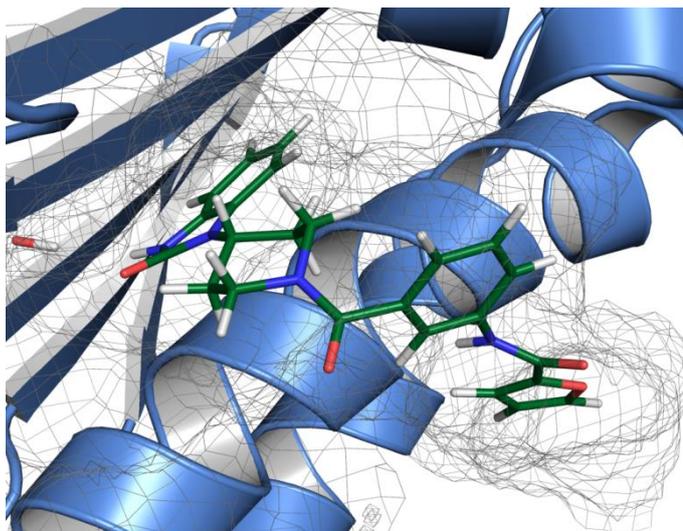


Abbildung 5.24 – Glide-Pose von ZINC59368683 (**80**). Auffällig ist die Bindung der Substanz in der tiefen P-Tasche.

Im A-Bereich wurden ähnlich wie beim 3G7E-Docking seltener Reste platziert, da die P-Tasche vor dem Ausgang des Proteins bevorzugt wurde. Dieser Effekt war hier noch deutlich größer, da die P-Tasche noch viel besser für diese Interaktionen ausgeprägt war. Hinzu kam, dass ein wichtiges Interaktionsmuster in dieser Struktur nur schwer zustande kam: Das Arg76 war nicht planar zum Ausgang der Tasche ausgerichtet, so dass das häufig vorkommende π -Kationen-Stacking nicht möglich war. Dies wird an

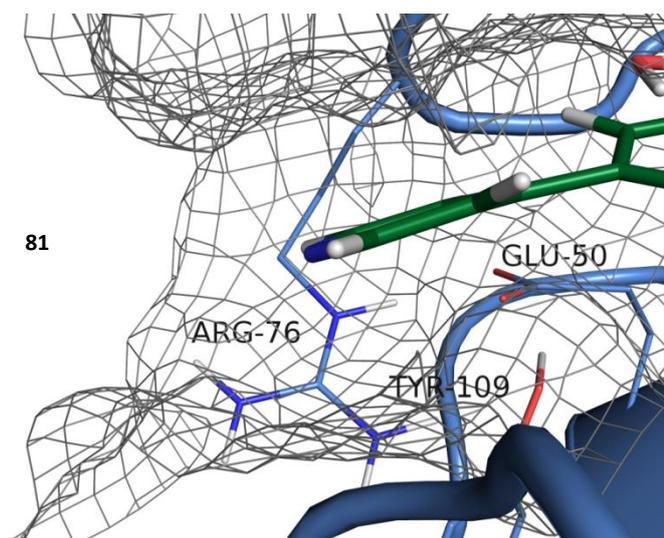


Abbildung 5.25 – Glide-Pose von ZINC19582299 (**81**). Der Pyridylrest bildete kein π -Kationen-Stacking mit Arg76 aus. Dies ist an der nichtvorhandenen Coplanarität zu erkennen.

Substanz **81** (s. Abbildung 5.25) deutlich. Obwohl sich am Ausgang ein Pyridin-Ring befand, welcher auch ein bekanntes und gutes Pharmakophor für diesen Bereich darstellte, war das typischerweise vorkommende π -Kationen-Stacking nicht möglich. Dafür gab es zwei Gründe. Einerseits drehte sich das Arg76 leicht aus der Tasche Richtung Lösungsmittel heraus (normalerweise befand sich das Arg76 planar zum Glu50) und andererseits drehte sich das Tyr109 vom Loop B in die Tasche Richtung Glu50 rein und stellte damit eine sterische Behinderung zum Arg76 dar. Damit erschwerte der Loop B also die Ausbildung der Interaktionen im A-Bereich bei 1E11, obwohl er besser zugänglich aussah als bei 3G7E.

In Abbildung 5.26 sind Posen von diesem Docking dargestellt. Sofort auffällig ist die ausgeprägte Bindung in dem P-Bereich von jeder Substanz. Neben den schon bekannten Pyrazolen (**152– 156**) und 2-Aminopyrimidinen (89 – 91), kamen das erste Mal Indole und Indolderivate (82– 87) vor. Dies gehörte auch zu Fragmenten die von AstraZeneca identifiziert wurden [76]. Daneben traten jedoch noch viele unterschiedliche Posen auf. Erwähnenswert sind insbesondere die Xanthine (97 und 98), die auch bei 3G7E_1 (s. Abschnitt 5.6.1 auf S. 128) beobachtet wurden.

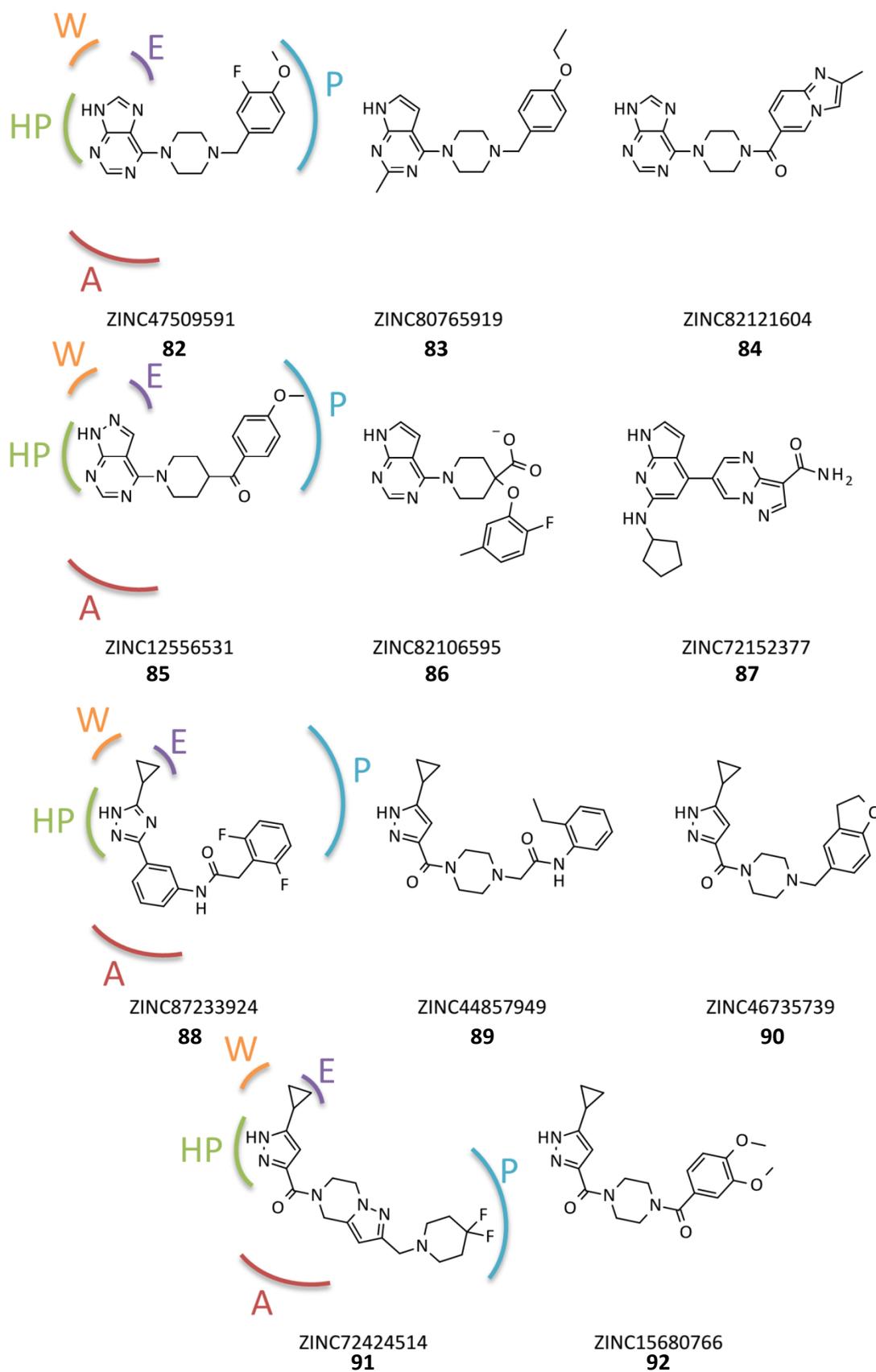


Abbildung 5.26 – Erste Auswahl interessanter Vertreter vom 1E11-Docking mit Glide. Die Substanzen sind entsprechend ihrer Docking-Pose ausgerichtet. Zur Orientierung sind die Bindungsbereiche der Bindetasche dargestellt.

Fortsetzung nächste Seite

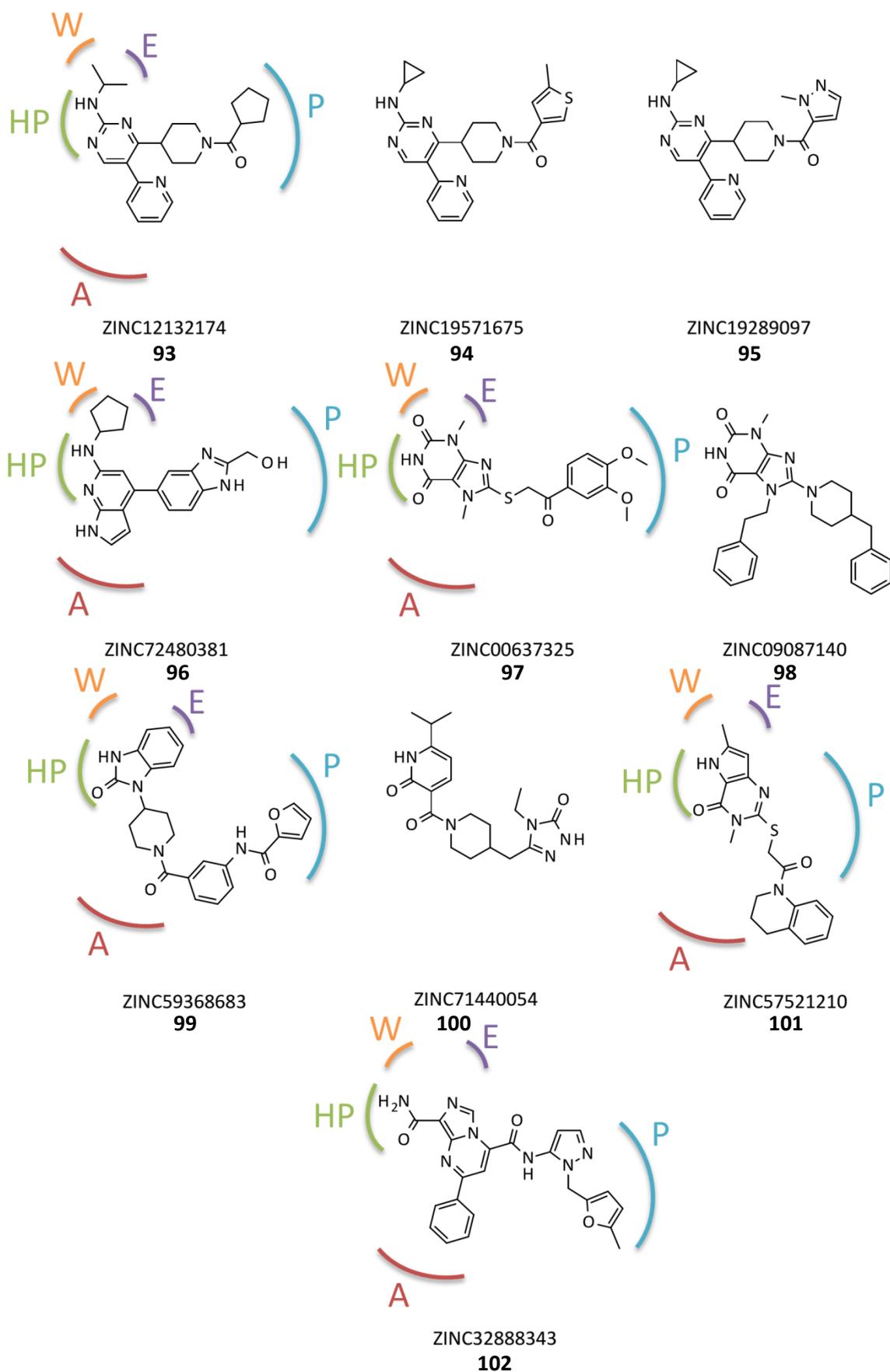


Abbildung 5.26 – Fortsetzung.

5.6.4 1AJ6

Die beste Pose des Glide-Dockings in 1AJ6 ist in Abbildung 5.27 gezeigt. Im Unterschied zu den anderen Proteinstrukturen handelte es sich hier um die minimierte Struktur einer MD-Simulation (s. Abschnitt 2.2.2 auf S. 44). Die W-Tasche war ähnlich ausgebildet wie bei den anderen Kristallstrukturen. Die E-Tasche wies die größte Ähnlichkeit mit der von 1AJ6 auf und war nur sehr schwach ausgeprägt. Die P-Tasche war in dieser Struktur am schlechtesten zugänglich. Dies wurde vor allem durch Asn46 verursacht, das im Unterschied zu den anderen Strukturen verdreht war und den Zugang erschwerte. Der A-Bereich hingegen war deutlich offener. Aufgrund der Konformation des B-Loops, gab es in diesem Bereich sterisch mehr Platz (s. Abschnitt 2.1 auf S. 32).

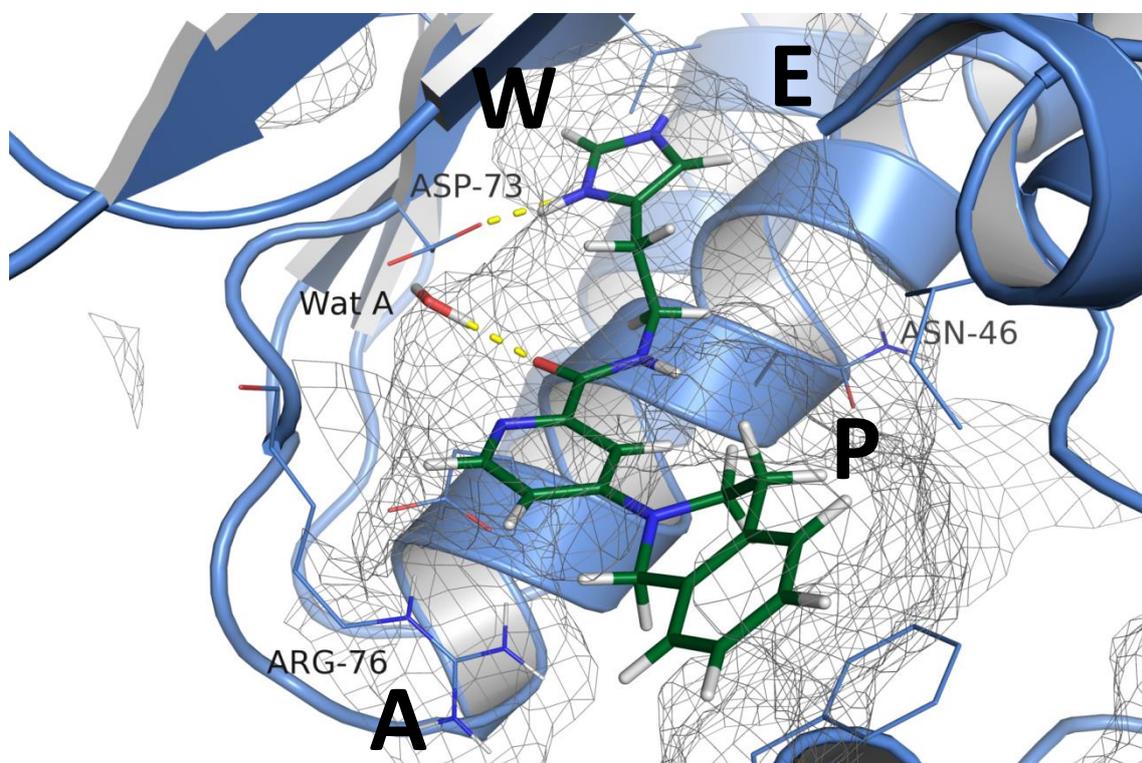


Abbildung 5.27 – Glide-Pose von ZINC72165365 (**103**) in 1AJ6. Die zentralen sind in der Abbildung gekennzeichnet. Loop A ist zur besseren Übersicht ausgeblendet.

Die Hauptpharmakophore verhielten sich ähnlich zu den bisherigen Docking-Ergebnissen (s. Abbildung 5.28). Sowohl die schon aus dem 3G7E_1-Docking bekannten Pyrrolamide als auch die in allen Dockings gefundenen

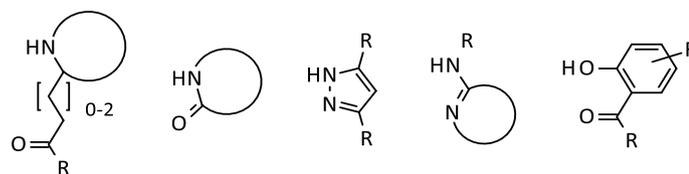


Abbildung 5.28 – Häufig vorkommende Hauptpharmakophore beim Glide-Docking von 1AJ6.

Pyrazole konnten wiedergefunden werden. Dennoch gab es auch neue Interaktionsmuster. Diese bestanden aus längeren Strukturen, die bspw. über die W-Tasche eine Interaktion mit dem Asp73

ausgebildet. Ein solcher Fall ist in Abbildung 5.27 zu sehen, bei dem der Wasserstoffbrückendonator (Imidazolkörper) zum –akzeptor (Amid) beim Liganden einen Abstand von 5 Atomen hatte.

Die W- und E-Tasche waren häufig, wie bei 1E11, mit einem einzigen Strukturelement besetzt. Dabei waren die vorkommenden Reste in der Regel größer als in der 1E11-Struktur aber noch immer deutlich kleiner als bei den Dockings mit 3G7E. Es kamen häufiger aromatische Systeme vor, die diesen Bereich besetzen (s. Abbildung 5.29), während gesättigte Strukturen, wie Isopropyl oder *tert.*-Butyl, seltener auftraten. Zusätzlich wurden vermehrt hydrophile Interaktionen von der W-Tasche zum Asp73 beobachtet.

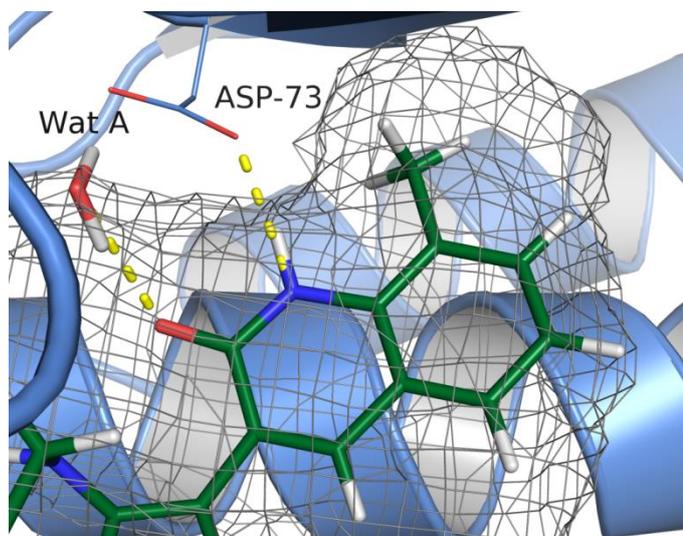


Abbildung 5.29 – Glide-Pose von ZINC01330094 (**104**). Die W/E-Tasche ist mit einem aromatischen System besetzt.

Möglicherweise führte die MD-Simulation mit der Minimierung dazu, dass das Asp73 besser zu diesem Bereich ausgerichtet wurde.

Die P-Tasche war im Vergleich zu allen anderen Dockings am seltensten besetzt. Verursacht wurde dies durch das Asn46, das leicht verdreht war und den Zugang erschwerte. In diesem Bereich zeigte auch Novobiocin keine Interaktion auf. In Abbildung 5.30 sind die 500 besten Posen des Dockings übereinandergelegt dargestellt. Nur wenige Strukturen zeigten Interaktionen in der tiefen A-Tasche. In der Regel traten leicht basische bis basische Amine (z. B. Piperidin) auf, die mit den hydrophilen Bereichen interagierten.

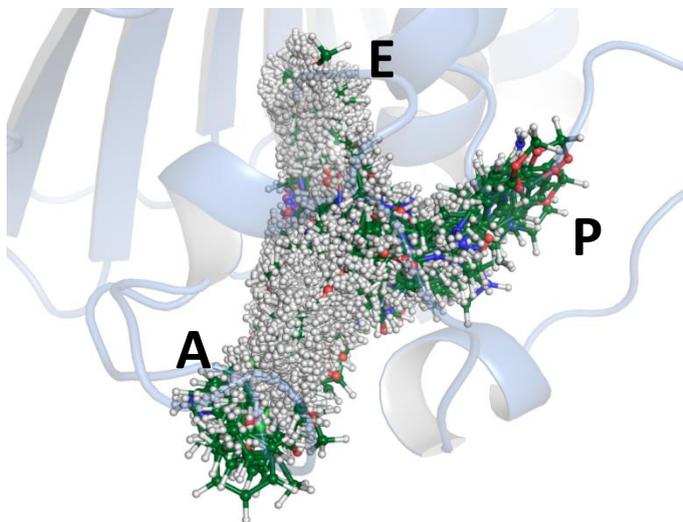


Abbildung 5.30 – Top 500 Glide-Posen vom Docking in 1AJ6. Das Enzym ist zur Orientierung transparent dargestellt. Es befanden sich nur sehr wenige Dockingergebnisse in der P-Tasche.

Der Ausgangsbereich wurde dafür im Vergleich zu den anderen Dockings sehr häufig besetzt (s. auch Abbildung 5.30). Es gab Substanzen, die nur den A-Bereich belegten, Interaktionen mit dem Loop A ausbildeten oder weiter in Richtung des B-Loops banden. Dabei trat die Interaktion mit Loop A am häufigsten auf. Diese kann auch bei Novobiocin (Ligand der Kristallstruktur) beobachtet werden. Bei der Beurteilung der Posen ist zu beachten, dass beide Loops gemodelt werden mussten, da sie in der Röntgenstrukturanalyse nicht beobachtet werden konnten (s. o.).

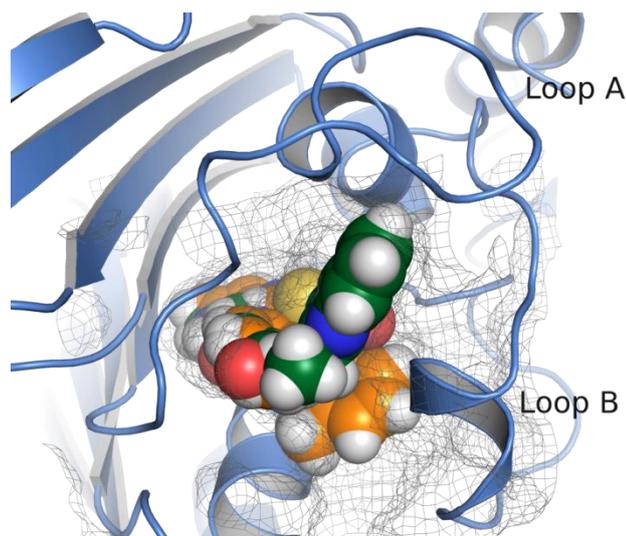


Abbildung 5.31 – Vergleich der verschiedenen Posen am Ausgang der Bindetasche. ZINC11665033 (**105**; sphärisch dargestellt, grün) bindet wie Novobiocin Richtung Loop A. ZINC67237052 (**106**; orange) hingegen interagiert mit Loop B.

Bei der Auswahl der Substanzen in Abbildung 5.32 ist die dominante Bindung im A-Bereich deutlich. Des Weiteren besitzen viele Substanzen in der Bindetasche eine weitere hydrophile Interaktion mit dem HP über die W-Tasche (z. B. sehr ausgeprägt bei **111** – **115**). Insgesamt traten bei diesem Docking sehr viele unterschiedliche und neue Posen auf. Dies ist auch das erste Docking, bei dem Phenole (**107** und **108**) mit guten Geometrien beobachtet werden konnten.

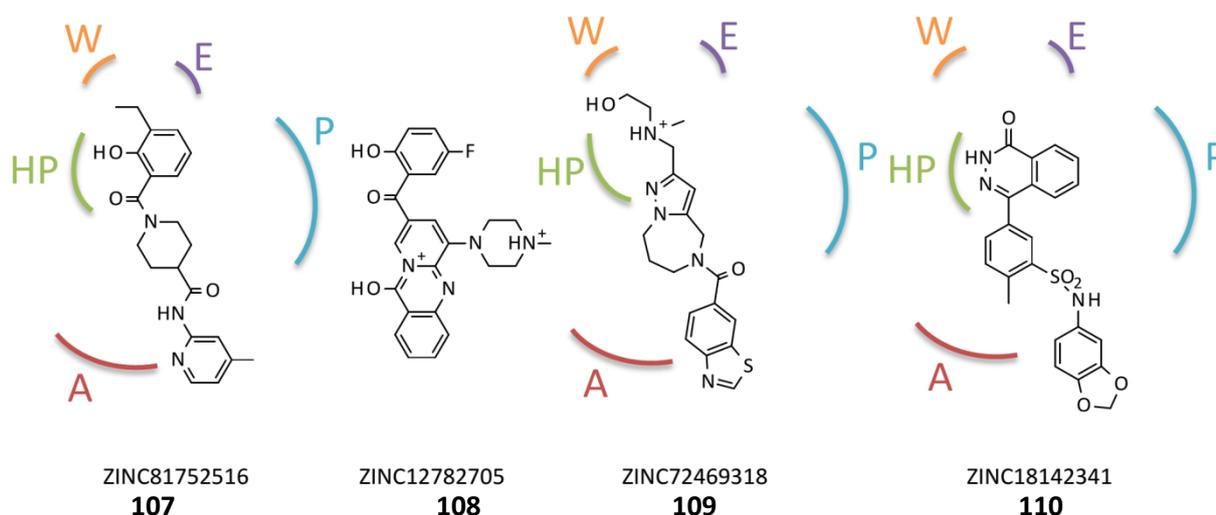


Abbildung 5.32 – Erste Auswahl interessanter Vertreter vom 1AJ6-Docking mit Glide. Die Substanzen sind entsprechend ihrer Docking-Pose ausgerichtet. Zur Orientierung sind die Bindungsbereiche der Bindetasche dargestellt.

Fortsetzung nächste Seite

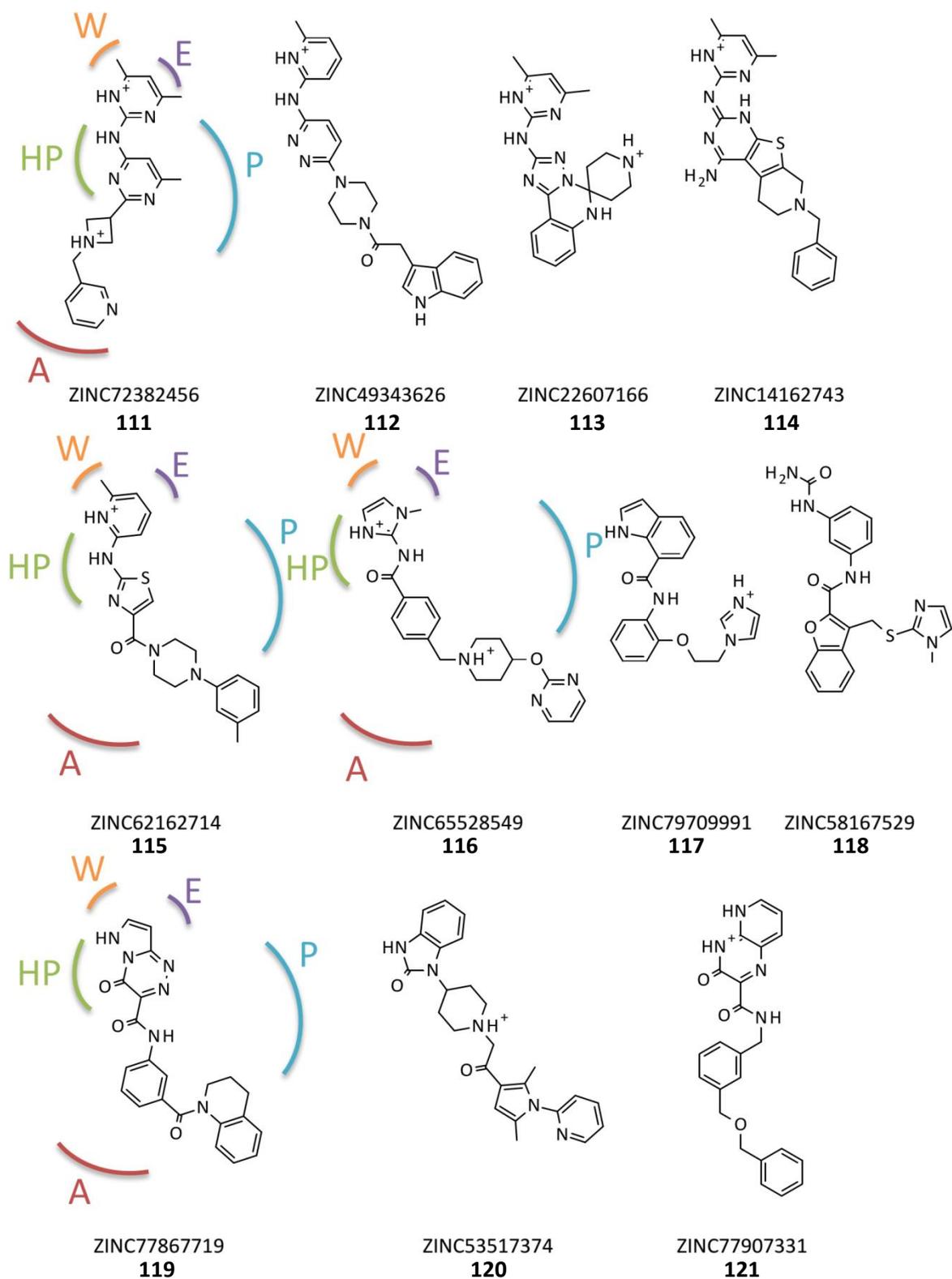
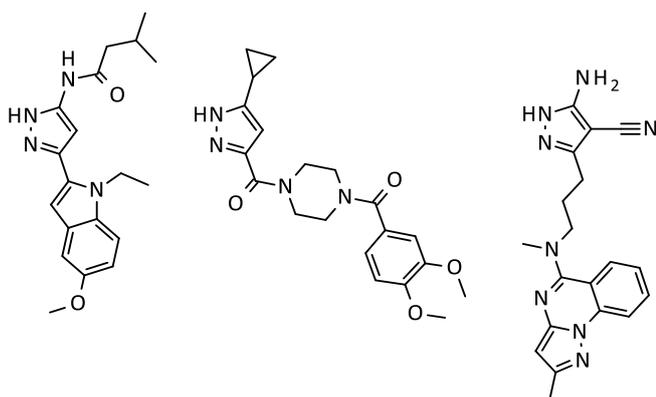


Abbildung 5.32 – Fortsetzung.

5.6.5 Erste Auswahl

Für die finale Auswahl der Substanzen waren neben den beschriebenen Vor- und Nachteilen der einzelnen Posen auch die Erwerbbarkeit und der Vergleich zu den im Cluster vorhandenen Strukturen ausschlaggebend. Insgesamt wurden 20 Substanzen käuflich erworben. Die einzelnen Substanzen sind nach dem Hauptpharmakophor sortiert.

Pyrazole waren eine häufig vorkommende Substanzgruppe. Es wurden drei verschiedene Derivate mit einem ähnlichen Hauptpharmakophor, aber ansonsten sehr unterschiedlichen Substitutionen ausgewählt (s. Abbildung 5.33). **65** und **69** wurden erfolgreich in 3G7E_2 gedockt und bildeten daher eine Wasserstoffbrücke mit dem in der Bindetasche vorhandenen Wat B. **92** wurde in 1E11 gedockt.



ZINC75544398

65

Abbildung 5.33 – Erworbene Substanzen mit einem Pyrazol-Körper als Grundkörper.

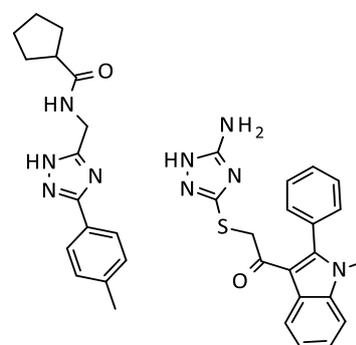
ZINC15680766

92

ZINC40071530

69

Eine sehr ähnliche Gruppe bestand aus Triazol-Körpern (s. Abbildung 5.34), die auch erfolgreiche Dockings von 3G7E_2 darstellten. Hier wurde ein Derivat (**66**) mit einem relativ großen Rest in der E-Tasche und ein Derivat (**122**) mit einem kurzen Rest in dieser Region. Bei **122** handelte es sich um ein Clustermittglied von **76**, das nicht verfügbar war.



ZINC13565291

66

ZINC06507886

122

Abbildung 5.34 – Erworbene Substanzen mit einem Triazol-Körper als Grundkörper.

Interessante Pyrrolamide (s. Abbildung 5.35) kamen vor allem beim Docking in 3G7E_1 vor. Alle Derivate besaßen an Position 4 eine Methylgruppe und eine Carbonylfunktion an Position 5 vom Pyrrol. Es wurden Derivate, die die E-Tasche nur schwach besetzten und unterschiedliche weitere Interaktionen ausbildeten, ausgewählt.

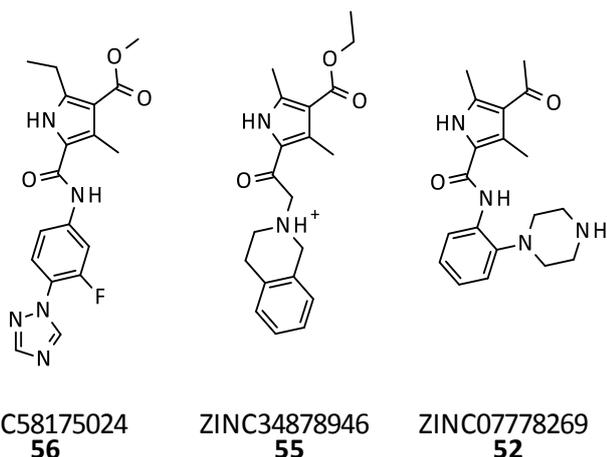


Abbildung 5.35 – Erworbene Substanzen mit einem Pyrrolamid-Körper als Grundkörper.

Von den Xanthin-Derivaten (s. Abbildung 5.36) wurden zwei Strukturen aus verschiedenen Docking-Ergebnissen ausgesucht. **57** wurde erfolgreich in 3G7E_1 gedockt und **123** gehörte zum Cluster von **97** des 1E11-Dockings. Beide Ergebnisse zeigten sehr unterschiedliche Substitutionsmuster.

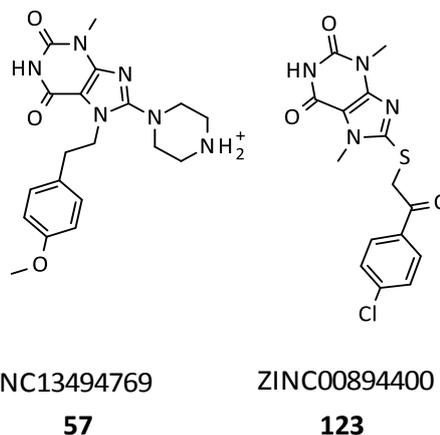


Abbildung 5.36 – Erworbene Substanzen mit einem Xanthin-Körper als Grundkörper.

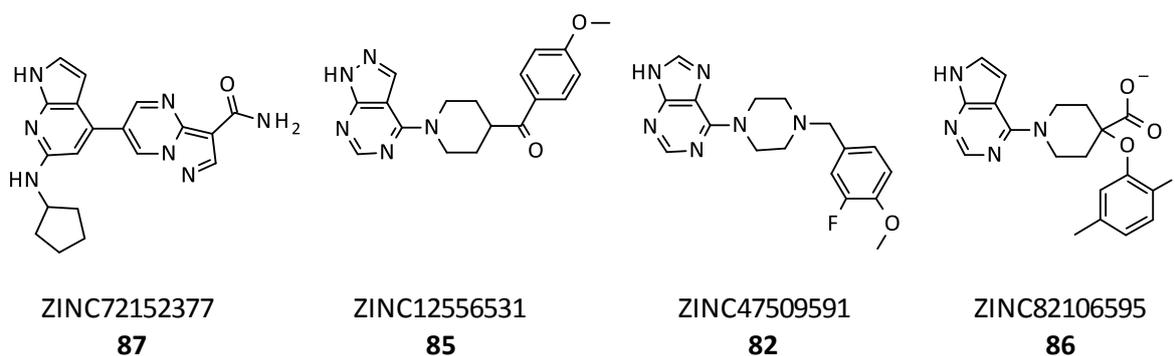


Abbildung 5.37 – Erworbene Substanzen mit einem Azaindol-Körper als Grundkörper.

Interessante Vertreter mit einem Azaindol-Körper, sind in Abbildung 5.37 dargestellt. Diese Stoffklasse kam vor allem bei dem Docking in 1E11 vor und zeigte verschiedene Interaktionsmuster (z. B. **86** mit einem Carboxylat). Es wurden sowohl Substanzen mit einer tiefen Bindung in der P-Tasche aus als auch Substanzen, die den A-Bereich mit adressierten, gewählt.

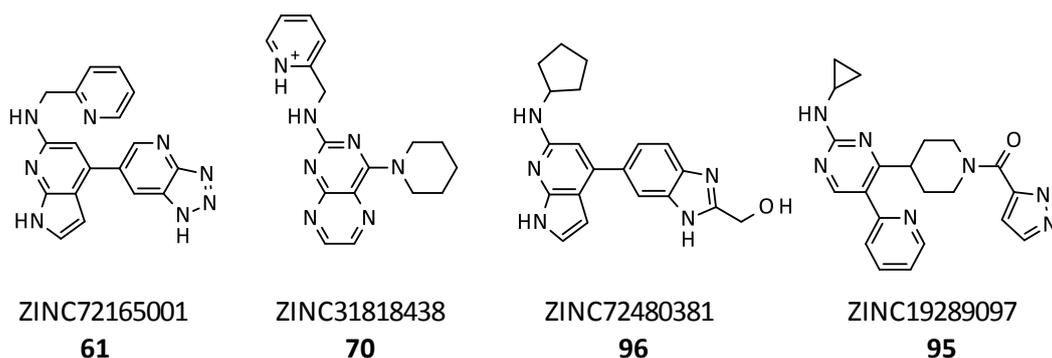


Abbildung 5.38 – Erworbene Substanzen mit einem Aminopyridin-Körper als Grundkörper.

Eine aufschlussreiche Gruppe stellten die, im weitesten Sinne, 2-Aminopyridin-Derivate dar. Obwohl die Interaktion mit dem Hauptpharmakophor ähnlich sind und die ersten drei Substanzen ein ähnliches Substitutionsmuster besaßen, wurden diese Substanzen in verschiedenen Dockings gefunden (3G7E_1, 3G7E_2 und 1E11; Abbildung 5.38 v. l. n. r.). Bei genauerer Betrachtung zeigten diese Strukturen dennoch die jeweils typischen Charakteristika der Bindetasche: **61** besaß einen Pyridin-Ring, der tief in der E-Tasche band (Merkmal von 3G7E_1), **70** hatte einen protonierten Pyridinium-Ring, der mit Asp73 interagierte (3G7E_2), und **96** zeigte die typische tiefe Bindung eines hydrophilen Restes in der P-Tasche (1E11). **95** (ebenfalls vom Docking in 1E11) unterschied sich von den andern untersuchten Derivaten. Der Cyclopropyl-Rest in der E-Tasche und der Pyridyl-Rest im A-Bereich waren vielversprechend.

Als Strukturen die keinem anderen Muster entsprachen wurden **101** und **77** ausgewählt (s. Abbildung 5.39). Die erstgenannte Struktur wurde in 1E11 gedockt und wies einen neuen Grundkörper auf. **77**, welches erfolgreich in 3G7E_2 gedockt wurde, war aufgrund seiner Protonierung ein gutes Beispiel, um die Vorhersage von dem cRAISE-Docking zu überprüfen.

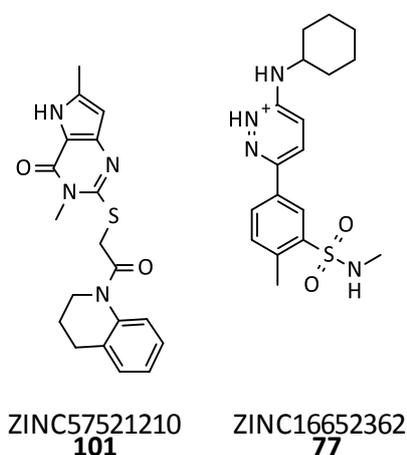


Abbildung 5.39 – Erworbene Substanzen mit verschiedenen Grundkörpern.

5.6.6 Zweite Auswahl

Aufgrund der ersten biologischen Ergebnisse (s. Abschnitt 7.1 auf S. 183) wurde eine zweite Auswahl von Substanzen vorgenommen, die anschließend erworben und getestet werden sollte. Ausgehend von der ersten Auswahl und den dort verwendeten Hypothesen, ergaben sich folgende Denkansätze:

1. Die in 3G7E vorhandene tiefe E-Tasche stellt ein zentrales Charakteristikum dieser Kristallstruktur dar. Die meisten Dockingergebnisse in den beiden verwendeten Screening-Ansätzen (3G7E_1 und 3G7E_2) wiesen eine ausgeprägte Bindung in diesem Bereich auf. Möglicherweise stellte diese tiefe Tasche nicht nur eine Besonderheit der Kristallstruktur von *E. coli* dar, sondern eine besondere Anpassung für diesen Liganden selbst. Dieser *Induced Fit* erforderte dann eine Ethinyl-Struktur oder eine Struktur mit ähnlichen Eigenschaften. Andere Reste, insbesondere raumfordernde Ringsysteme wie Pyridin, können diese Konformation des Proteins nur unzureichend stabilisieren und sollten daher vermieden werden.
2. In 1EI1 gibt es ebenfalls einen stark angepassten Bereich der Bindetasche: die P-Tasche. Diese Seitentasche wird in der Kristallstruktur von einem extrem hydrophilen, mehrfach negativ geladenen Rest belegt. Dies führt zu großen strukturellen Anpassungen innerhalb dieses Bereichs (s. Abschnitt 2.1 auf S. 31). Die im Screening verwendeten und auch schließlich erworbenen Substanzen waren demgegenüber deutlich hydrophober. Möglicherweise tritt dann ein ähnliches Problem, wie bei 3G7E auf: Die Substanzen sind bei einem einfachen Docking nicht in der Lage, das entsprechende *Induced Fit* hervorzurufen.
3. Eine generelle Tendenz bei Docking-Tools besteht darin, die gesamte Bindetasche nach dem Schlüssel-Schloss-Prinzip [231] möglichst vollständig auszufüllen. Diese ist darauf zurückzuführen, dass viele Docking-Prozesse oder Scoring-Funktionen darauf ausgelegt sind, additiv enthalpische Interaktionen zu bewerten, aber entropische Effekte häufig unterbewertet werden (s. Abschnitt 5 auf S. 100). Dieser Zusammenhang wurde bereits von KLEBE diskutiert [120]. Wie schon anhand von ATP und Novobiocin deutlich wird (s. Abschnitt 2.1 auf S. 27), ist es jedoch nicht nötig, die gesamte Bindetasche auszufüllen, um nanomolare Bindungskonstanten zu erreichen. Dennoch traten Posen von Liganden, die die Bindetasche nicht vollständig ausfüllen, während dieser Docking-Experimente nur selten auf. Dies bedeutet zwar nicht, dass die erfolgreich gedockten Liganden falsch sind, es fehlt jedoch ein Teil der Ergebnisse. Diese Tatsache verstärkt weiterhin die zuvor beschriebenen Probleme mit den sehr spezifisch ausgebildeten Seitentaschen erheblich. Dies wird insbesondere in der Bindetasche von 1EI1 deutlich, da im P-Bereich sehr viele gut bewertbare Wasserstoffbrückenbindungen möglich sind. Sehr unterschiedliche Strukturen können daher in diesem Bereich viele Posen generieren und werden auch gut bewertet, weil sie ein oder zwei gute hydrophile Kontakte

besitzen. Diese Interaktionen würden diesen Bereich aber noch lange nicht so gut stabilisieren wie der vierfach negativ geladene Phosphatrest.

Die Hauptproblematik bestand damit in dem strukturbasierten Vorgehen und der hohen Flexibilität und Anpassungsfähigkeit der Bindetasche. Enzymflexibilität stellt, wie bereits beschrieben (s. Abschnitt 4 auf S. 79), eine wichtige, in den letzten Jahren immer mehr im Fokus geratene, Eigenschaft für das Docking dar. Für den in dieser Arbeit verwendeten Docking-Vorgang wurden starre Zielstrukturen verwendet, die einen möglichst großen konformativen Raum abdecken sollten. Trotzdem ist die Enzymstruktur in der Realität jedoch nicht starr, sondern befindet sich in einer stetigen Bewegung und Anpassung an mögliche Liganden. Jede Struktur sollte vielmehr, ähnlich wie bei der Mesomerie, als konkreter Zustand in einem fließenden Gleichgewicht verstanden werden. Für die hier erarbeiteten Problemstellungen könnte, rein hypothetisch, eine Kombination der verschiedenen Enzymstrukturen mit einem Phosphatbereich wie bei 3G7E und einem zentralen Bereich wie bei 1E11 vorteilhaft sein. Andererseits ginge bei einem solchen Vorgehen auch charakteristische Merkmale für die selektive Hemmung der Gyrase verloren.

Ein anderer Lösungsansatz für die oben beschriebenen Probleme besteht darin, mithilfe anderer Bewertungskriterien die Dockingpose neu zu analysieren. Bisher wurden in dieser Arbeit hauptsächlich strukturbasierte Methoden verwendet, um neue Liganden zu finden und zu bewerten. Mithilfe literaturbekannter Inhibitoren wäre es jedoch auch möglich, die Posen ligandenbasiert zu untersuchen. Der Vorteil von dieser Vorgehensweise (bspw. QSAR, ligandenbasiertes Screening) besteht darin, dass sie nicht von einem einzigen Protein-Ligand-Komplex dominiert werden wie beim strukturbasierten Screening. Vielmehr gehen Informationen von bekannten Komplexen, bzw. der Posen der Liganden, mit in die Bewertung ein. Dadurch werden insbesondere selten vorkommende Konformationen abgeschwächt. Dabei besteht auch bei diesen Methoden eine große Abhängigkeit vom Datensatz: z. B. Anzahl der Substanzen, chemische Diversität, Bereich der biologischen Aktivität [232–234]. Diese Eigenschaften können für unterschiedliche Ansätze sehr verschieden sein. Dennoch stellten die ligandenbasierten Methoden die optimale Ergänzung zu den hier erarbeiteten Problemstellungen dar: Die möglichen Extrema in den bekannten Strukturen werden nur abgeschwächt bewertet, während allgemein zur Aktivität beitragende Substrukturen verstärkt in die Bewertung eingehen.

Aus diesen Gründen wurde als ein zusätzliches Bewertungskriterium 3D-QSAR-Modelle verwendet. Dafür wurden zwei verschiedene 3D-QSAR-Modelle^{xiii} entwickelt: ein auf K_D - und ein auf IC_{50} -Werten basierendes Modell (s. Tabelle 5.4). Diese beiden Modelle wurden schließlich

Tabelle 5.4 – Statistische Kenndaten der verwendeten 3D-QSAR-Modelle^{xiii}

	K_D -Model	IC_{50} -Model
Substanzen	56	87
Q²	0.920	0.644
R²	0.997	0.840
Komponenten	7	3

genutzt, um die Glide-Redocking-Ergebnisse der Clusterleader aus Abschnitt 5.5 auf S. 120 neu zu bewerten. Anhand der Priorisierung durch die beiden unterschiedlichen Modelle und mit einer abschließenden visuellen Bewertung der Posen, bei der die Ergebnisse der ersten Testung mit berücksichtigt wurden, fand eine erneute Auswahl von Substanzen (s. u.) statt. Prinzipiell handelte es sich um die Analyse der gleichen Docking-Ergebnisse, wie bei der ersten visuellen Auswahl im Abschnitt 5.6 auf S. 123. Daher sollen im Folgenden nur die am Ende erworbenen Strukturen gezeigt und beurteilt werden. Neben dieser Priorisierung wurden von jedem Docking mit cRAISE (s. Abschnitt 4.2.3 auf S. 98) nochmal die fünf besten Posen visuell verglichen und vier weitere Substanzen zusätzlich ausgewählt.

Da sich die meisten Posen nicht mehr von den zuvor untersuchten und dargestellten unterschieden, wird im Folgenden nur auf die wichtigsten Unterschiede zur ersten Auswahl (s. Abschnitt 5.6.5 auf S. 141) eingegangen. Die ausgewählten Substanzen (s. Abbildung 5.40) enthielten zumeist nur kleine hydrophobe Reste, die in die W- oder E-Tasche zeigten, bspw. ein Cyclopropyl-Rest bei **124**, ein Isopropyl-Rest bei **129** oder ein Methyl-Rest bei **134**. Teilweise besetzten die Substanzen diesen Bereich auch gar nicht mehr: z. B. **133**, **140** und **143**. Es gab jedoch auch noch einige Ausnahmen, wie bei den Substanzen **138**, **141** und **146**. Dies waren häufig Substanzen, die aus der Auswahl aus den TOP5-cRAISE-Ergebnissen hervorgingen.

Da der W- und E-Bereich nicht mehr häufig besetzt waren, traten auch seltener Interaktionen mit dem dort vorhandenen katalytisch wichtigen Asn46 auf. Ausnahmen bildeten Substanzen wie **126**, **139** oder **141**. In der P-Tasche gab es eine ähnliche Tendenz: Die meisten Strukturen füllten diesen Bereich deutlich weniger aus als die Substanzen der ersten Auswahl. Zumeist befanden sich dort kleine aromatische Reste (z. B. **133**, **139** oder **143**) oder aliphatische Gruppen (**135** oder **146**). Es traten keine raumfüllenden Reste mehr in diesem Bereich auf. Im A-Bereich der Bindetasche änderte sich bei der Auswahl wenig von dem bisher gewählten Ansatz. Wurde dieser besetzt, handelte es sich meistens um kurze aromatische Reste, wie bei **128**, **134**, **137** oder **146**.

^{xiii} Die Modelle wurden von Pia Schieback aus der Arbeitsgruppe von Thomas Lemcke vom Institut für Pharmazie der Universität Hamburg erstellt und auf den Datensatz angewendet.

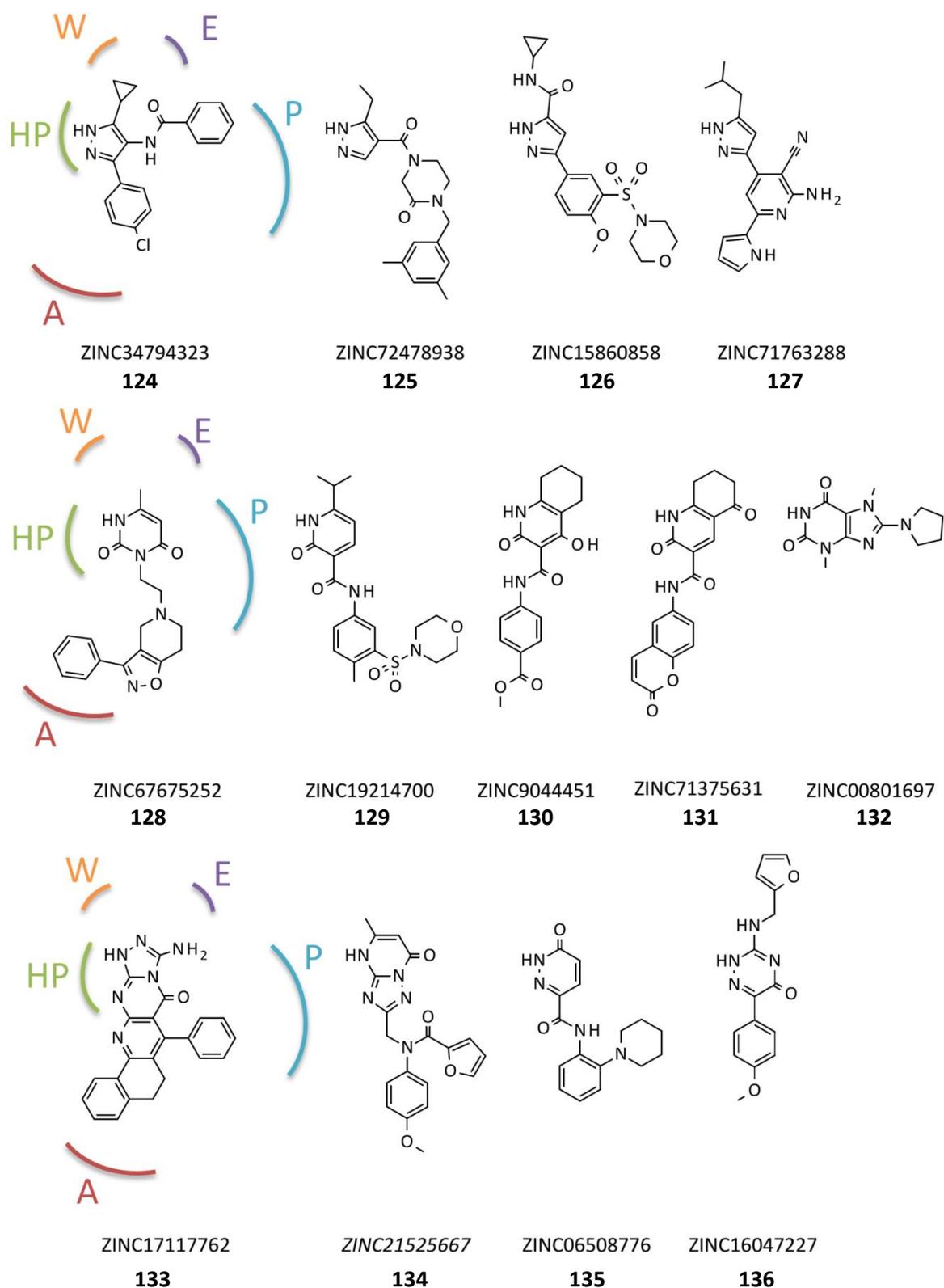


Abbildung 5.40 – Erworbene Substanzen für die zweite biologische Testreihe. Die Substanzen sind entsprechend ihrer Docking-Pose ausgerichtet. Zur Orientierung sind die Bindungsbereiche der Bindetasche dargestellt. Kursiv geschriebene Substanzen sind aufgrund ihres cRAISE-Rankings ausgewählt worden.

Fortsetzung nächste Seite

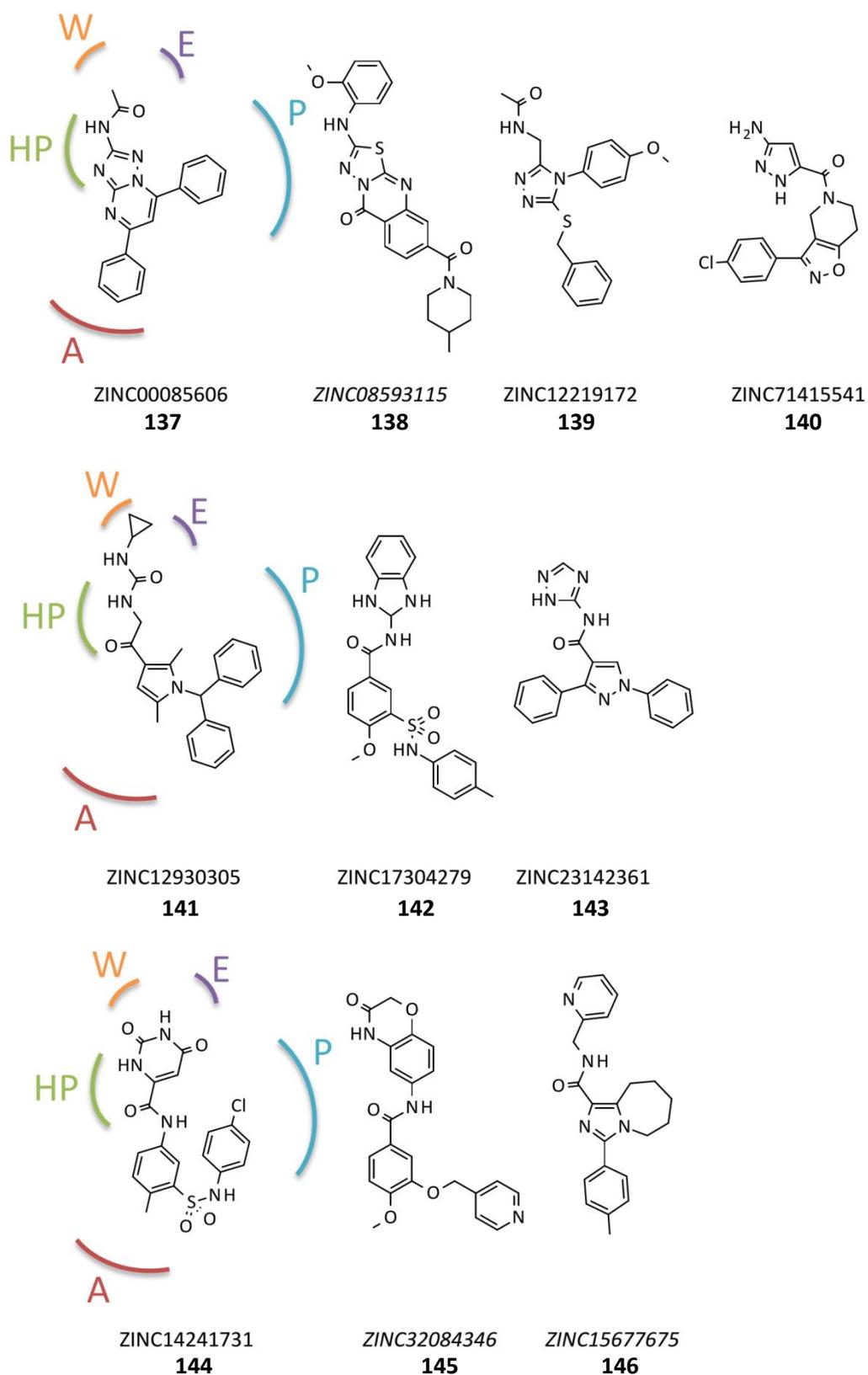


Abbildung 5.40 – Fortsetzung.

6 Assay-Entwicklung

Die Bestimmung der Aktivität von Inhibitoren der Gyrase kann mit zahlreichen direkten und indirekten Methoden erfolgen. Voraussetzung für die Durchführung jedes Assays ist das Verständnis über die Grundlagen der enzymatischen Reaktion. Bei der durch die Gyrase katalysierten Reaktion liegt eine Besonderheit darin, dass es sowohl zwei Edukte (ATP und DNA) als auch Produkte unterschiedlicher Umsetzungsart gibt: ATP wird in einer chemischen Reaktion zu ADP und anorganischem Phosphat (P_i) gespalten, während bei der DNA die Topologie verändert wird (s. Abschnitt 1.1 auf S. 5). Seit der ersten Isolierung der Gyrase wurden zahlreiche Assays entwickelt, die jeweils Produkte (ADP, P_i oder topologisch veränderte DNA) zur Auswertung benutzen oder mithilfe indirekter Methoden die Aktivität des Enzyms bestimmen (z. B. Verdrängung eines fluoreszenzmarkierten Inhibitors). Während ältere Methoden entweder hohen Aufwand erfordern oder keine hohe Präzision besitzen, sind neuere Assay-Methoden apparativ oder aufgrund der Materialien anspruchsvoller aber meist präziser.

Bei dem Assay liegt der Fokus darauf, schnell, einfach und sicher Substanzen zu identifizieren, die das zu untersuchende Target inhibieren. Da Assay-Methoden häufig auf der Messung physikalischer Parameter basieren, ist es in der Regel nicht sicher möglich, das Target des möglichen Inhibitors mit einer einzigen Messmethode zu verifizieren. Um falsch-positive Hits, z. B. durch Eigenfluoreszenz oder Inhibition von anderen beteiligten Enzymen, zu eliminieren, werden deshalb bei Screenings orthogonale Assays durchgeführt. Orthogonale Assays basieren auf einer unterschiedlichen Methode und sollten, z. B. bei falsch positiven Hits, nicht auf die gleiche Weise beeinflusst werden. Bei der Gyrase bietet es sich an, einen auf ADP oder P_i und orthogonal dazu einen auf der Topologie der DNA basierenden Assay zu verwenden.

Im Folgenden wird ein kurzer Überblick über die bestehenden Assay-Methoden für die Aktivitätsbestimmung der Gyrase gegeben und deren Vor- und Nachteile diskutiert.

6.1 Assays

- Bestimmung der Aktivität über Phosphat oder ATP/ADP

Die einfachsten Assays befassen sich mit der Bestimmung des bei der enzymatischen Reaktion freigesetzten Phosphats. Phosphat ist als Abspaltungsprodukt ein ubiquitär vorkommendes Molekül in biologischen Systemen. Es tritt bei vielen Enzymreaktionen auf und daher gibt es auch viele etablierte Methoden, die sich mit dessen Bestimmung befassen. So ist es nicht überraschend, dass die ersten beschriebenen Aktivitätsmessungen, die je mit aufgereinigter Gyrase durchgeführt wurden, mithilfe einer Phosphatbestimmung über Autoradiographie durchgeführt wurden [235]. Neben dieser Me-

thode hat sich vor allem die Absorptionsmessung von Phosphomolybdat-Malachitgrün-Komplexen etabliert. Diese sind, neben ihrer besseren Handhabbarkeit, wirtschaftlicher und unkomplizierter durchzuführen [236].

Bei der Autoradiographie wird [γ - ^{32}P]-ATP verwendet. Dieses setzt bei der Hydrolyse das radioaktiv markierte endständige Phosphat frei, das nach Aufreinigung mittels einer Dünnschichtchromatografie auf einer Polyethylenimincellulose-Platte bestimmt werden kann [235]. Der Vorteil dieser Methode ist die relativ direkte und sichere Bestimmung des Produktes. Demgegenüber steht die teure und vergleichsweise risikoreiche Handhabung von radioaktivem Material. Aufgrund der relativ direkten Messung eignet sich die Methode, um mechanistische Aufklärungen, wie bspw. Produktspezifität, zu untersuchen. Die Handhabung und Entsorgung von radioaktiven Abfällen erschwert jedoch eine routinemäßige, mit großem Durchsatz verbundene Testung.

Die Bestimmung mithilfe des Phosphomolybdat-Farbstoff-Komplexes von freiem Phosphat wurde erstmals 1947 von SOYENKOFF *et al.* beschrieben [237]. Hierbei wird Phosphat mithilfe von Molybdensäure nachgewiesen. Molybdat bildet mit Phosphat anionische Komplexe unterschiedlicher Zusammensetzung (s. Formel 6.1).



Formel 6.1 – Reaktionsgleichung für den Phosphomolybdatokomplex. Die Zusammensetzung des polyanionischen Komplex kann sehr stark variieren [238]. Durch einen großen Überschuss an Molybdat und einem niedrigen pH-Wert im Assay wird das Gleichgewicht möglichst weit zum Produkt verschoben.

Dieser qualitative Nachweis hätte für die geringen freigesetzten Mengen von Phosphat bei Enzymreaktionen eine zu geringe Sensitivität. Durch Zugabe eines kationischen Farbstoffes (z. B. Chinaldinrot oder Malachitgrün) bildet sich ein weiterer Komplex mit dem bereits vorhandenen, wodurch die Sensitivität signifikant erhöht wird. Die Wahl des Farbstoffes richtet sich dabei nach der Löslichkeit des Farbstoffes, der Komplexe und ihrer Stabilitäten. Dabei haben sich insbesondere Chinaldinrot und Malachitgrün bewährt. Über die photometrische Messung der Komplexe kann auf die freigesetzte Menge an Phosphat zurückgeschlossen werden [236, 239].

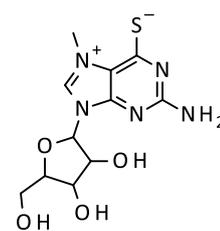
Ein Nachteil dieser Methode besteht darin, dass die Aktivität nur indirekt gemessen wird. Damit steigt die Fehleranfälligkeit, so können z. B. hydrophobe Amine mit dem Phosphomolybdat-Komplex interagieren und niedrigere Phosphatkonzentrationen vortäuschen [240] oder die getesteten Substanzen können im Messbereich eine starke Eigenabsorption aufweisen, wodurch fälschlicherweise eine hohe Phosphatkonzentration suggeriert wird. Problematisch ist auch die Sensitivität

gegenüber Phosphat: Verunreinigungen täuschen eine höhere Enzymaktivität vor. Phosphat ist gerade in biologischen Systemen ein weit verbreitetes Anion (z. B. bei der biotechnologischen Synthese von Novobiocin und seiner Derivate [241]). Demgegenüber steht ein einfach durchführbares System, welches robust, schnell und unkompliziert zu handhaben ist.

Bei dem enzymgekoppelten Assay mit der Pyruvatkinase und Lactatdehydrogenase wird das freigesetzte ADP aus der enzymatischen Reaktion der Gyrase bestimmt. Das ADP wird dabei von der Pyruvatkinase mit Phosphoenolpyruvat zu Pyruvat und ATP umgesetzt. Die Lactatdehydrogenase reduziert im nächsten Schritt, unter gleichzeitiger Oxidation von NADH zu NAD⁺, Pyruvat zu Lactat (s. Abbildung 6.1). Die Abnahme von NADH kann photometrisch gemessen werden [242]. Die Vorteile dieser Methode liegen darin, dass ATP durch die Rückreaktion unbegrenzt zur Verfügung steht. Weiterhin kann direkt die Kinetik der Freisetzung bestimmt

werden, weil die Umsetzung, also die Abnahme von NADH, fortlaufend mitverfolgt wird (Bei den bisher vorgestellten Assays wird die enzymatische Reaktion durch die Messung gestoppt). Die wichtigsten Nachteile ergeben sich durch die Verwendung von zwei weiteren Enzymsystemen. Das Gesamtsystem wird empfindlicher und anfälliger für Fehler. Weiterhin sind die Kosten dadurch auch deutlich höher.

Ein anderer enzymgekoppelter Assay basiert auf der Purin-Nukleosid-Phosphorylase. Das Enzym setzt 7-Methyl-8-thioguanosin (**147**) (s. Abbildung 6.2) in Anwesenheit von Phosphat zu einem stark lichtabsorbierenden Produkt um [243]. Das System hat eine höhere Sensitivität als der zuvor beschriebene Assay, in den Grundzügen jedoch die gleichen Vor- und Nachteile.

**147****Abbildung 6.2** – 7-Methyl-8-thioguanosin.

– Bestimmung der Aktivität über die Topologie der DNA

Die Topologie von DNA wird üblicherweise mit Hilfe der Gelelektrophorese analysiert. GELLERT *et al.* bewiesen 1976 bei der Entdeckung der Gyrase deren Aktivität ebenfalls anhand von Supercoiling-Assays [244]. Erst 2006 wurde von der Arbeitsgruppe um MAXWELL ein weiterer Assay auf einer DNA-Triplex-basierenden Methode entwickelt, die die Anwendung in einem High-Throughput Screening ermöglicht [245].

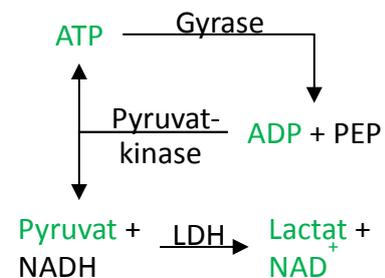


Abbildung 6.1 – Schematische Darstellung des Pyruvatkinase-Lactatdehydrogenase gekoppelten Assays. Es sind nur die für die Messung relevanten Edukte (schwarz) und Produkte (grün) dargestellt. ATP ist sowohl Edukt des Assays als auch Produkt der Pyruvatkinase. PEP: Phosphoenolpyruvat; LDH: Lactatdehydrogenase

Beim Supercoiling-Assay wird ein relaxiertes Plasmid mit der Gyrase umgesetzt. Dabei wird das Plasmid bei einer aktiven Gyrase in seine supercoiled Form überführt. Wird das Enzym gehemmt, findet keine Umsetzung statt und das Plasmid bleibt relaxiert. Im nächsten Schritt findet eine Agarosegelelektrophorese statt. Bei der Gelelektrophorese werden große Biomoleküle (z. B. DNA, RNA oder Proteine) durch ein Polymer aufgrund ihrer Größe in einem elektrischem Feld getrennt. Die hier als Polymer verwendete Agarose ist ein Polysaccharid, das aus glykosidisch verknüpfter D-Galaktose und 3,6-Anhydro-L-Galaktose zusammengesetzt ist. Die durch die zahlreichen Phosphatgruppen negativ geladene DNA wandert im elektrischen Feld zur Anode und wird durch die Poren des Polymers in Abhängigkeit ihrer Größe zurückgehalten [246, 247]. Das Volumen des supercoiled Plasmids ist deutlich geringer als das des relaxierten Plasmids. Daher wandert das supercoiled Plasmid schneller im elektrischen Feld als das relaxierte und die beiden sequenzgleichen aber topologisch unterschiedlichen Plasmidformen trennen sich auf. Zur Auswertung wird das Gel mit dem DNA-interkalierenden Farbstoff Ethidiumbromid angefärbt.

Ähnlich wie bei der Autoradiographie findet eine sehr direkte Messung des Produktes statt. Weiterhin handelt es sich um eine gut standardisierte Methode. Die Handhabung ist für einzelne Messungen praktikabel; größere Durchsätze können mit diesem Assay jedoch nicht erreicht werden. Der größte Nachteil besteht darin, dass eine IC_{50} -Bestimmung zwar möglich, aber aufwendiger und ungenauer als bei anderen Methoden ist [64]. Für eine rein qualitative Aussage, ob die Gyrase gehemmt wird, ist dieser Assay jedoch aufgrund seiner hohen Spezifität die Methode der Wahl.

Bei der auf der DNA-Triplex-basierenden Methode wird der Umstand ausgenutzt, dass in Abhängigkeit von der Verdrillung der DNA DNA-Triplexe ausgebildet werden. DNA-Triplexe stellen eine besondere Struktur der DNA dar. Die häufigsten Strukturen der DNA bestehen aus zwei DNA-Strängen (z. B. B-, A- und Z-Form). Unter besonderen Umständen kann sich noch ein dritter Strang anlagern, der meist aus Polypyrimidin- oder Polyuridinbasen besteht, um eine dreisträngige DNA (Triplex-DNA, H-Form) zu bilden [248]. In der Natur übernehmen diese Strukturen Aufgaben in der Genregulation und -expression [249]. Für die Aktivitätsbestimmung wird ein Oligonucleotid auf einer Mikrotiterplatte immobilisiert, das bevorzugt Triplexe mit einem supercoiled Plasmid formt. In den Kavitäten der Mikrotiterplatte findet dann die Inkubation der Gyrase mit einem relaxierten Plasmid statt. Wird die Gyrase nicht gehemmt, setzt sie das Plasmid in seine supercoiled Form um. Das supercoiled Plasmid bildet dann ein stabiles Triplex mit dem Oligonucleotid und kann, nach einem Waschschrift, bspw. mit einem Fluoreszenzfarbstoff angefärbt werden. Diese Methode erlaubt mithilfe eines sequentiellen Verfahrens, das auch gut automatisiert durchgeführt werden kann, eine schnelle und einfache Bestimmung. Es handelt sich aber ebenfalls um eine indirekte Methode, die nicht geeignet ist, die Kinetik zu bestimmen [245, 250, 251].

– Andere Methoden zur Bestimmung der Aktivität

Bei der Fluoreszenzpolarisation wird die Aktivität mithilfe eines fluoreszenzmarkierten Novobiocinderivats („*novobiocin-texas red probe*“ (**148**)) bestimmt (s. Abbildung 6.3). Der Ligand wird mit linear polarisiertem Licht angeregt. Das emittierte Licht des Fluorophors ist deswegen auch prinzipiell linear polarisiert. Ist der Ligand klein (wie bei **148**) und kann sich frei bewegen und drehen, ist das emittierte Licht trotzdem unpolarisiert, weil die Moleküle zufällig rotieren. Bei Bindung an das Enzym wird die Bewegung stark eingeschränkt, so dass ein größerer Teil des fluoreszierenden Lichts weiter linear polarisiert ist. Bei einem Screening verdrängt ein Inhibitor den fluoreszierenden Farbstoff und die Fluoreszenzpolarisation nimmt ab [252]. Diese Methode hat große Vorteile gegenüber den anderen Methoden, da die Verdrängung des Novobiocins aus der Bindetasche gemessen wird. Damit ist der Bindungsort viel klarer definiert und gleichzeitig können die Assoziations- und Dissoziationskinetiken bestimmt werden. Als größte Nachteile erweisen sich die Kosten für den Liganden und der große apparative Aufwand für das Verfahren.

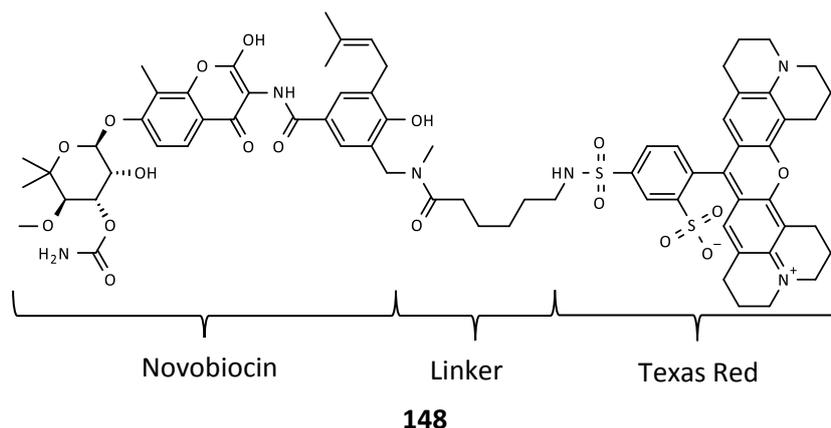


Abbildung 6.3 – Fluoreszenzmarkiertes Novobiocin. Am Novobiocin wird dafür durch eine Mannichreaktion ein sekundäres Amin eingeführt. Dieses wird dann über einen Linker mit dem fluoreszierenden Farbstoff Texas Red verknüpft.

Eine weitere Möglichkeit für die Bestimmung der Hemmung der Gyrase basiert auf der Genexpression. Dieser Assay nutzt im Gegensatz zu den bisher beschriebenen Assays ganze Bakterien und zählt somit zu den *in vivo*-Verfahren. Dabei besitzen *E. coli*-Zellen ein Plasmid, das ein Luciferase-Reportergen unter Kontrolle des *pgyA*-Promotors enthält. Wird die Gyrase gehemmt, versucht die Bakterienzelle die fehlende Aktivität der Gyrase, durch eine verstärkte Expression der Gyrase auszugleichen. Daher wird vermehrt Luciferase gebildet, da diese von dem Gyrase-Promoter reguliert wird. Antibiotika mit anderem Wirkmechanismus haben keinen Einfluss auf die Expression des Reportergens. Die Luciferase ermöglicht damit eine Lumineszenzmessung, die auf die Gyrasehemmung zurückzuführen ist [253]. Ein Screening bei *in vivo*-Experimenten ist ambivalent zu bewerten. Einerseits erhält man durch das spezifische Reportergen genaue Informationen über den Einfluss auf die Gyr-

se, andererseits ist die Fehleranfälligkeit und Interaktionsmöglichkeit bei lebenden Bakterien im Vergleich zu einem *in vitro* Assay sehr hoch. Dadurch werden hohe Anforderungen an den Inhibitor gestellt, da er in die Bakterienzelle hineingelangen muss, nicht wieder ausgeschleust und modifiziert werden darf. Dies ist andererseits auch ein Argument für dieses Verfahren, weil ein Hit schon viele Voraussetzungen einer Leitstruktur erfüllt, die später hilfreich für die Weiterentwicklung sind.

Es gibt eine Vielzahl an bereits bekannten Verfahren für die Bestimmung von Inhibitoren der Gyrase und deren ATPase-Untereinheit. Probleme verursachen bei vielen Assays die zusätzlichen Kosten durch Reaktanden, Materialien und Geräte. Für die Auswahl von zwei orthogonalen Testsystemen war im Rahmen dieser Arbeit die Verlässlichkeit, Handhabbarkeit und Durchführbarkeit von großer Bedeutung. Der Phosphomolybdat-Malachitgrün-Komplex (s. S. 152) und der Supercoiling-Assay (s. S. 154) erfüllten die gewünschten Voraussetzungen und dienten daher als Ausgangspunkte für weitere Untersuchungen.

6.2 Material und Methoden

6.2.1 Supercoiling Assay

Die Inkubation wurde nach den Vorgaben des Herstellers (New England Labs) mit 500 ng relaxiertem DNA-Plasmid ausgeführt (s. Experimenteller Teil 11.2.5 auf S. 218).

6.2.2 Absorption des Phosphomolybdat-Farbstoff-Komplexes

Dieser Assay verwendet die Absorption von Licht eines Phosphomolybdat-Farbstoff-Komplexes. Obwohl diese Messung schon lange bekannt ist, wurde sie, unseres Wissens nach, noch nicht bei der Gesamtgyrase angewendet. Daher mussten die Bedingungen angepasst werden. Die verwendete Methode basierte auf den Bedingungen von ANDERLE *et al.* (verwendeten die isolierte GyrB-Untereinheit) [63] und ROWLANDS *et al.* (High-Throughput-Screening von HSP90) [254].

Die endgültig optimierte Methode ist im Experimentellen Teil aufgeführt (s. 11.2.6 auf S. 218).

6.2.3 Fluoreszenzauslöschung des Phosphomolybdat-Farbstoff-Komplexes

ZUCK *et al.* konnten 2005 die Empfindlichkeit des Phosphomolybdat-Farbstoff-Komplex Assay bei der Alkalischen Phosphatase deutlich verbessern, indem sie die Fluoreszenzeigenschaften von weißen Platten benutzten, um die Reaktion zu miniaturisieren [255]. Diese Methode konnte auch erfolgreich beim HSP90 verwendet werden [256].

Die Methode basiert darauf, dass weiße Mikrotiterplatten fluoreszieren. Dies ist keine beabsichtigte Eigenschaft der Platte, daher kann die Fluoreszenz bei unterschiedlichen Herstellern und auch bei verschiedenen Mikrotiterplatten-Chargen variieren. Die Messung verläuft analog zu einer normalen Fluoreszenzmessung bei einer Mikrotiterplatte. Der zur Anregung verwendete Lichtstrahl verursacht allerdings keine Fluoreszenz in der Lösung. Diese entsteht erst, wenn der Lichtstrahl auf den Boden der Kavität trifft. Anschließend wandert das fluoreszierte Licht wieder durch die Lösung zum Detektor (s. Abbildung 6.4). Bei diesem Prozess kann die Fluoreszenz dadurch abgeschwächt werden, dass das zur Anregung verwendete Licht oder das fluoreszierte Licht absorbiert wird. Im Vergleich zu einer einfachen Absorptionsmethode ist die Methode wesentlich flexibler: Bei einer Absorptionsmessung funktioniert die Messung nur, wenn beim Absorptionsmaximum gemessen wird; bei der Fluoreszenzauslöschung, kann das anregende Licht oder das fluoreszierte Licht absorbiert und daher für die Bestimmung verwendet werden. Zusätzlich können die Messbedingungen durch die Wahl der Platten (s. u.) noch weiter angepasst werden (z. B. Wellenlänge des anregenden Lichts). Schließlich konnten

sowohl ZUCK *et al.* als auch MIYATA *et al.* eine deutlich höhere Sensitivität für den Phosphatnachweis zeigen.

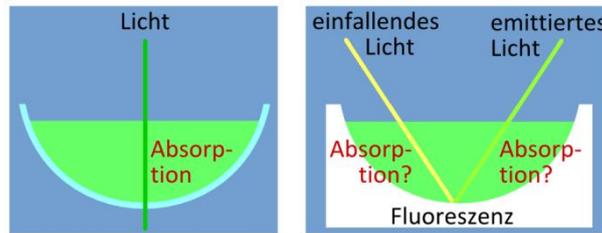


Abbildung 6.4 – Vergleich der Fluoreszenzauslöschungs- mit der Absorptionsmethode. Bei der normalen Absorptionsmethode (**links**), wird der Lichtstrahl durch die farbige Lösung geleitet. Dabei wird eine transparente Mikrotiterplatte verwendet, die den Lichtstrahl anschließend ungehindert passieren lässt. Die Abnahme der Lichtintensität wird durch die Absorption von der Lösung verursacht. Bei der Fluoreszenzmethode (**rechts**) durchquert der einfallende Lichtstrahl die Lösung und führt zur Fluoreszenz. Das emittierte Licht passiert erneut die Lösung und wird vom Detektor gemessen. Bei diesem Vorgang kann die Lichtintensität des einfallenden oder des emittierten Lichtstrahls von der Lösung durch Absorption verringert werden. Dies äußert sich in einer verringerten Fluoreszenz. Diesen Vorgang bezeichnet man als Fluoreszenzauslöschung oder Quenching.

Obwohl bei dieser Bestimmung die Fluoreszenz gemessen wird, ist die bestimmte Veränderung in der Energie nur auf die Absorption zurückzuführen (s. Abbildung 6.4), da sich die Fluoreszenzeigenschaften der Platte nicht verändern. Damit gelten für die Auswertungen die bekannten Gesetze für die Absorption. Für die Bestimmung der Konzentration mithilfe der Absorption ist das Lambert-Beersche Gesetz geeignet [257, 258]. Dieses beschreibt für ein definierten Bereich den Zusammenhang zwischen dem Absorptionskoeffizienten ε , der Konzentration c und der Schichtdicke d zum Verhältnis der gemessenen Lichtintensität I_t zur einfallenden Lichtintensität I_0 . Damit lässt sich nach Formel 6.2 die Absorption A berechnen:

$$A = -\lg\left(\frac{I_t}{I_0}\right) = \varepsilon \cdot c \cdot d \quad \text{Formel 6.2}$$

Dabei ist die Lichtintensität die Leistung, die pro Flächeneinheit erbracht wird. Die Leistung beschreibt wiederum die Energie, die pro Zeiteinheit verbraucht wird. Können alle anderen Parameter konstant gehalten werden, ist es daher möglich von der Energie der gemessenen Fluoreszenz (Blindwert E_0 und Probe E_t) auf die Absorption zurückzuschließen. Da dies einige Vereinfachungen einschließt und um Verwechslungen vorzubeugen, wird diese im Folgenden die berechnete Absorption A_{calc} genannt. Diese sollte sich im Idealfall wieder, wie im Lambert-Beerschen Gesetz, proportional zur Konzentration der Lösung verhalten (s. Formel 6.3).

$$A_{calc} = -\lg\left(\frac{E_t}{E_0}\right) \sim c \quad \text{Formel 6.3}$$

6.2.4 Regressionen für IC₅₀-Werte

Bei einem Inhibitionsassay treten typische Dosis-Wirkungsbeziehungen auf (s. Abbildung 6.5). Mathematisch kann dieser Verlauf mit der 4-Parameter-Logarithmische-Funktion (4PL; s. Formel 6.4), besser bekannt als die Hill-Funktion, oder der 5-Parameter-Logarithmische-Funktion (5PL; Formel 6.5) beschrieben werden [259]. Bei asymmetrischen Kurvenverläufen in dem IC₅₀-Bereich ist die 5PL-Funktion der 4PL-Funktion überlegen, insbesondere wenn der Exponent im Nenner (e in Abbildung 6.5) zwischen 0.2 und 0.6 liegt (in Abbildung 6.5 beträgt der Exponent für die blaue Linie 0.6). Demgegenüber steht ein eventuelles Fitting-Problem durch die kleinere Anzahl an Freiheitsgraden und der zusätzlichen exponentiellen Variabel. Dies erschwert die Auswertung bei der Anwesenheit von wenigen oder hauptsächlich in den Asymptoten vorliegenden Datenwerten.

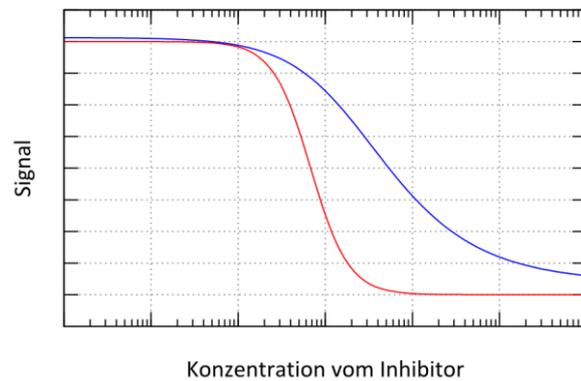


Abbildung 6.5 – Typischer Verlauf von Dosis-Wirkungsbeziehungen (rot symmetrisch, blau asymmetrisch). Der Logarithmus der Konzentration vom Inhibitor wird gegen das Signal aufgetragen. Bei sehr niedrigen Inhibitor-Konzentrationen nähert sich die Kurve asymptotisch dem Signal des ungehemmten Enzyms, im Bereich des IC₅₀-wertes fällt die Kurve stark ab und nähert sich schließlich, wieder asymptotisch, dem Signal des komplett gehemmten Enzyms.

$$f(x) = \min + \frac{\max - \min}{1 + \left(\frac{x}{IC_{50}}\right)^{-Hillslope}} \quad \text{Formel 6.4}$$

$$f(x) = \min + \frac{\max - \min}{\left(1 + \left(\frac{x}{c}\right)^d\right)^e} \quad \text{Formel 6.5}$$

Für die Auswertung wurde R [260] mit dem DRC-Paket („Analysis of dose-response curves“) [261] verwendet.

6.3 Absorptionsmethode

Als Ausgangspunkt für den Assay wurde das Protokoll von ANDERLE *et al.* gewählt [63]. Dieser Assay basierte auf der Bestimmung des freigesetzten Phosphats mithilfe eines Phosphomolybdat-Malachitgrün-Komplexes. Dafür inkubierten sie die GyrB-Untereinheit mit ATP und bestimmten anschließend mit einem Farbreagenz das freigesetzte Phosphat.

Der Reaktionspuffer (im Folgenden Anderle-Puffer genannt) besaß folgende Bestandteile: Tris(hydroxymethyl)-aminomethan (TRIS), Kaliumchlorid, Magnesiumchlorid, Dithiothreitol (DTT), Ethylendiamintetraessigsäure (EDTA) und ATP (s. Tabelle 6.1). Für Anpassungen ist das Verständnis der einzel-

Tabelle 6.1 – Zusammensetzung der Reaktionspuffer [63].

Substanz	Anderle-Puffer	NEB-Puffer
TRIS	50 mM	50 mM
Kaliumchlorid	100 mM	24 mM
Magnesiumchlorid	5 mM	4 mM
DTT	5 mM	2 mM
EDTA	1 mM	-
Glycerol	-	6.5 %
ATP	1 mM	1.75 mM
BSA	-	0.1 mg/mL
Spermidin	-	5 mM
Plasmid (relaxiert)	-	500 ng

nen Bestandteile relevant. TRIS stellt das Puffersystem für die Reaktion dar (pH = 7.6). Bei Kaliumchlorid und Magnesiumchlorid sind beide Kationen der Salze für die Hydrolyse von ATP wichtig. Während dies bei Magnesiumionen schon seit der Erstbeschreibung der Gyrase bekannt war [235], wurde die katalytische Beteiligung von monovalenten Kationen speziell bei ATPasen der GHKL-Klasse erst 2003 beschrieben [262]. Bei der Gyrase gibt es eine Präferenz zu Kaliumionen, das Enzym wird sogar durch höhere Konzentrationen von Natriumionen gehemmt [64]. Erst 2015 fanden HEARNshaw *et al.* heraus, dass es zwei Bindungsstellen für monovalente Kationen gibt, wobei eine von Natriumionen und die andere von Kaliumionen bevorzugt wird [30]. Das Kaliumkation übernimmt dabei möglicherweise eine katalytische Funktion. DTT wird als Antioxidans und EDTA zur Stabilisierung der Reaktionslösung verwendet. ATP ist eines der Substrate für die Reaktion. Da die Arbeitsgruppe um ANDERLE *et al.* mit der isolierten GyrB-Untereinheit arbeitete, die unter diesen Bedingungen schon ATPase-Aktivität zeigt, wird keine DNA benötigt. Aus Kostengründen konnte in dieser Arbeit ausschließlich mit der Gesamtgyrase gearbeitet werden. Daher mussten die Bedingungen des Assays angepasst werden. In Tabelle 6.1 ist auch die Zusammensetzung für den Supercoiling-Assay vom Hersteller (New England Biolabs; im Folgenden NEB-Puffer genannt) angegeben. Die wichtigsten Unterschiede waren die niedrigere Kaliumchloridkonzentration, die Verwendung von Glycerol statt EDTA zur Stabilisierung und der Zusatz von Bovinen Serumalbumin (BSA), Spermidin und relaxiertem Plasmid. Die Zielsetzung bestand zunächst darin, die Absorptionsmethode als Assay mit der Gesamtgyrase zu etablieren.

Dafür wurde die Aktivität der Gyrase im Anderle- und NEB-Puffer verglichen. Dabei wurde im Anderle-Puffer zusätzlich zu den bisherigen Bestandteilen noch ein relaxiertes DNA-Plasmid in der gleichen Menge wie beim NEB-Puffer verwendet. Dies waren zunächst die einzigen Änderungen, um die Experimente nachvollziehbar zu gestalten. Die zugehörige Messung bestand damit aus jeweils einem Ansatz mit dem modifizierten Anderle-Puffer, dem NEB-Puffer, den zugehörigen Blindwerten, die identisch behandelt wurden, aber kein Enzym enthielten, und einer Phosphatstandardreihe als Referenz. Mit einem Teil der Reaktionslösung wurde die Absorption mithilfe des Phosphomolybdat-Malachitgrün-Komplexes bestimmt (Vorgehen analog zu s. Experimenteller Teil 11.2.6 auf S. 218), mit einem weiteren Teil wurde eine Gelelektrophorese (s. Experimenteller Teil 11.2.5 auf S. 218) durchgeführt. Bei der Absorptionsmessung, s. Abbildung 6.6, zeigte keine der Proben Freisetzung von Phosphat: Beide Messungen mit Enzymen zeigten eine niedrigere Absorption als die zugehörigen Blindwerte. Auffällig war, dass der Blindwert vom NEB-Puffer schon einen Absorptionswert von 0.897 besaß, also bei einem Wert von fast 225 μM Phosphat.

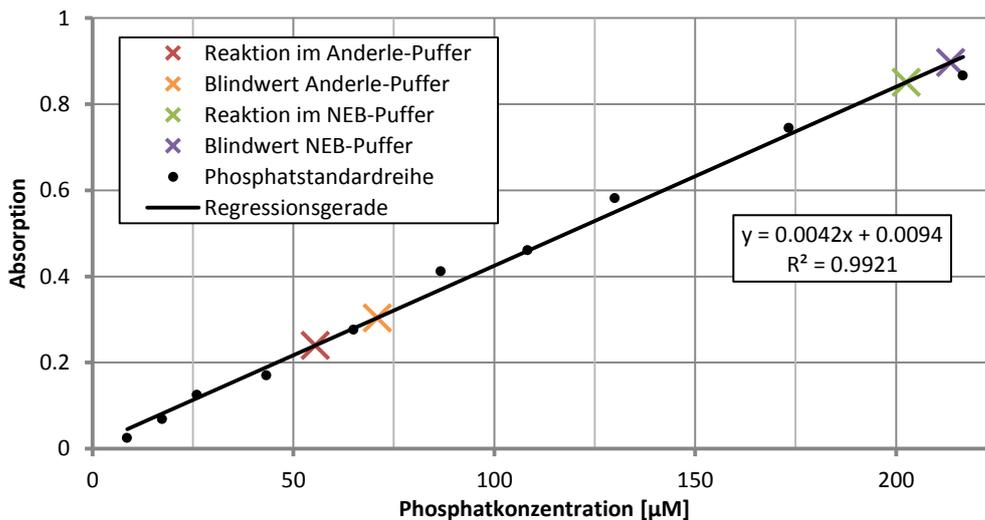


Abbildung 6.6 – Vergleich der Puffer mittels Absorptionsmessung. Bei beiden Reaktionsansätzen war der Messwert kleiner als der zugehörige Blindwert, daher ergaben sich rein rechnerisch negative Phosphatkonzentrationen. Folglich fand keine nachweisbare Freisetzung von Phosphat statt. Der Blindwert vom NEB-Puffer war dreimal so hoch wie der selbst hergestellte Puffer nach Anderle.

Die Ergebnisse der Gelelektrophorese sind in Abbildung 6.7 gezeigt. Nur im NEB-Puffer konnte die Aktivität der Gyrase nachgewiesen werden. Der Anderle-Puffer zeigte keine Umsetzung.

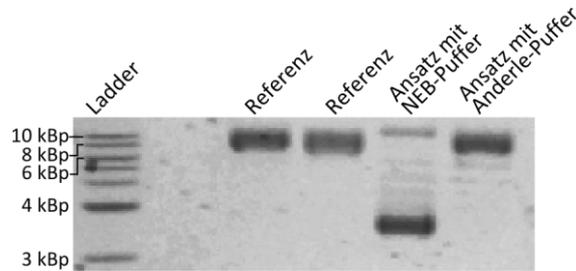


Abbildung 6.7 – Vergleich der Puffer mittels Gelelektrophorese. Bei beiden Referenzen wurden 500 ng relaxiertes DNA-Plasmid verwendet. Im Ansatz des NEB-Puffers ist die Hauptbande weiter gelaufen als in der Referenz; das Plasmid wurde in die supercoiled Form umgesetzt. Im Ansatz nach Anderle fand keine signifikante Reaktion statt. Es sind nur ganz schwache Banden zwischen 8 kBp und 3 kBP zu erkennen. (In dieser Gelelektrophorese wurde der Quick-Load 1 kb DNA Ladder verwendet.)

Mit diesen Vorversuchen konnten drei wichtige Rückschlüsse gezogen werden:

1. Im Anderle-Puffer findet keine Umsetzung von der DNA statt und es kann keine ATP-Hydrolyse gemessen werden. Die fehlende Umsetzung der DNA ist nachvollziehbar, da z. B. Bestandteile wie Spermidin und BSA fehlen, dies sind wichtige Komponenten, die für die Bindung von DNA an der Gyrase notwendig sind [24, 244].
2. Der NEB-Puffer ist mit Phosphat verunreinigt. Der Assay reagiert sehr empfindlich auf Phosphat. Daher kann der Puffer des Herstellers in den weiteren Experimenten zur Bestimmung der Phosphatkonzentration mit dem Farbstoff-Komplex nicht verwendet werden und der zu verwendende Puffer muss selbst angesetzt werden. Für die Herstellung der eigenen Lösungen wurden folgende Anpassungen durchgeführt: Verwendung von bidestilliertem Wasser, Aufbewahrung alle Lösungen und Chemikalien, soweit wie möglich, in Plastikgefäßen und Vermeidung von Kontakt mit Glasgeräten.
3. Obwohl im NEB-Puffer die DNA umgesetzt wurde, konnte kein erhöhter Phosphatwert gemessen werden. Der Umsatz von dem im Reaktionsvolumen enthaltenen 2 µg relaxiertem DNA setzte nicht genügend Phosphat frei, der im Absorptionsassay gemessen werden konnte. Die Alternativen bestanden darin, eine höhere Konzentration des Plasmides oder ein anderes Substrat für die Gyrase zu verwenden. Höhere Konzentrationen des Plasmides wären allerdings unrentabel, da die Herstellung eines relaxierten DNA-Fragments aufwendig und teuer ist. Interessanterweise wurden schon in den ersten Aktivitätsuntersuchungen festgestellt, dass die Gyrase lineare DNA-Fragmente (im konkreten Fall das Colicin E1 Plasmid) umsetzt [235]. Unter gleichen Bedingungen sogar besser als das korrespondierende relaxierte Plasmid. Dies könnte dadurch erklärt werden, dass sich ein lineares DNA-Fragment von alleine entwindet. Die Gyrase müsste dann nie gegen die interne Spannung des coiled oder supercoiled Plasmids anarbeiten. Es handelt sich dadurch um ein unerschöpfliches Produkt. Möglicherweise vereinfachen lineare Fragmente auch die Anlagerung an die Gyrase (das zir-

kuläre, nicked Plasmid, welches sich auch von selbst entwinden kann, wird auch langsamer als das lineare Fragment umgesetzt (235). Das Colicin E1 Plasmid besteht aus ca. 6.6 kBp. Als preiswerte Alternative für ein lineares DNA-Fragment bot sich DNA-Ladder (im Folgenden auch Ladder genannt) an. Dieser besteht aus linearen DNA-Fragmenten und steht, da er standardmäßig in der Gelelektrophorese eingesetzt wird, auch in größeren Mengen zur Verfügung. Darum wurde der „GeneRuler 1kb DNA Ladder“ (Fisher Scientific, Schwerte), der aus DNA-Fragmenten von 250 bis 10,000 BP besteht (mit Anreicherungen bei 1,000, 3,000 und 6,000 Bp), verwendet.

Zunächst wurde ein geeignetes Puffersystem benötigt, bei dem die Gyrase Aktivität zeigte und das nicht mit Phosphat verunreinigt war. Die wichtigsten Unterschiede bei den Bestandteilen vom Anderle- und NEB-Puffer bestanden im Zusatz von Spermidin und BSA im NEB-Puffer. Wie oben beschrieben sind beide Substanzen dafür notwendig, damit die Gyrase mit der DNA interagieren kann. Daher wurden zunächst die beiden Bestandteile in den Konzentrationen, die auch im NEB-Puffer vorhanden waren, verwendet. Die Aktivität wurde mit einem Supercoiling-Assay bestimmt (s. Abbildung 6.8).

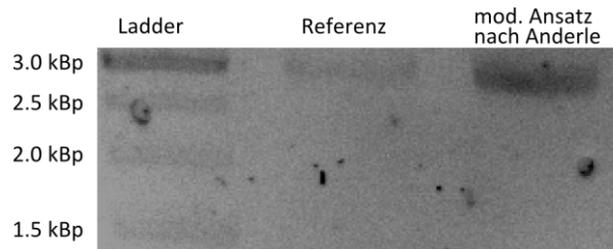


Abbildung 6.8 – Gelelektrophorese mit modifiziertem Anderle-Puffer. Auch nach dem Zusatz von BSA und Spermidin zeigt die Gyrase keine Aktivität. Die Bande mit dem modifizierten Ansatz befindet sich auf gleicher Höhe wie die verwendete Substrat-DNA.

Wiederum zeigte die Gyrase keine Aktivität. Neben marginalen Konzentrationsdifferenzen bestanden die größten Unterschiede in der Kaliumchlorid-Konzentration, Glycerol und EDTA (s. Tabelle 6.2). Da weder Glycerol noch EDTA großen Einfluss auf die Reaktion haben sollten, stellte der wichtigste Unterschied die Kaliumchlorid-Konzentration dar. In der Literatur lässt sich für diese Hypothese eine Bestätigung finden:

Tabelle 6.2 – Vergleich der verwendeten Puffer. Die Gyrase zeigt nur beim NEB-Puffer Aktivität. Der entscheidende Unterschied besteht in der Konzentration von Kaliumchlorid.

Substrat	mod. Anderle-Puffer	NEB-Puffer
TRIS	50 mM	50 mM
Kaliumchlorid	100 mM	24 mM
Magnesiumchlorid	5 mM	4 mM
DTT	5 mM	2 mM
EDTA	1 mM	-
Glycerol	-	6.5 %
ATP	1 mM	1.75 mM
BSA	0.1 mg/mL	0.1 mg/mL
Spermidin	5 mM	5 mM
DNA (relaxiert)	500 ng	500 ng

ALT *et al.* konnten ebenfalls nachweisen, dass die Gesamtgyrase ab einer Konzentration von

100 mM Kaliumchlorid ihre Aktivität verliert [64]. Die Ursachen dafür sind noch nicht geklärt. Dies verdeutlicht das komplexe Gleichgewicht in der enzymatischen Reaktion. Obwohl die Bedingungen für die GyrB-Untereinheit optimiert waren, kann schon die höhere Konzentration eines Elektrolyts die Aktivität der Gesamtgyrase massiv beeinflussen.

Dementsprechend wurde im Folgenden die Zusammensetzung des NEB-Puffers (s. Tabelle 6.2) verwendet, der wegen der Phosphatverunreinigung des NEB-Puffers selbst hergestellt werden musste. Die Gelelektrophorese konnte belegen, dass die Gyrase mit diesem Puffer die gleiche Aktivität wie mit dem original NEB-Puffer zeigte (s. Abbildung 6.9).

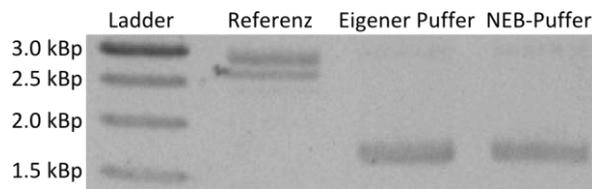


Abbildung 6.9 – Test des eigenen Puffers. Nach Anpassung des Puffers zeigte die Gyrase auch im selbst hergestellten Puffer die gleiche Aktivität wie im NEB-Puffer.

Weiterhin wurde überprüft, ob unter diesen Bedingungen auch die Bestimmung mit dem Phosphat-Assay möglich war. Dazu wurden Blindwertmessungen mit dem eigenen Puffer durchgeführt. Neben dem Puffer selbst wurde auch der Einfluss von ATP überprüft, der wahrscheinlichsten Quelle für Phosphatverunreinigungen (durch eine einfache Hydrolyse des ATPs).

Alle Werte zeigten eine niedrige Phosphatkonzentration und bestätigten eine geringere Verunreinigung der Reagenzien (s. Tabelle 6.3). Der Einfluss von ATP auf den Puffer war gering und stellte keine externe Phosphatquelle dar.

Dies bestätigte sich auch für die Absorptionsmessung mit einer Phosphatstandard-Reihe (s. Experimenteller Teil 11.2.4 auf S. 217). Wie in Abbildung 6.10 zu sehen, zeigten die Blindwerte, die alle Bestandteile des Assays bis auf das Enzym und DNA-Substrat enthielten, nur eine geringe Phosphatverunreinigung.

Tabelle 6.3 – Vergleich der Verunreinigung der Standards mit Phosphat.

Zusammensetzung	A
30 µL bidestilliertes Wasser	0.074
30 µL eigener Puffer	0.288
30 µL eigener Puffer + ATP	0.296
30 µL Phosphatstandard (Endkonzentration: 62 µM) + ATP	0.446

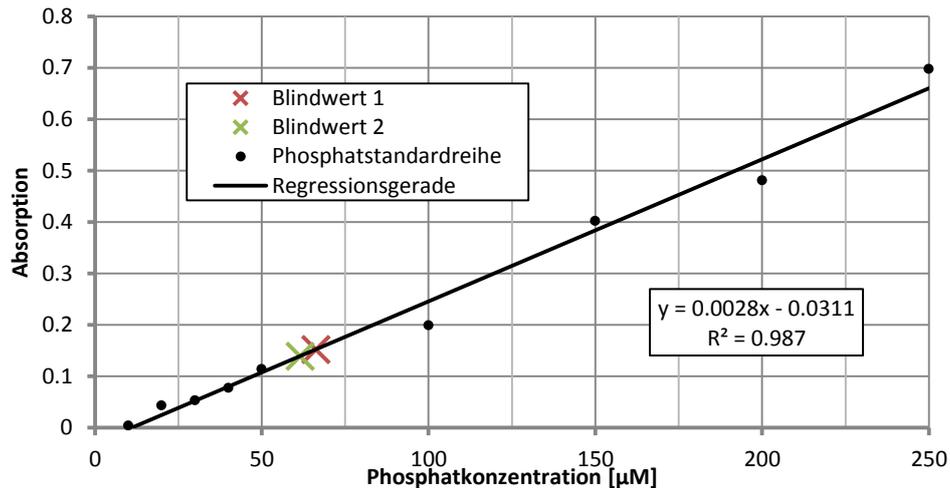


Abbildung 6.10 – Überprüfung des eigenen Puffers. Der Blindwert enthielt alle Bestandteile des Puffers, die für die Inkubation verwendet wurden. Er befindet sich in einem niedrigen Bereich und hat daher auch keinen Einfluss auf die weitere Messung.

Damit erfüllt das Puffersystem alle Ansprüche: Die Gyrase zeigt die gleiche Aktivität wie beim Puffer des Herstellers. Gleichzeitig beeinflusst es weder die Absorptionsmessung noch zeigte es starke Verunreinigungen mit Phosphat.

Somit konnte das System dafür optimiert werden, mehr Phosphat freizusetzen, so dass eine Bestimmung mit dem Phosphat-Assay möglich ist. Für den Assay wurde 1 µg des DNA Ladders als Substrat verwendet. Durch den großen Überschuss an ATP und dem theoretisch unendlich zur Verfügung stehenden Substraten (s. S. 162) konnten die bisher begrenzenden Parameter beseitigt werden. Zunächst wurde eine Inkubationszeit von 4 h gewählt. Dabei setzte die Gyrase 28.7 µM Phosphat frei (s. Abbildung 6.11). Da diese Konzentration noch nicht ausreicht, um eine Abnahme der Aktivität zu messen, wie es bei IC_{50} -Bestimmungen nötig wäre, wurde die Messung im nächsten Schritt auf 20.5 Stunden verlängert (s. Abbildung 6.11). Erwartungsgemäß erhöhte sich die Freisetzung von Phosphat um ein Vielfaches.

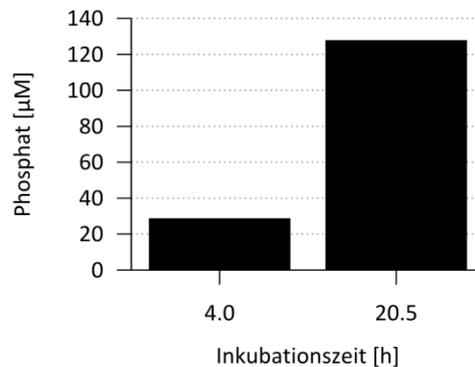


Abbildung 6.11 – Freisetzung von Phosphat in Abhängigkeit der Inkubationszeit.

Zum Abschluss des „*Proof of Concept*“ wurde geprüft, ob die Freisetzung von Phosphat durch die Zugabe eines Inhibitors der ATPase-Untereinheit auch gehemmt wird und dies im Experiment nachweisbar ist. Dafür wurde die Absorptionsmethode mit zwei verschiedenen Konzentrationen von Novobiocin durchgeführt, eine im Bereich des IC_{50} -Wertes und eine in einem Bereich, bei dem die Gyrase vollständig inhibiert werden soll. Wie in Abbildung 6.12 zu sehen, verringerte sich die

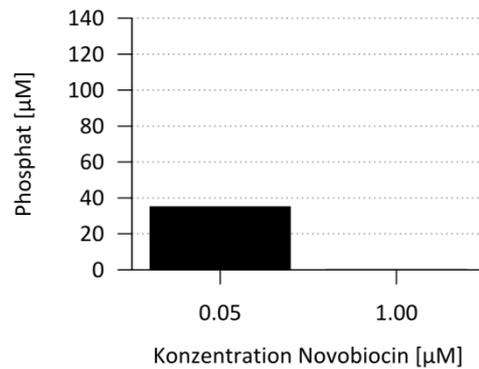


Abbildung 6.12 – Freisetzung von Phosphat in Anwesenheit von Novobiocin (Inkubationszeit 24 h).

Phosphatfreisetzung in Abhängigkeit von der Novobiocinkonzentration. Novobiocin besitzt einen IC_{50} -Wert von ca. 0.06 µM (s. Abschnitt 6.4 auf S. 179). Daher wurde bei einer Konzentration von 0.05 µM eine teilweise Inhibition der Gyrase erwartet. Bei einer Konzentration von 1.00 µM Novobiocin wurde kein Phosphat mehr freigesetzt und daraus folgend keine Enzymaktivität gemessen.

Damit entsprach der Assay, den an ihn gestellten Anforderungen für eine qualitative Bestimmung. Das Vorgehen war wie folgt:

1. In einem Reaktionsgefäß wird die Gyrase in einem Reaktionsvolumen von 30 µL über Nacht (20 h) bei 37 °C mit der in Tabelle 6.4 aufgeführten Zusammensetzung inkubiert.

Tabelle 6.4 – Zusammensetzung des Reaktionsansatzes (s. Experimenteller Teil 11.2 auf S. 216) für die Phosphatbestimmung (Gesamtvolumen 30 µL).

Substrat	Zusammensetzung
bidestilliertes Wasser	17.9 µL
6x-RP	5.0 µL
60x-DTT-Lösung	0.5 µL
25x-Spermidin-Lösung	1.2 µL
DMSO (+ ggf. Inhibitor)	1.2 µL
25x-ATP-Lösung	1.2 µL
DNA-Ladder (0.5 µg/µL)	2.0 µL
Gyrase (5 U/µL)	1.0 µL

2. Die Reaktion wird durch Zugabe von 96 µL Farbreagenz (s. Experimenteller Teil 11.2.3 auf S. 216) gestoppt und die Lösung sorgfältig gemischt.
3. Nach ca. 5 Minuten wird die Farbentwicklung mit 12 µL einer 30 %igen Zitronensäure-Lösung gestoppt.
4. 115 µL werden in eine transparente Mikrotiterplatte überführt und die Absorption bei 620 nm vermessen.

Für quantitative Bestimmungen wurden weitere Untersuchungen der Enzymaktivität benötigt. Dafür wurde zunächst den Einfluss der Inkubationszeit auf die Phosphatfreisetzung ermittelt. Es wurde die Aktivität der Gyrase nach 17, 20 und 23 Stunden Inkubation bestimmt. In Abbildung 6.13 ist zu erkennen, dass in diesem Bereich keine Korrelation zwischen der Inkubationszeit und der freigesetzten Phosphatmenge bestand. Bei 23 Stunden Inkubation wurde weniger Phosphat ermittelt als nach 17 oder 20 Stunden Inkubation.

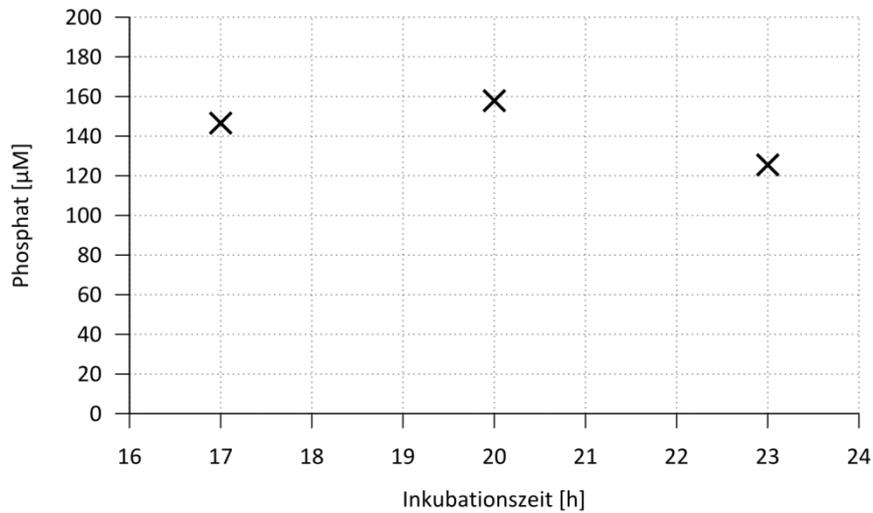


Abbildung 6.13 – Vergleich verschiedener Inkubationszeiten. Zwischen den verschiedenen Inkubationszeiten verändert sich die freigesetzte Phosphatmenge nur wenig. Nach 23 h wurde die kleinste Phosphatmenge bestimmt.

Aus diesen Gründen wurde die Inkubationszeit auf einen Bereich von 8 bis 23 h variiert (s. Abbildung 6.14).

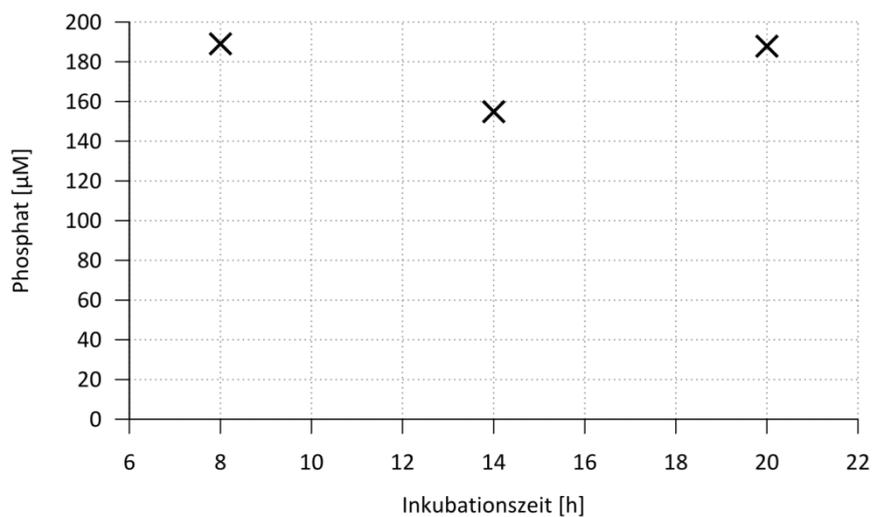


Abbildung 6.14 – Vergleich verschiedener Inkubationszeiten. Bei einer Inkubationszeit von 8 h wurde nahezu die gleiche Menge an Phosphat freigesetzt wie nach 20 h. Bei 14 h Inkubation wurde die geringste Menge bestimmt.

Wieder lagen alle Werte auf einem ähnlichen Niveau und schwankten zwischen 155 und 189 μM Phosphatfreisetzung.

Eine mögliche Erklärung wäre, dass die Inkubationszeit zu lang ist, also schon eine Plateauphase für die Reaktion erreicht worden ist. Obwohl noch ausreichend Produkte für die Umsetzung vorhanden sind, findet keine mehr statt. Dies kann auf den Verlust der Aktivität des Enzyms, durch bspw. Denaturierung, zurückzuführen sein. Dabei handelt es sich um ein bekanntes Phänomen bei Enzymkinetiken und wurde von BISSWANGER ausführlich diskutiert [263]. Eine weitere Problematik bestand in den hohen Phosphatkonzentrationen, weil mit 250 μM schon kritische Konzentrationsbereiche erreicht wurden, bei denen der Phosphomolybdat-Malachitgrün-Komplex ausfällt.

Deswegen bestand das Hauptziel darin die Aktivität der Gyrase zu steigern und die Inkubationszeit deutlich zu verkürzen. 2011 konnten ALT *et al.* zeigen, dass sich die Aktivität der Gyrase in *E. coli* durch die Zugabe von Kaliumglutamat moderat steigern lässt [64]. Eine Inkubation mit zusätzlich 300 mM Kaliumglutamat für 3 h konnte zwar eine höhere Aktivität zeigen, die Messergebnisse waren jedoch noch immer nicht ausreichend reproduzierbar (s. Abbildung 6.15).

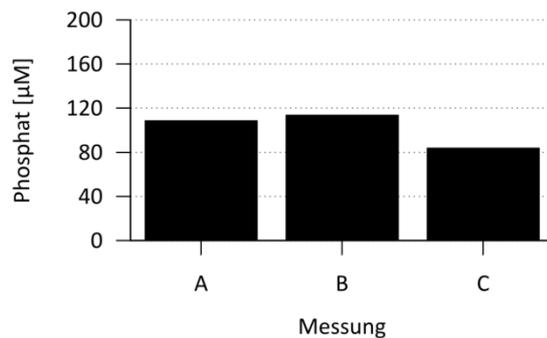


Abbildung 6.15 – Inkubation über 3 h mit 300 mM Kaliumglutamat. Die Ergebnisse schwankten stark. Die Präzision war für weitere Bestimmungen ungenügend ($s_{\text{rel}} = 15.7\%$).

Quantitative Auswertungen waren damit nicht zuverlässig möglich. Für eine weitere Optimierung musste die Inkubationszeit noch weiter reduziert werden. Dafür eignete sich die Absorptionsmethode jedoch nicht, da die Messwertschwankungen in niedrigen Phosphatkonzentrationsbereich zu hoch waren (s. auch Tabelle 6.6 auf S. 176). Dies war ein limitierender Faktor der Messmethode selbst.

6.4 Fluoreszenzauslöschungsmethode

Ein deutlich sensitiveres Verfahren stellte die Fluoreszenzauslöschung dar [255, 256, 264], welches allerdings bisher noch nicht mit der Gyrase durchgeführt wurde. Um diesen Assay zu testen, mussten zunächst einige grundlegende Fragen geklärt werden:

1. *Wie sind die Fluoreszenzeigenschaften der Mikrotiterplatten?* Da, wie in 6.2.3 auf S. 157 beschrieben, die Fluoreszenz der Platte keine vom Hersteller beabsichtigte Eigenschaft ist, muss zunächst überprüft werden, welche Wellenlängen zur Anregung benötigt werden und welche Wellenlängen das fluoreszierende Licht hat.
2. *Wie homogen ist die Fluoreszenz?* Aus den gleichen Gründen, wie bei 1, muss auch sichergestellt werden, dass die Platte über alle Kavitäten gleichmäßig fluoresziert.
3. *Welche Absorptionseigenschaften besitzt der Farbstoff?* In Abhängigkeit der Fluoreszenzeigenschaften der Platte, muss der Farbstoff möglicherweise angepasst werden. In der Erstbeschreibung der Methode von Zuck *et al.* konnte die Messung sowohl mit Chinaldinrot als auch mit Malachitgrün durchgeführt werden. Die Güte der Ergebnisse kann dadurch maßgeblich beeinflusst werden [255].

– Fluoreszenzeigenschaften der Mikrotiterplatte

Der Vorteil dieser Methode besteht darin, dass die Messungen sehr variabel an die Bedingungen, wie Farbreagenzien, Platte oder auch Reaktionsbedingungen, angepasst werden können. Zuck *et al.* verwendete folgende Bedingungen: für Chinaldinrot 430 nm Anregungswellenlänge und 530 nm gemessene Emissionswellenlänge, für Malachitgrün 430 nm/630 nm. Da sich die Anregungs- und Emissionswellenlängen gegenseitig beeinflussen können, sind die Möglichkeiten sehr vielfältig. Im Folgenden werden die wichtigsten Ergebnisse dargestellt.

Zunächst wurde bei einem konstanten Anregungslicht mit einer Wellenlänge von 430 nm verschiedenen Emissionsfilter getestet. Dabei lag das Maximum der Fluoreszenz sehr deutlich bei 568 nm (s. Abbildung 6.16). Mit einem absoluten Wert von 1.8 Mio. RFU befand sich die Fluoreszenz dabei auf einem sehr hohen Niveau. Damit verhielten sich die hier verwendeten Mikrotiterplatten der Firma NUNC deutlich anders als die von Zuck *et al.*, die eine maximale Fluoreszenz bei 500 nm besaßen. Dieser Bereich stellte bei den hier verwendeten Platten ein Minimum dar.

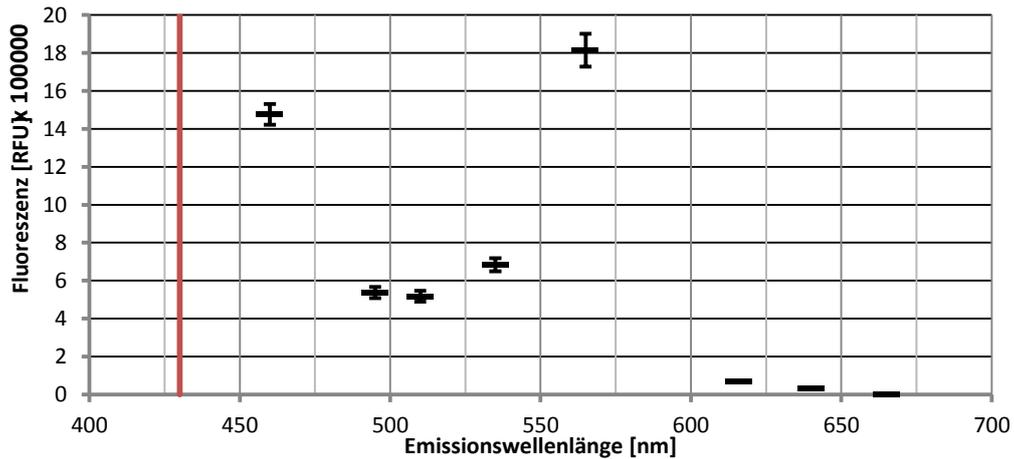


Abbildung 6.16 – Fluoreszenz der weißen Platte bei einer Anregungswellenlänge von 430 nm (rote Linie) und verschiedenen Emissionswellenlängen der zur Verfügung stehenden Filter. Alle Werte sind mit der Standardabweichung als Fehlerbalken angegeben (eine Platte mit 96 Wells). Bei einer Wellenlänge von 460 nm ist der Filter noch zu nahe an der Anregungswellenlänge, wodurch sich das erhöhte Signal erklärt. Die höchste Fluoreszenz wurde bei einer Emissionswellenlänge von 568 nm gemessen.

Damit war der Emissionsfilter für 568 nm am erfolgversprechendsten. Die Fluoreszenz bei dieser Wellenlänge wurden nun mit Variation der Anregungswellenlängen getestet, wie in Abbildung 6.17 zu sehen. Hierbei zeigte sich eine gleichmäßig hohe Fluoreszenz zwischen 430 und 496 nm. Nur bei Anregungswellenlängen von < 420 nm wurde eine geringe Fluoreszenz gemessen. Bei Wellenlängen über 496 nm war die Fluoreszenz höher als der Messbereich des Detektors (> 2 Milliarden RFU). Sofern die Messung nicht zu nahe an der Wellenlänge der fluoreszierten Wellenlänge (568 nm) erfolgte, war also eine variable Anregungswellenlänge (zwischen 430 und 496 nm) möglich.

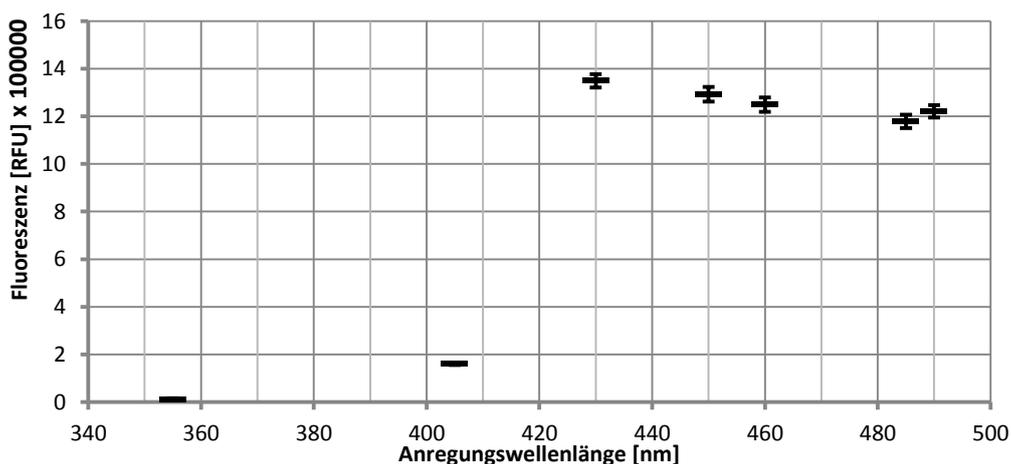


Abbildung 6.17 – Fluoreszenz einer weißen Platte bei konstant gemessener Emissionswellenlänge von 568 nm und verschiedenen Anregungswellenlängen. Die Fehlerbalken geben die Standardabweichung der Messwerte wieder ($n = 96$). Bei Werten unter 420 nm war die gemessene Fluoreszenz sehr niedrig. Bei allen anderen Messungen befand sich die Fluoreszenz auf ähnlich hohem Niveau. Das Maximum befindet sich bei 430 nm. Bei einer Anregungswellenlänge von 530 nm wird der maximale Wert für die messbare Fluoreszenz überschritten (> 2 Milliarden RFU), weil das zur Anregung verwendete Licht selbst als Fluoreszenz gemessen wurde.

Diese Werte gaben einen ersten Anhaltspunkt, in welchem Bereich die Messung erfolgen kann. Für die letztendliche Festlegung sind noch zusätzlich die Absorptionseigenschaften der Farbkomplexe wichtig. Festzuhalten galt jedoch, dass es bei den Emissionsfiltern nur wenig Anpassungsmöglichkeiten gab, da das Maximum bei 568 nm sehr stark ausgeprägt war. Eine andere Wellenlänge würde zu einem erheblichen Verlust in der Sensitivität führen. Die Anregungswellenlänge könnte deutlich flexibler angepasst werden. Für die weitere Analyse wurden die beiden Maxima verwendet: Wenn nicht anders aufgeführt, wurde die Fluoreszenz mit einer Wellenlänge von 430 nm angeregt und bei 568 nm gemessen.

– Homogenität der Fluoreszenz

Für das Wertepaar 430 nm/568 nm (Anregungswellenlänge/Emissionswellenlänge) enthielten die vorherigen Messungen schon erste Ergebnisse über die Homogenität. Die dort gezeigten Werte (Abbildung 6.16 und Abbildung 6.17) wurden unter den genannten Wellenlängen bei derselben Platte, aber an verschiedenen Tagen gemessen. Bei den Werten der einzelnen Kavitäten einer Platte war die Standardabweichung gering (unter 5 %). Dennoch zeigten die Mittelwerte der verschiedenen Messreihen (von jeweils 96 Wells) einen Unterschied von fast 30 % in der Fluoreszenz (1.8 Mio. RFU gegenüber 1.3 Mio. RFU). Die absoluten Werte können je nach Umweltbedingungen (Sonneneinstrahlung, Lagerung der Mikrotiterplatten usw.) variieren. Erklärbar wäre dieses Phänomen z. B. durch unspezifische Phosphoreszenz. Dieser Umstand ist für die Bestimmung der berechneten Absorption nicht entscheidend, da dieses Phänomen durch die Berücksichtigung eines Blindwerts (E_0) wieder eliminiert wird (s. Formel 6.3 auf S. 158). Dennoch müssen diese Einflüsse reduziert werden, da sie potentielle Fehlerquellen darstellen. Die Platte sollte also vor äußerer Strahlung geschützt werden (bei Raumtemperatur im Dunkeln lagern) und während der Arbeit nicht offen in der Sonne liegen. Auch wenn es zwischen den beiden Messungen teilweise große Schwankungen gab, waren die Messwerte innerhalb einer Messung selbst stabil.

Um im Folgenden die Homogenität der Fluoreszenz von den Platten zu ermitteln, wurden an vier verschiedenen Tagen bei vier verschiedenen Platten die Fluoreszenz bestimmt (s. Tabelle 6.5). Auch wenn die Mittelwerte sich unterschieden, waren die relativen Standardabweichungen auf einem vergleichbaren Niveau. Lediglich bei Platte III war die relative Standardabweichung mit über 5 % sehr hoch.

Tabelle 6.5 – Statistische Daten für die Präzision der Messung. Vier verschiedene Mikrotiterplatten (mit 96 Wells) wurden an verschiedenen Tagen vermessen.

Platte	I	II	III	IV
Mittelwert [RFU]	2455976	2485789	2350587	2654225
Standardabweichung [RFU]	33412	35582	154610	44114
Relative Standardabweichung [%]	1.36	1.43	6.58	1.66

Daher ist es erforderlich jede Mikrotiterplatte vor dem Experiment zu überprüfen, ob und in welchem Bereich deren Verwendung im Assay möglich ist (s. Abbildung 6.18).

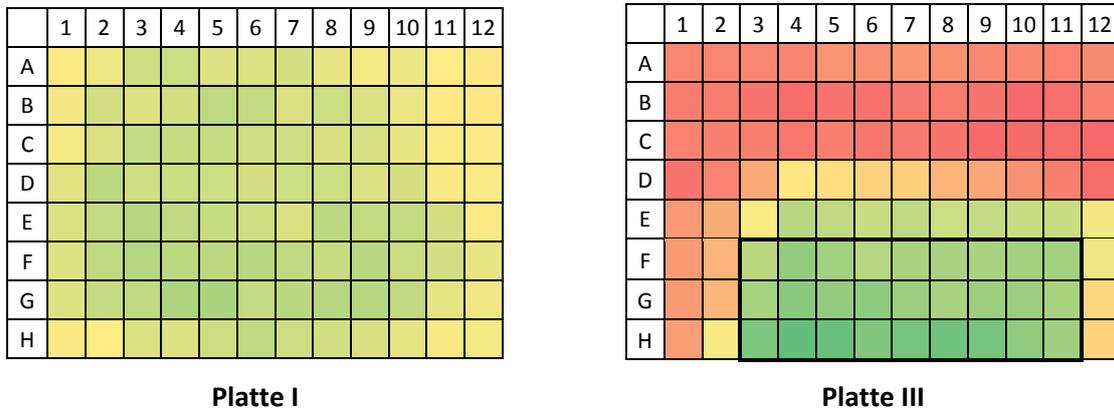


Abbildung 6.18 – Vergleich der RFU-Werte von Platte I (links) und Platte III (rechts). Die einzelnen Wells in den Grafiken sind nach ihrem Wert eingefärbt: von rot (2,6 Mio. RFU) über gelb (2,4 Mio. RFU) nach grün (2,2 Mio. RFU). Während Platte III ab Reihe E inhomogene Fluoreszenzeigenschaften aufwies, verhielt sich Platte I in allen Bereichen homogen.

Typischerweise traten die Schwankungen zeilenweise auf, so dass einzelne Zeilen dennoch verwendet werden konnten. Randspalten zeigten häufiger abweichende Werte als andere Spalten (s. Spalte 12 in Platte III, Abbildung 6.18). Im Fall von Platte III wären Zeilen A bis C noch in einem akzeptablen Bereich für eine Messung. Die Wells des Rechtecks, das sich von F3 bis H11 erstreckt (fett umrandet), wäre ebenfalls für eine Messung geeignet.

– Absorptionseigenschaften des Farbstoffes

Für die Wahl der optimalen Anregungs- und Emissionswellenlängen sind die Absorptionseigenschaften der verwendeten Farbstoffe von entscheidender Bedeutung. Dafür wurden die UV/Vis-Absorptionsspektren von den beiden am häufigsten verwendeten Farbstoffkomplexen (mit Chinaldinrot und Malachitgrün) bestimmt; bestehend aus Farbstoff, Molybdensäure und unterschiedlichen Phosphatkonzentrationen (s. Abbildung 6.19). Die Chinaldinrot-Komplexe besaßen hohe langgezogene Maxima bei 510 nm, während die Malachitgrün-Komplexe zwei Maxima bei 435 und 627 nm aufwiesen.

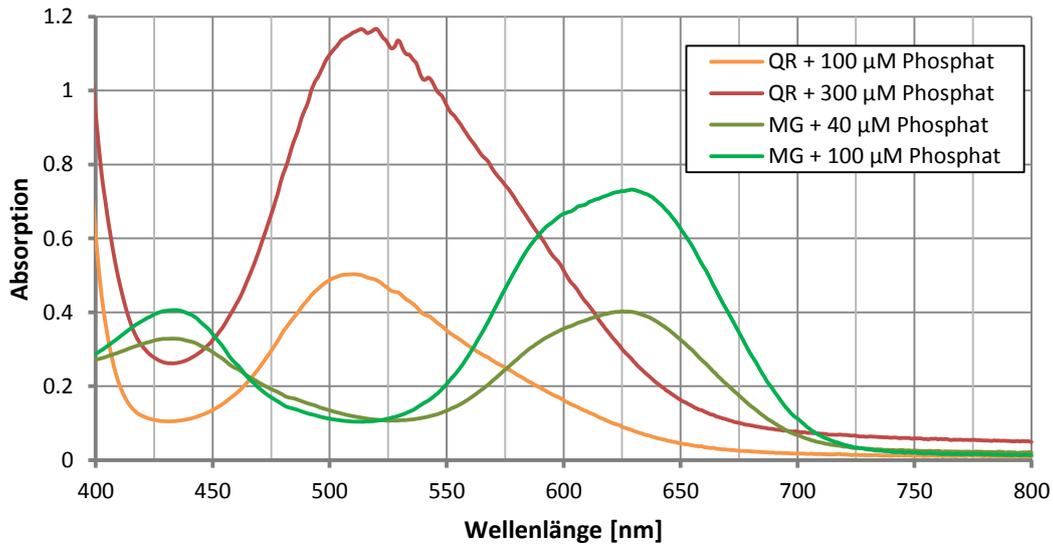


Abbildung 6.19 – Absorptionsspektren von Chinaldinrot (QR) und Malachitgrün (MG) bei verschiedenen Phosphatkonzentrationen. Das Maximum bei QR lag bei 510 nm. MG hatte zwei Maxima: ein kleineres bei 433 nm und ein höheres bei 627 nm. Bei allen Messungen nehmen die Maxima mit steigender Phosphatkonzentration zu.

Bei der Fluoreszenzmessung wurden als optimale Parameter das Paar 430 nm/568 nm ermittelt. Beide Wellenlängen liegen in stark absorbierenden Bereichen des Malachitgrün-Komplexes, so dass bei einer Messung der Fluoreszenz, deren Intensität sowohl von der Absorption des einfallenden Licht als auch von der Absorption des fluoreszierenden Lichts reduziert wird. Der Weg des Lichtstrahls, der absorbiert wird, ist also doppelt so lang. Dies könnte die sowieso schon überlegene Präzision der Fluoreszenzmessung noch weiter vergrößern. Bisher wurde diese Methode nur mit Chinaldinrot durchgeführt, da die Maxima der Fluoreszenz der dort verwendeten weißen Mikrotiterplatten in einem Bereich um 500 nm lagen [255, 256]. Für die in dieser Arbeit verwendeten weißen Platten stellte Malachitgrün jedoch die bessere Wahl dar.

Damit waren die optimalen Parameter für die Fluoreszenzauslöschung ermittelt: Weiße Mikrotiterplatten von NUNC wurden mit einer Anregungswellenlänge von 430 nm bei einer Wellenlänge von 568 nm mit Malachitgrün als Farbstoff vermessen.

Der Fluoreszenz-Assays wurde zunächst mit einer Phosphatstandardreihe geprüft, die Konzentrationen von 0 – 50 µM Phosphat enthielt. Die bisherigen Absorptionsmessungen wurden bei Konzentrationen bis 250 µM Phosphat durchgeführt.

Abbildung 6.20 A zeigt die Fluoreszenzminderung bei steigender Phosphatkonzentration. Die Messwerte zeigten eine logarithmische Abnahme der Fluoreszenz. Zuck *et al.* rechneten diese Werte mit der angegebenen Formel (s. Formel 6.3 auf S. 158) in eine blindwertkorrigierte Absorption um [255]. Diese Auswertung ist in Abbildung 6.20 B dargestellt. Der gemessene Blindwert des Puffers lag im Bereich der Fluoreszenz leerer Wells. Die Pufferlösung hatte also keinen Einfluss auf die Fluores-

zenauslöschung. Durch die Division mit einem konstanten Pufferblindwert wird die letztendlich erhaltene Gerade nur um einen festen Betrag verschoben (Quotientenregel für Logarithmen: $\lg(a/b) = \lg(a) - \lg(b)$). Mathematisch gibt es daher für die Berechnung der Phosphatkonzentration und den Korrelationskoeffizienten kein Unterschied, ob der Blindwert des Puffers berücksichtigt wird oder nicht.

In einem weiteren Auswertungsschritt wurden die Fluoreszenzschwankungen der einzelnen Wells berücksichtigt (s. S. 171). Dafür wurde die Fluoreszenz der zuvor leer vermessenen Wells (E_{leer}) als Referenzwert verwendet (s. Formel 6.6). Im Vergleich zu einer Absorptionmessung wäre dies ein logischer Schritt: Dort geht die Lichtintensität des einfallenden Strahls in die Berechnung ein; dieser ist, in Analogie, für die Fluoreszenz die Energie des Leerwertes von jedem Well. Dies war möglich, da der Puffer, wie oben beschrieben, keinen Einfluss auf den Fluoreszenzwert hatte. Abbildung 6.20 C zeigt diese Auswertung. Dies verbesserte die Regression weiter. Die Verbesserung ist davon abhängig, wie homogen die Fluoreszenz der leeren Wells ist (s. Abbildung 6.18 auf S. 172). Je inhomogener die Fluoreszenz ist, desto größer ist der Einfluss auf die Auswertung. In diesem Beispiel besaß die Fluoreszenz der leeren Wells eine relative Standardabweichung von 1.96 %. Um diese mögliche Fehlerquelle von vornherein zu eliminieren, wurde im weiteren Verlauf die auf jedem Leerwert der Wells basierende Auswertung verwendet.

$$A_{calc} = -\lg\left(\frac{E_t}{E_{leer}}\right)$$

Formel 6.6

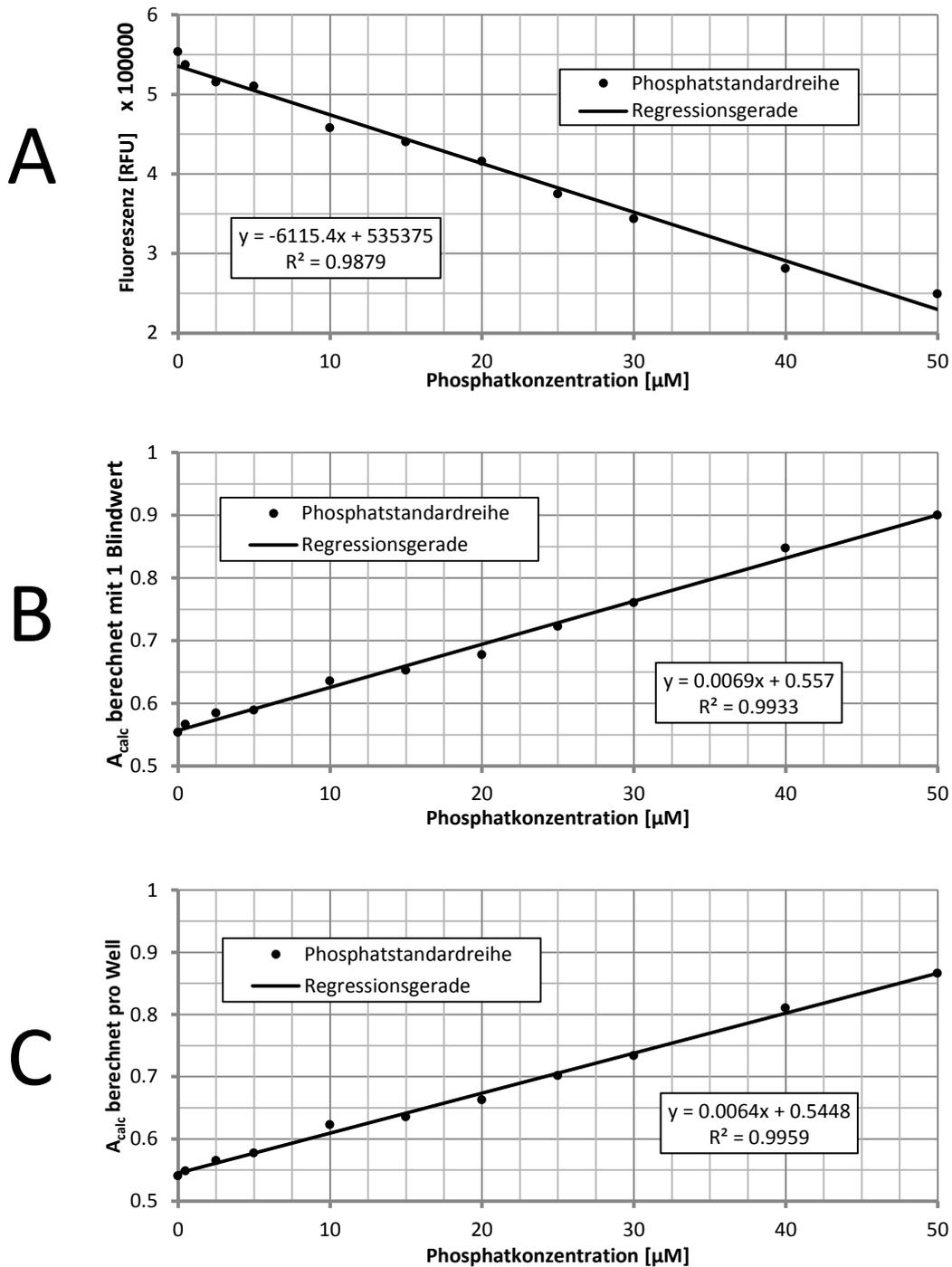


Abbildung 6.20 – Auswertung der Fluoreszenzauslöschung. **A)** In dieser Abbildung ist die Phosphatkonzentration gegen die gemessene Fluoreszenz aufgetragen. Es besteht kein linearer Zusammenhang zwischen diesen beiden Werten. **B)** Dies zeigt die Auswertung mithilfe eines Blindwertes, wie es von Zuck *et al.* beschrieben wurde. Durch die Verwendung des negativ dekadischen Logarithmus verhält sich die Phosphatkonzentration bei der berechneten Absorption direkt proportional. Das Bestimmtheitsmaß (0.9933) ist deutlich besser als bei A (0.9879). **C)** Wird bei der Bestimmung der berechneten Absorption die Fluoreszenz von jedem Well berücksichtigt, ist die Korrelation mit einem Bestimmtheitsmaß von 0.9959 weiter verbessert.

Zur Überprüfung, ob mit dieser entwickelten Fluoreszenzmethode eine höhere Präzision als mit der Absorptionsmethode erzielt werden kann, wurden bei vier verschiedenen Phosphatkonzentrationen jeweils fünffach-Bestimmungen durchgeführt. In Tabelle 6.6 sind die Ergebnisse aufgeführt. Hierbei zeigte sich die deutliche Überlegenheit der Fluoreszenzmethode. Die Präzision war in allen Konzentrationsbereichen höher als bei der Absorptionsmethode. Bei geringen Phosphatkonzentrationen $\leq 10 \mu\text{M}$ zeigte sich dies sehr deutlich, da bei der Absorptionsmethode die relative Standardabweichung noch über 10 % betrug. Demgegenüber waren die bestimmten Präzisionen mit der Fluoreszenzmethode im gesamten Bereich von $2.5 \mu\text{M}$ bis $100 \mu\text{M}$ auf einem gleich hohen Niveau von ca. 1 %.

Tabelle 6.6 – Vergleich der Präzision der Absorptions- und Fluoreszenzmethode. Auf einer Mikrotiterplatte wurde jede Phosphatkonzentration (P_i) im Reaktionspuffer jeweils fünfmal entsprechend der Methode bestimmt. Die Fluoreszenzmethode zeigte über alle Bereiche eine bessere Präzision.

P _i -Konzentration	Absorptionsmethode			Fluoreszenzmethode		
	\bar{x}	s	s _{rel} [%]	\bar{x}	s	s _{rel} [%]
2.5 μM	0.023	0.003	12.71	0.748	0.005	0.69
10 μM	0.044	0.005	11.98	0.789	0.009	1.13
40 μM	0.145	0.008	5.64	1.001	0.011	1.09
100 μM	0.445	0.013	2.86	1.326	0.010	0.74

Neben der Präzision zeigte der Vergleich von Fluoreszenz- und Absorptionsmessung bei der Linearität klare Vorteile für die Fluoreszenzmessung: Hierzu wurde die Verfahrensstandardabweichung S_{x0} und den Verfahrensvariationskoeffizienten V_{x0} ermittelt. Diese treffen im Vergleich zum Korrelationskoeffizienten eine bessere Aussage über die Güte der Korrelation, da die Empfindlichkeit in Form der Steigung berücksichtigt wird. Die Verfahrensstandardabweichung wird mithilfe der Reststandardabweichung S_y berechnet, die die Abweichung jedes Messwertes y_i zur Korrelationsgerade ($f(x) = m \cdot x + c$) beschreibt (s. Formel 6.7) [265].

$$S_y = \sqrt{\frac{\sum [y_i - (m \cdot x + c)]^2}{n-2}}$$

$$S_{x0} = \frac{S_y}{m} \quad \text{Formel 6.7}$$

$$V_{x0} = \frac{S_{x0}}{\bar{x}} \cdot 100 \%$$

Die Verfahrensstandardabweichung ist damit ein Maß dafür, wie groß die Schwankung für die Werte der Abszisse ist und gibt damit auch Informationen über die Empfindlichkeit und Fehleranfälligkeit der Messung.

Für die Bestimmung wurde eine Messung bei sieben Phosphatkonzentrationen (2.5 µM bis 50 µM) mit der Fluoreszenz- und der Absorptionsmethode parallel durchgeführt (s. Abbildung 6.21). Die Verfahrensstandardabweichung (Verfahrensvariationskoeffizient) betrug für die Absorptionsmethode 3.3 µM (14.5 %) und für die Fluoreszenzmethode 1.4 µM (6.35 %). Dies verdeutlichte, dass die Absorptionsmethode für diesen Konzentrationsbereich ungeeignet war. Die Schwankungen für diesen Bereich waren größer als der kleinste Messwert selbst (2.5 µM). Die Fluoreszenzmethode reagiert deutlich sensibler auf Änderungen in der Phosphatkonzentration und lieferte akzeptable Werte für beide Kennzahlen.

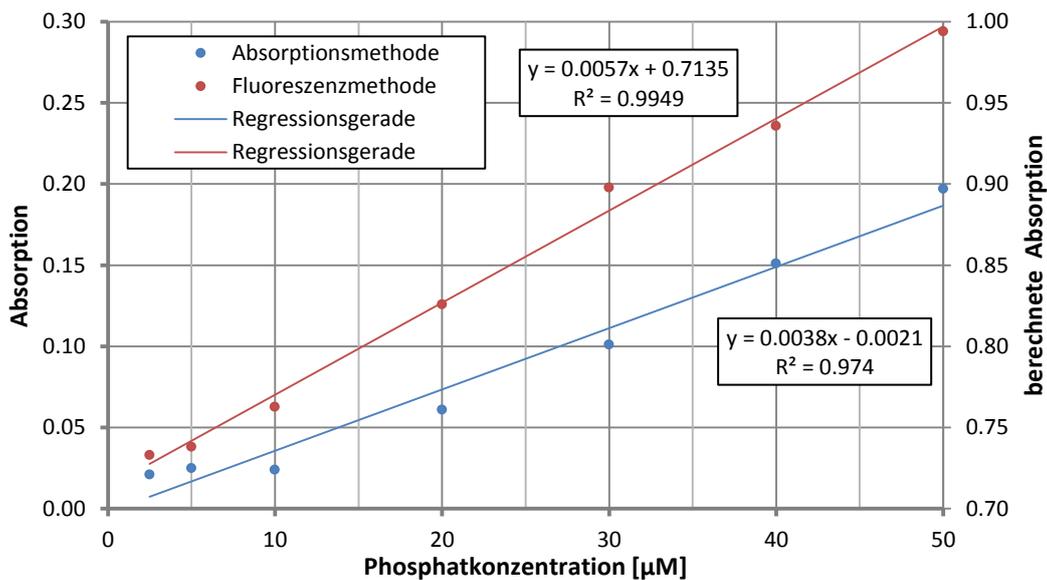


Abbildung 6.21 – Bestimmung der Verfahrensstandardabweichung. Die linke Ordinate zeigt die Absorptionswerte (blau), die rechte die berechnete Absorption für die Fluoreszenzmethode (rot).

Damit konnten nachgewiesen werden, dass die Fluoreszenzauslöschungsmethode durchführbar und der Absorptionsmethode deutlich überlegen ist. Durch die Wahl der Mikrotiterplatten und des geeigneten Farbstoffes konnte die Methode schließlich weiter optimiert werden. Dies ließ sich auch durch eine verbesserte Präzision nachweisen.

– Enzymmessung mit Fluoreszenzmethode

Nach der Etablierung der Messmethode wurde die Übertragbarkeit auf die Enzymmessung getestet. Zunächst wurden wiederum verschiedene Inkubationszeiten verglichen, um den Assay mit dieser Methode zu prüfen (s. Abbildung 6.22). Damit konnte gezeigt werden, dass mit dieser Methode die Bestimmung der Aktivität möglich war und auch deutlich niedrigere Phosphatfreisetzungen unproblematisch waren. Diese Methode eignete sich also, um Phosphat mit einer deutlich kürzeren Inkubationszeit als bei der Absorptionsmethode präzise bestimmen zu können.

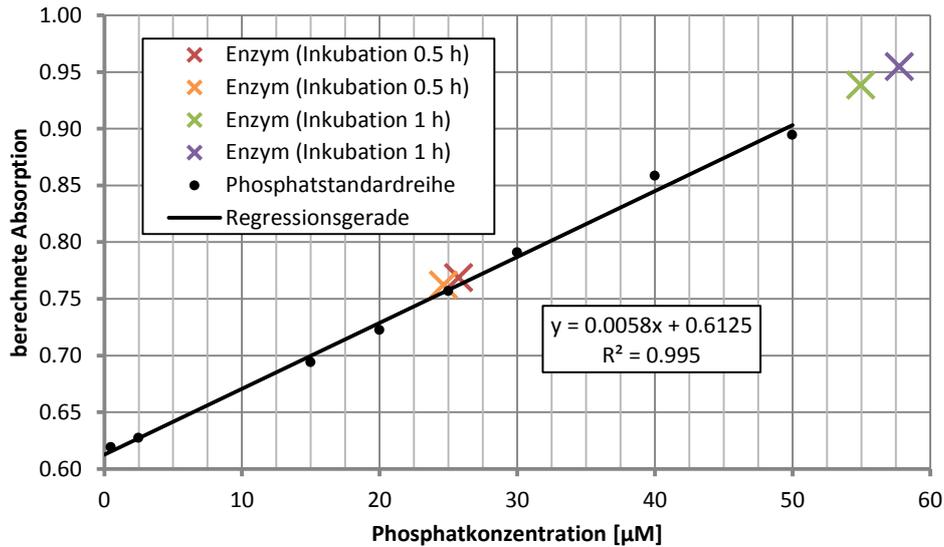


Abbildung 6.22 – Vergleich der Inkubationszeiten. Nach einer halben Stunde Inkubation lag die freigesetzte Menge an Phosphat im Bereich der Phosphatstandardreihe. Die Aktivität des Enzyms lag bei einer Stunde außerhalb der Standardreihe, verhielt sich aber ungefähr proportional zur halbstündigen Messung.

Im Folgenden wurde die Methode für eine Optimierung der reproduzierbaren Bestimmung der Aktivität verwendet. Insbesondere folgende Probleme sollten bei der Aktivitätsbestimmung berücksichtigt werden:

1. Die Aktivität der Gyrase konnte in Abhängigkeit von der Charge, eine unterschiedliche Aktivität zeigen. Die zuerst verwendeten Chargen zeigten deutlich höhere Aktivitäten als die später erhaltenen.
2. Wiederholte Auftau- und Einfriervorgänge verringerten die Aktivität der Gyrase kontinuierlich.
3. Die wichtigste praktische Fehlerquelle stellt die mangelhafte Vermischung des Enzyms mit dem Reaktionspuffer dar.

Daraus folgend entstand das für die folgenden Messungen verwendete Protokoll (s. Experimenteller Teil 11.2.6 auf S. 218) mit diesen Kernelementen:

- Verwendung eines Mastermixes (enthält alle Bestandteile außer: ATP, DNA-Ladder und DMSO (+ evtl. Inhibitor)) und eines Startermixes (enthält ATP und DNA-Ladder)
- Equilibrierung des Master- und Startermixes bei 37 °C für ca. 15 min
- variable Inkubationszeit der Ansätze (in Abhängigkeit von der Aktivität der Gyrase)

Mithilfe dieser Vorgehensweise war die Aktivität der Gyrase mit guter Präzision bestimmbar. Wie in Abbildung 6.23 zu sehen, zeigte der Assay eine hohe Präzision und auch die Hemmung durch Novobiocin konnte klar nachgewiesen werden.

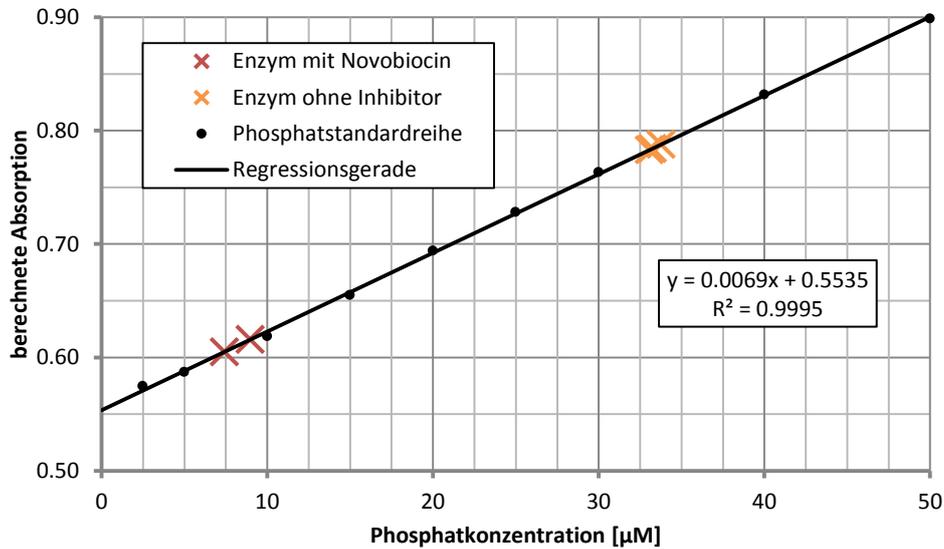


Abbildung 6.23 – Bestimmung der Enzymaktivität und der Hemmung durch Novobiocin. Beim Enzym ohne Inhibitor war die Präzision sehr hoch. Die relative Standardabweichung betrug bei einer Dreifachbestimmung 1.1 %. Sehr deutlich ist auch die erfolgreiche Inhibition der Gyrase durch Novobiocin.

Zur abschließenden Überprüfung des Assays wurde der IC_{50} -Wert von Novobiocin ermittelt. Dafür wurde eine Konzentrationsreihe von Novobiocin mit insgesamt acht Konzentrationen verwendet, die von 10^4 nM bis 10^{-1} nM reichte (s. Abbildung 6.24).

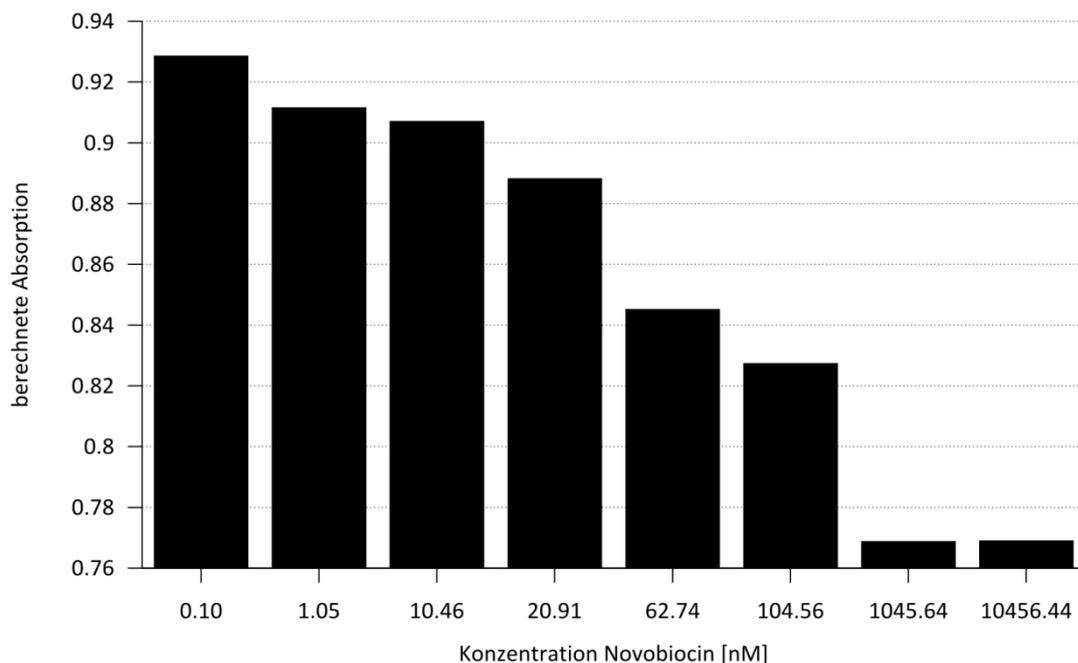


Abbildung 6.24 – Messreihe für die Bestimmung des IC_{50} -wertes für Novobiocin. Mit steigender Konzentration von Novobiocin nimmt die Aktivität der Gyrase ab.

Anhand der berechneten Absorptionswerte kann der IC_{50} -Wert mit der 4PL- oder 5PL-Funktion bestimmt werden (s. Abschnitt 6.2.4 auf S. 159). Für beide Regressionen sind im Folgenden die statistischen Kennzahlen angegeben:

Tabelle 6.7 –Regression des 4PL-Modells. Alle Parameter haben valide statistische Daten mit Signifikanzwerten unter 0.5 % (Null-Hypothese: Parameter hat keinen Einfluss auf die Regression).

	Wert	Std. Fehl.	p-Wert
min	0.766	0.005	0.0000
max	0.921	0.005	0.0000
IC_{50}	67.9	8.910	0.0016
Hillslope	1.14	0.176	0.0030

Tabelle 6.8 –Regression des 5PL-Modells. Durch den zusätzlichen Parameter e , verringert sich der Freiheitsgrad um 1. Dies hatte weitreichende Konsequenzen, da der Standardfehler für Parameter c und e sehr hoch war und nur unzureichende Signifikanzwerte ($p > 60\%$) lieferte.

	Wert	Std. Fehl.	p-Wert
min	0.768	0.006	0.0000
max	0.965	0.308	0.0520
c	202.3	548.3	0.7367
d	0.92	0.006	0.0000
e	2.34	4.74	0.6554

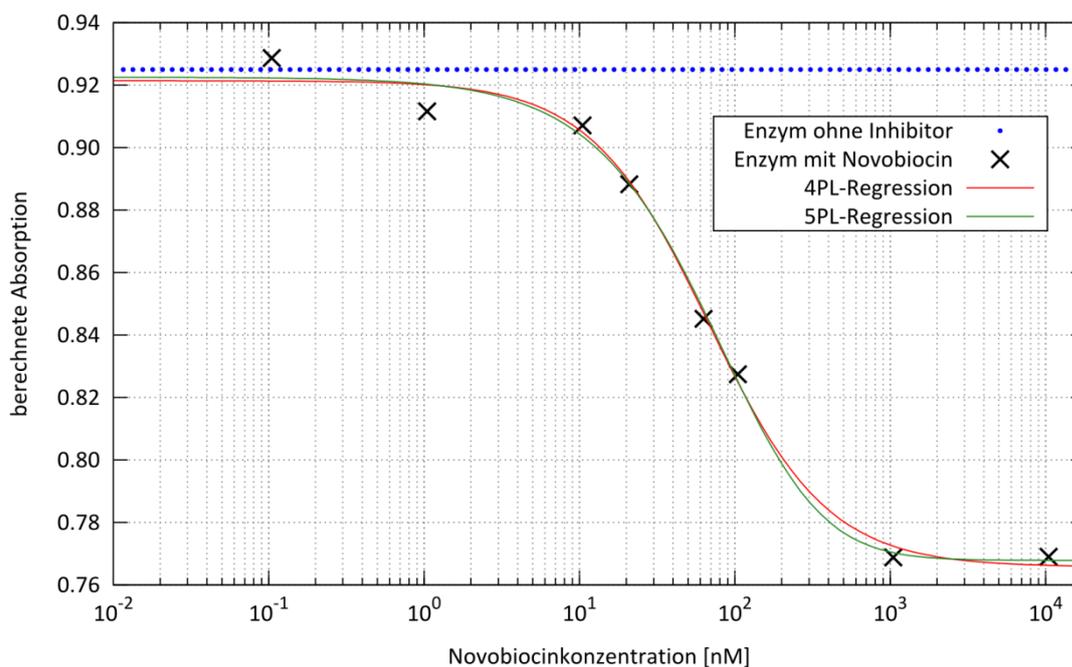


Abbildung 6.25 – Bestimmung des IC_{50} -Wertes für Novobiocin. Beide Regressionsmodelle zeigen einen sehr ähnlichen Verlauf.

Anhand der statistischen Auswertungen der Modelle (Tabelle 6.7 und Tabelle 6.8) zeigte sich die Problematik des 5PL-Modells. Die beiden Parameter c und e hatten sehr hohe Signifikanzwerte. Dies sprach für einen niedrigen Einfluss der Parameter auf die Regression. Weiterhin befand sich der Parameter e nicht im kritischen Bereich (0.2 – 0.6) [259], der ein Indikator für eine hohe Fehleranfälligkeit des 4PL-Modells wäre. Der visuelle Vergleich der beiden Regressionen (s. Abbildung 6.25) verdeutlichte die Ähnlichkeit des Verlaufs. Schließlich bestätigte auch eine ANOVA-Analyse, dass eine Reduktion vom 5PL- auf das 4PL-Modell möglich war (s. Tabelle 6.9).

Tabelle 6.9 – ANOVA-Analyse des 4PL- und 5PL-Modells. Die Nullhypothese, das 5PL-Modell ist dem 4PL-Modell nicht überlegen, kann nicht widerlegt werden. Der Signifikanzwert liegt bei über 60 %.

Modell	Freiheitsgrade	RSS	Df	F-Wert	p-Wert
4PL	4	1.62E-04			
5PL	3	1.47E-04	1	0.324	0.61

Alle statistischen Kennzahlen sprachen daher für die Verwendung des 4PL-Modells. Damit betrug der für Novobiocin ermittelte IC_{50} -Wert $67.9 \text{ nM} \pm 8.9 \text{ nM}$.

ALT *et al.* bestimmten unter ähnlichen Bedingungen (Verwendung der Gesamtgyrase, Pufferzusammensetzung, ATP-Konzentration), allerdings mithilfe eines Supercoiling-Assays, einen IC_{50} -Wert von 80 nM . Dieser Wert ist in guter Übereinstimmung mit dem ermittelten IC_{50} -Wert von $67.9 \text{ nM} \pm 8.9 \text{ nM}$.

7 Biologische Testung

Die im Screening gefilterten und gekauften Substanzen (s. Abschnitt 5.6.5 auf S. 141) wurden schließlich in dem im Rahmen dieser Arbeit entwickelten Assays (s. Experimenteller Teil 11.2.6 auf S. 218) getestet.

7.1 Erste Auswahl

Für jede Substanz wurde die inhibitorische Aktivität in der ATP-Bindetasche der Gyrase in einer Doppelbestimmung mithilfe der Fluoreszenzauslöschungsmethode bestimmt. In allen Messreihen wurde jeweils eine Negativkontrolle (-) – Messung ohne Zugabe einer weiteren Substanz – durchgeführt. Zusätzlich wurden jeweils zwei Positivkontrollen (+) verwendet, die Novobiocin in einer Konzentration enthielten, die die Gyrase vollständig inhibierte. In Abbildung 7.1 bis Abbildung 7.4 sind die Ergebnisse der Messung in der gleichen Reihenfolge dargestellt, wie sie auch während des Experiments getestet wurden. Dadurch können bspw. Trends oder Verschleppungen besser nachverfolgt werden. Für das erste *in vitro* Screening lag der Fokus auf dem Auffinden von neuen Hitstrukturen. Viel wichtiger als die quantitative Einschätzung war daher die qualitative Bewertung, ob eine Hemmung der Gyrase stattfand oder nicht. Daher waren kleinere Schwankungen während der Messreihe auch akzeptabel und vernachlässigbar. Alle Substanzen wurden auf einem Niveau von 100 μM getestet. Für jede Substanz wurde zusätzlich überprüft, ob sie durch Beeinflussung des Phosphomolybdat-Malachitgrün-Komplex falsch-negative Ergebnisse lieferte oder den Assay, z. B. durch Eigenfluoreszenz, verfälschten.

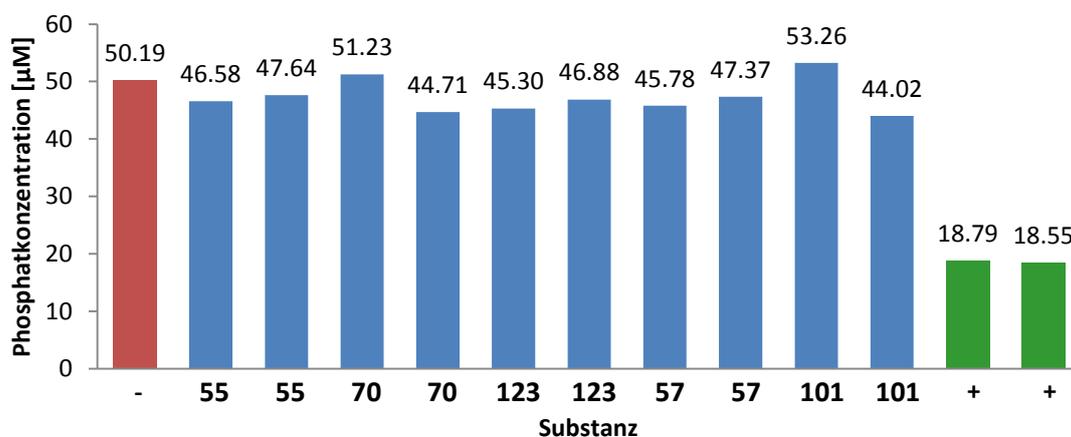


Abbildung 7.1 – Messung 1. Keine Substanz inhibierte die Phosphatfreisetzung auf dem Niveau von der Positivkontrolle. +: Positivkontrolle; -: Negativkontrolle.

7.1 Biologische Testung

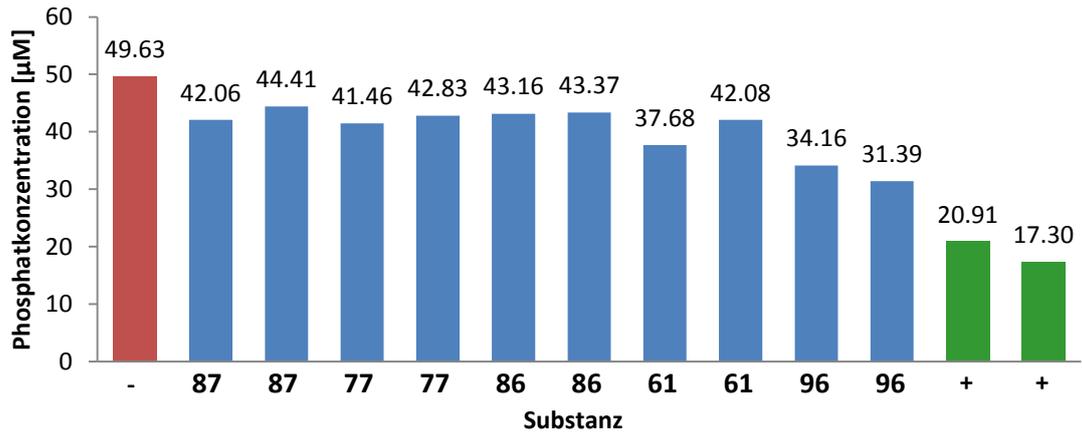


Abbildung 7.2 – Messung 2. Keine Substanz inhibierte die Phosphatfreisetzung auf dem Niveau von der Positivkontrolle. Bei 96 war die Freisetzung jedoch deutlich niedriger und bei 61 leicht niedriger als bei der Negativkontrolle. +: Positivkontrolle; -: Negativkontrolle.

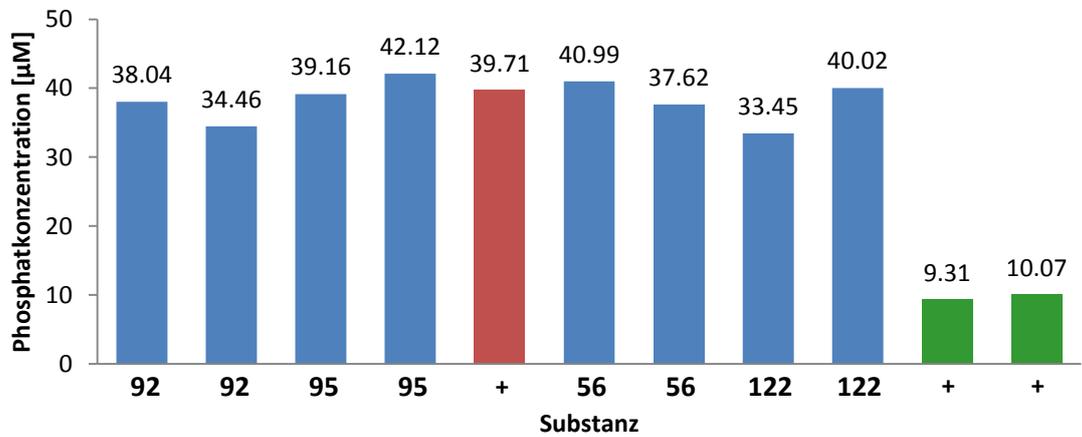


Abbildung 7.3 – Messung 3. Keine Substanz inhibierte die Phosphatfreisetzung auf dem Niveau von der Positivkontrolle. +: Positivkontrolle; -: Negativkontrolle.

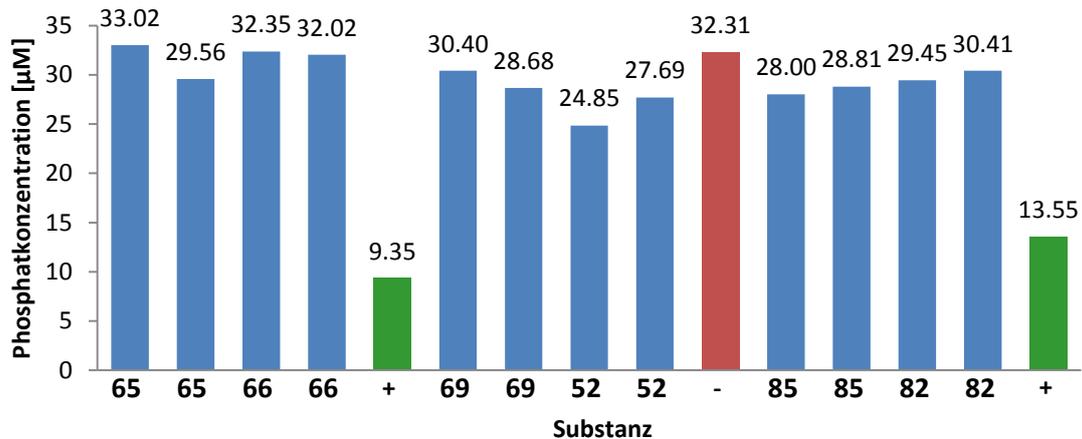


Abbildung 7.4 – Messung 4. Keine Substanz inhibierte die Phosphatfreisetzung auf dem Niveau von der Positivkontrolle. +: Positivkontrolle; -: Negativkontrolle.

Keine der untersuchten Substanzen (s. Abbildung 7.1 bis Abbildung 7.4) zeigte eine Hemmung auf dem Niveau von Novobiocin. Lediglich die Substanzen **96** und **61** in Abbildung 7.2 zeigten Tendenzen, die auf eine schwache Hemmung hindeuteten. In diesem Fall lagen jedoch auch die Positiv-Werte von Novobiocin schon bei einer hohen Phosphatkonzentration (im Durchschnitt bei 19.1 μM). Für beide Substanzen wurde zur Verifizierung ein Supercoiling-Assay (s. Experimenteller Teil 11.2.5 auf S. 218) durchgeführt. Dafür wurde **96** bei einer Konzentration von 100 und 200 μM und **61** bei 100 μM getestet (s. Abbildung 7.5). Keine der Substanzen wies im Supercoiling-Assay eine nachweisbare Hemmung auf.

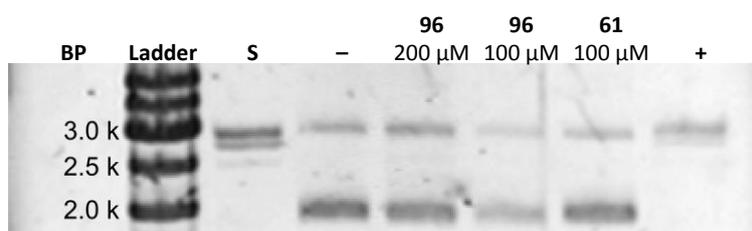


Abbildung 7.5 – Supercoiling-Assay der Hit-Kandidaten von der ersten Auswahl. Für die Substanzen konnte keine Inhibition der Gyrase nachgewiesen werden. BP: Basenpaare; S: reines Substrat ohne Inkubation; -: Inkubation ohne Inhibitor (Negativkontrolle); +: Inkubation mit Novobiocin (Positivkontrolle).

7.2 Zweite Auswahl

Für die zweite Auswahl wurden die Substanzen aus Abschnitt 5.6.6 auf S. 144 getestet, die mithilfe eines 3D-QSAR Modells und den Erkenntnissen der ersten Messreihe ausgewählt wurden. In der zweiten Messung wurde mit diesen Substanzen genauso wie bei der ersten Messreihe verfahren, mit dem Unterschied, dass die Substanzen zunächst auf einen Niveau von 500 μM getestet wurden.

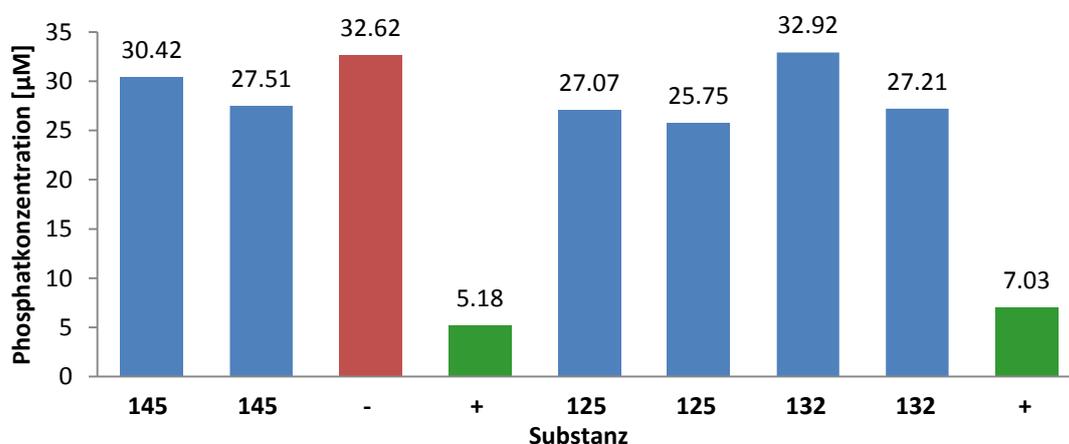


Abbildung 7.6 – 1. Messung. Keine Substanz wies eine signifikante Hemmung auf. +: Positivkontrolle; -: Negativkontrolle.

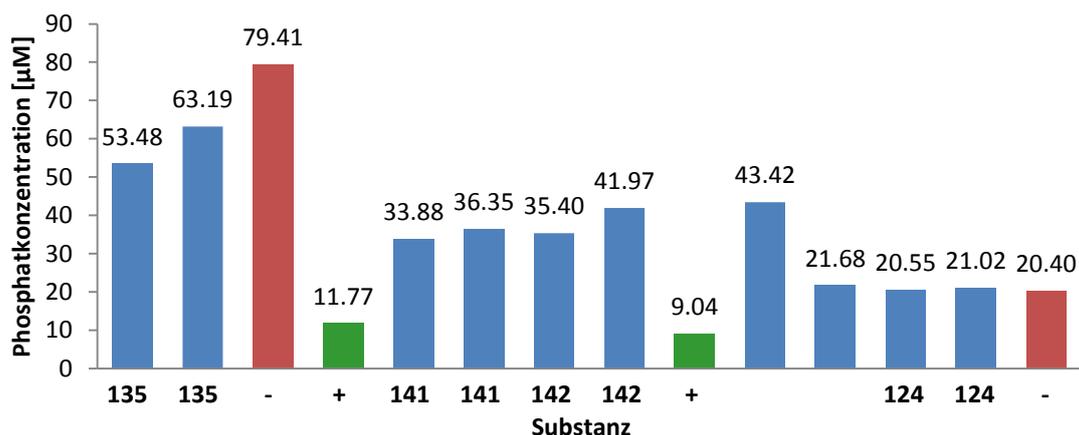


Abbildung 7.7 – 2. Messung. Die Enzymlyösung war unzureichend gemischt, so dass die Phosphatfreisetzung stark schwankte. Der Vergleich mit der Negativ- und Positivkontrolle ließ dennoch den Schluss zu, dass keine Substanz eine Aktivität besaß. Dies konnte in separaten Messungen^{xiv} bestätigt werden. +: Positivkontrolle; -: Negativkontrolle.

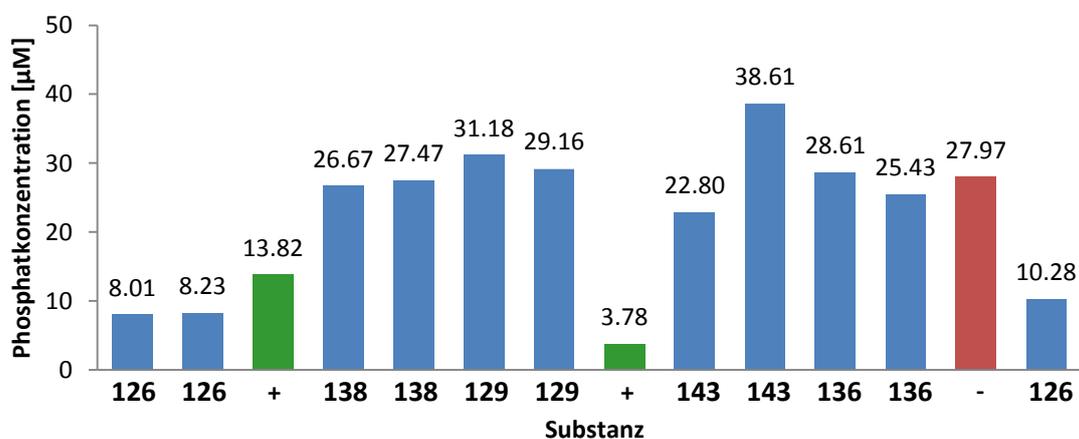


Abbildung 7.8 – 3. Messung. Die Substanz **126** inhibierte die Phosphatfreisetzung auf dem gleichen Niveau wie die Positivkontrolle. Damit war Substanz **126** ein Hit-Kandidat. +: Positivkontrolle; -: Negativkontrolle.

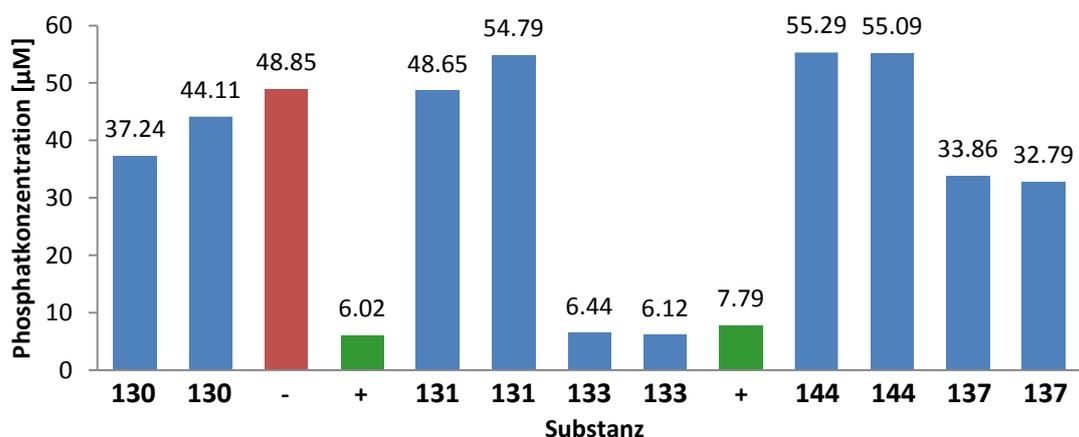


Abbildung 7.9 – Messung 4. Substanz **133** konnte die Freisetzung auf dem gleichen Niveau wie Novobiocin hemmen. Bei Substanz **137** konnte eine leicht verminderte Phosphatkonzentration bestimmt werden. Die Substanz zeigte aber in dem Kontrollversuch eine schwache Eigenfluoreszenz, die in der Folge zu einer leichten Abnahme der bestimmten Phosphatkonzentration führte. Daher handelte es sich bei **137** um keinen Hit-Kandidaten. +: Positivkontrolle; -: Negativkontrolle.

^{xiv} Persönliche Kommunikation mit Pia Schieback aus der Arbeitsgruppe von Thomas Lemcke Institut für Pharmazie der Universität Hamburg.

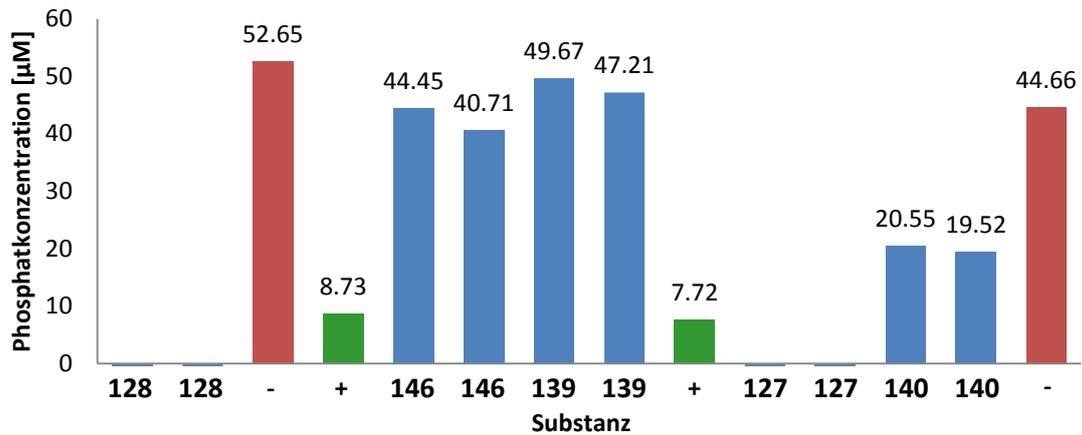


Abbildung 7.10 – Messung 5. Substanz **128** und **127** waren beide fluoreszierende Substanzen. Dies führte dazu, dass die scheinbar bestimmte Phosphatkonzentration negative Werte besaß (für **128**: $-2.21 \mu\text{M}$ und $-3.89 \mu\text{M}$; für **127**: $-111.19 \mu\text{M}$ und $-112.49 \mu\text{M}$). **140** wies eine partielle Inhibition der Gyrase auf. +: Positivkontrolle; -: Negativkontrolle.

In dieser Messreihe inhibierten drei Substanzen die Gyrase: **126** (s. Abbildung 7.8), **133** (s. Abbildung 7.9) und **140** (s. Abbildung 7.10). Sowohl **126** als auch **133** wiesen bei $500 \mu\text{M}$ eine vergleichbare Hemmung wie Novobiocin auf. Daher war bei diesen beiden Substanzen davon auszugehen, dass sie bei einer Konzentration von $500 \mu\text{M}$ die Gyrase komplett inhibieren. Bei **140** zeigte die Gyrase noch eine partielle Aktivität. Aufgrund ihrer starken Eigenfluoreszenz konnten die Substanzen **128** und **127** (s. Abbildung 7.10) nicht ausgewertet werden.

Die Gelelektrophorese bestätigte diese Ergebnisse (s. Abbildung 7.12). Alle Hit-Kandidaten aus der zuvor gefundenen Messung konnten verifiziert werden. Zusätzlich konnte auch für Substanz **127** nachgewiesen werden, dass sie das Supercoiling der DNA durch die Gyrase inhibieren kann. Ohne eine weitere Messmethode kann dieser Hit-Kandidat jedoch nicht validiert werden.

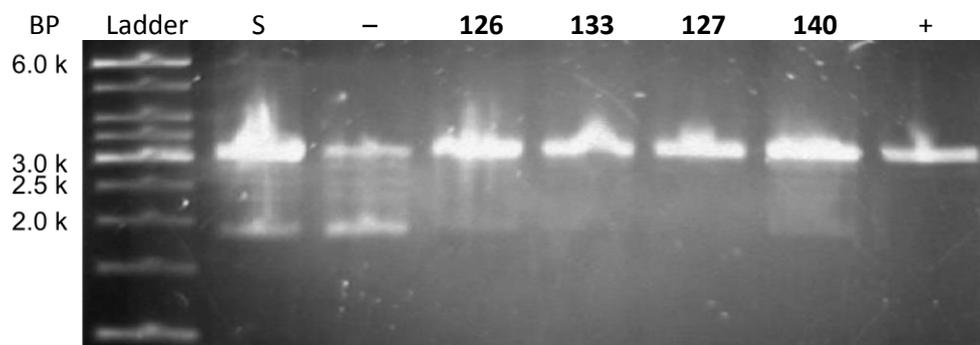


Abbildung 7.11 – Supercoiling-Assay der Hit-Kandidaten von der zweiten Auswahl. Alle Substanzen wurden bei einer Konzentration von $500 \mu\text{M}$ getestet und verhinderten das Supercoiling der DNA. BP: Basenpaare; S: reines Substrat ohne Inkubation; -: Inkubation ohne Inhibitor (Negativkontrolle); +: Inkubation mit Novobiocin (Positivkontrolle).

Zur abschließenden analytischen Charakterisierung der identifizierten Hits wurde schließlich mittels HPLC und LC-MS die Reinheit und Identität untersucht (s. Experimenteller Teil 11.1 auf S. 213). Alle Substanzen besaßen bei einer Wellenlänge von 250 nm eine über HPLC ermittelte Reinheit von > 95 % (s. Anhang 12.4 auf S. 232 ff.). **126** und **140** wurden anhand ihrer Masse eindeutig identifiziert.

Die Masse von **133** stimmte jedoch nicht mit dem erwarteten Wert überein: [**133** + H⁺] sollte eine exakte Masse von 381.1458 Da besitzen, experimentell ermittelt wurde aber eine Masse von 380.1504 Da (s. Anhang 12.4 auf S. 234). Die Differenz der Masse ließe sich durch Substitution einer Aminogruppe mit einer Methylgruppe erklären. Eine Elementaranalyse konnte diesen Verdacht erhärten (s. Anhang 12.4 auf S. 234). Es handelte sich bei der getesteten Substanz nicht um ZINC171177625 (**133**), sondern um dessen Derivat ZINC17117749 (**149**; s. Abbildung 7.12), welches beim Lieferanten unter sehr ähnlichem Namen verfügbar war: Amb21997829 (**133**) gegenüber Amb21997834 (**149**). Die exakte Masse von [**149** + H⁺] beträgt 380.1506 Da und liegt in guter Übereinstimmung mit dem experimentell bestimmten Wert von 380.1504 Da (s. o.).

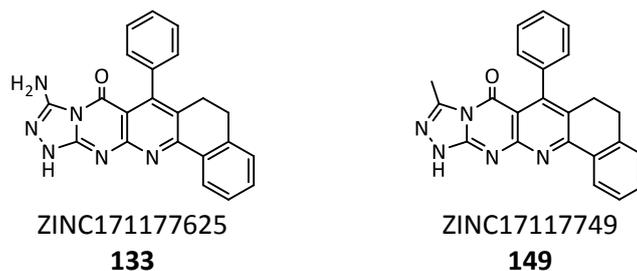


Abbildung 7.12 – Links: Die ursprünglich bestellte Substanz **133**. **Rechts:** Die erhaltene Substanz **149**.

8 Evaluierung

In dieser Arbeit konnten drei neue, bisher unbekannte Inhibitoren der ATPase der GyrB-Untereinheit gefunden werden: Substanzen **126**, **133** und **140** (s. Abschnitt 7.2 auf S. 187). Im Folgenden soll zunächst eine allgemeine Evaluierung der strukturbasierten Leitstruktursuche erfolgen. Im Anschluss werden die Substanzen, ihre Cluster, ihr potentieller Bindungsmodus und die Weiterentwicklungsmöglichkeiten diskutiert.

Im Rahmen dieser Arbeit wurden drei verschiedene Röntgenkristallstrukturen der Gyrase von *E. coli* identifiziert, für deren Bindetasche unterschiedliche Konformationen vorlagen (s. Abschnitt 2). Im weiteren Verlauf konnte nachgewiesen werden, dass die Substanzen, die in die jeweiligen Bindetaschen gedockt wurden, sich bei den einzelnen Strukturen sehr stark unterschieden (s. Abschnitt 5.6 auf S. 123). Dies bestätigte das Vorgehen, für das Screening unterschiedliche Proteinstrukturen zu verwenden, um eine höhere Diversität an potentiellen Hits zu erlangen. Die genaue Analyse der Wassermoleküle in der Bindetasche (s. Abschnitt 2.3.2 auf S. 50) erwies sich ebenfalls als essentiell. Sowohl die korrekte Vorhersage von Wat A, ohne das aufgrund der strukturellen Instabilität vermutlich kein Hit gefunden worden wäre, als auch die optionale Struktur von 3G7E mit Wat B waren wichtig. Der Hit **140** wurde aufgrund eines Dockingergebnisses in letzterer Struktur ausgewählt (s. u.).

Für die Substanzbibliothek wurde eine gefilterte Version der ZINC-Datenbank gewählt. Neben den Filterkriterien für chirale Substanzen waren vor allem die Molekülgröße und die Anzahl der drehbaren Bindungen für die erste Reduktion entscheidend (s. Abschnitt 3.3 auf S. 66). Während der biologischen Testung bereiteten jedoch einige Substanzen Probleme, da sie weder in Wasser noch in DMSO vollständig löslich waren (bspw. auch der Hit **133**). Die cLogP-Werte wurden während der Erstellung der Filterkriterien ebenfalls analysiert (s. Abschnitt 3.3 auf S. 71), aber aufgrund unplausibler Protonierungen und Tautomere nicht weiter verfolgt. Des Weiteren erschienen diese Werte auch innerhalb eines günstigen Bereichs verteilt zu sein. Möglicherweise favorisiert der Screening-Prozess auch lipophile Substanzen – ein bereits häufig beobachtetes Phänomen und Problem in der Leitstruktursuche [266]. Es kann eine weitere Optimierung auch in Richtung hydrophilerer, kleinerer Moleküle erfolgen, so dass auch aus retrospektiver Sicht auf eine weitere Filterung verzichtet werden könnte.

Das Docking mit cRAISE (s. Abschnitt 4.2.3 auf S. 98) ist grundlegend positiv zu bewerten. Mithilfe dieses Tools gelang es, innerhalb weniger Monate, in der Summe über 15 Mio. Protein-Ligand-Komplexe zu erzeugen sowie sie zu bewerten und den Datensatz dabei erheblich zu reduzieren. Ein besonderer Vorteil, der nicht geplant und vorhersagbar war, bestand in der sehr guten Kombination der Zielstruktur der Gyrase, der Docking-Methode von cRAISE und dem Rescoring mit PLP-Score während der Nachprozessierung (s. Abschnitt 0 auf S. 111). Diese ermöglichte eine außergewöhnlich gute

Anreicherung von aktiven Substanzen. Dadurch war es im Folgenden möglich, die Anzahl der Posen durch das Leader-Clustering (s. woanders) drastisch zu reduzieren.

Für die biologische Testung konnte einerseits ein Supercoiling-Assay etabliert werden. Andererseits konnte ein bekannter auf einer Absorptionsmessung basierender Assay, der bisher nur für die GyrB-Untereinheit durchgeführt wurde, auf die Gesamtgyrase übertragen und angepasst werden (s. Abschnitt 6.2.2 auf S. 157). In einer weiteren Optimierung konnte die Reproduzierbarkeit und Nachweisgrenze durch die Weiterentwicklung zu einem auf der Fluoreszenzauslöschung basierenden Assay, der noch nicht für die Gyrase beschrieben ist, erheblich gesteigert werden. Dies ermöglichte die schnelle und sichere Bestimmung der Aktivität einer relativ großen Anzahl an Substanzen (s. Abschnitt 6.2.4. auf S. 159).

Ein zunächst aufgetretener Kritikpunkt in dieser Arbeit bestand darin, dass bei der ersten Auswahl keine Substanz im getesteten Konzentrationsbereich eine inhibitorische Aktivität zeigte. Eine erste Evaluierung fand schon in Abschnitt 5.6.6 auf S. 144 statt. Neben der bereits diskutierten Problematik des strukturbasierten Designs lag ein weiterer wichtiger Grund in der begrenzten Anzahl an erworbenen Substanzen. Es wurden lediglich 20 Substanzen gekauft. Das entspricht nur einem Anteil von ca. 6 ppm von den ursprünglich 3.4 Mio. Substanzen, die mit cRAISE gedockt wurden. Zum Auffinden aktiver Substanzen musste bei 20 gekauften Substanzen die Erfolgsrate nach allen Reduktionsschritten bei 5 % liegen, um zumindest einen Hit zu finden. Zusätzlich erschwert die Auswahl einer kleinen Anzahl zu testender Substanzen die Wahl der Hypothesen, die für die erfolgreiche Vorhersage der Bindung angenommen werden können: Dies betrifft z. B. die Auswahl von protonierten Pyridinen in der tiefen E-Tasche von 3G7E, die womöglich ungeeignet sind, dort zu binden. Zur Überprüfung einer Bindungshypothese war es jedoch unzureichend nur eine einzige Struktur mit diesem Merkmal zu erwerben, da die Bindung von vielen verschiedenen Faktoren beeinflusst werden kann, wie die zusätzlichen Interaktionen in anderen Bereichen der Bindetasche.

Dennoch konnten mithilfe dieser ersten Auswahl wichtige Fragestellungen aufgegriffen und analysiert werden, so dass der Gesamtprozess letztendlich zu einem erfolgreichen Ergebnis führte. Dies lässt sich schon damit belegen, dass in der Folge drei Hits gefunden werden konnten. Der Hit **126** wurde bspw. explizit aufgrund der aus der ersten Auswahl getroffenen Hypothesen zum Kauf ausgewählt (s. u.).

– Substanz **126**

Bei **126** handelte es sich nicht um einen Clusterleader. Tatsächlich handelte es sich sogar um eine Substanz, die verhältnismäßig unähnlich zum Clusterleader **150** war (Tanimotokoeffizient 0.72 (der Cutoff-Wert betrug 0.70); s. Abbildung 8.1). Für diese Substanz wurden die für die Auswahl verwendeten Ansätze kombiniert: Zum einen wurde die im 3D-QSAR gut bewertete und gedockte Substanz **150** ausgewählt, zum anderen wurde die Substanz **126** erworben, welche einen deutlich kleineren Rest besaß, der in der E-Tasche binden würde. Anhand der Ergebnisse der ersten biologischen Testungen und den daraus geschlussfolgerten Erkenntnissen sollte die Substanz daher eine bessere Aktivität aufweisen (s. Abschnitt 5.6.6 auf S. 144). In dem zugehörigen Cluster waren insgesamt 28 Substanzen enthalten. Der in **126** vorhandene Cyclopropyl-Gruppe bildete dabei den kleinsten Rest für diesen Bereich. Bei einem anschließend durchgeführten Docking mit Glide in Analogie zum Docking mit den Clusterleadern (s. Abschnitt 5.5 auf S. 120) konnte für **126** eine vergleichbare Pose wie für **150** generiert werden (s. Abbildung 8.2).

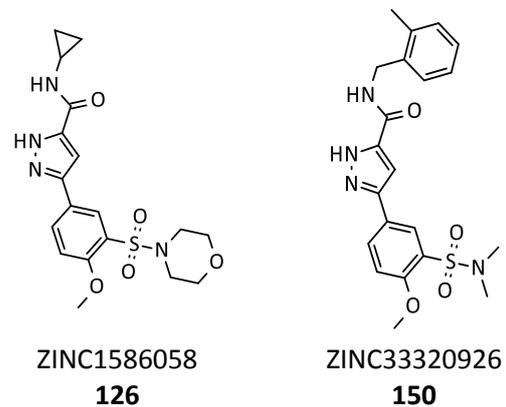


Abbildung 8.1 – Links: Der ausgewählte Clusterleader **150**. Rechts: Die erworbene Substanz **126**.

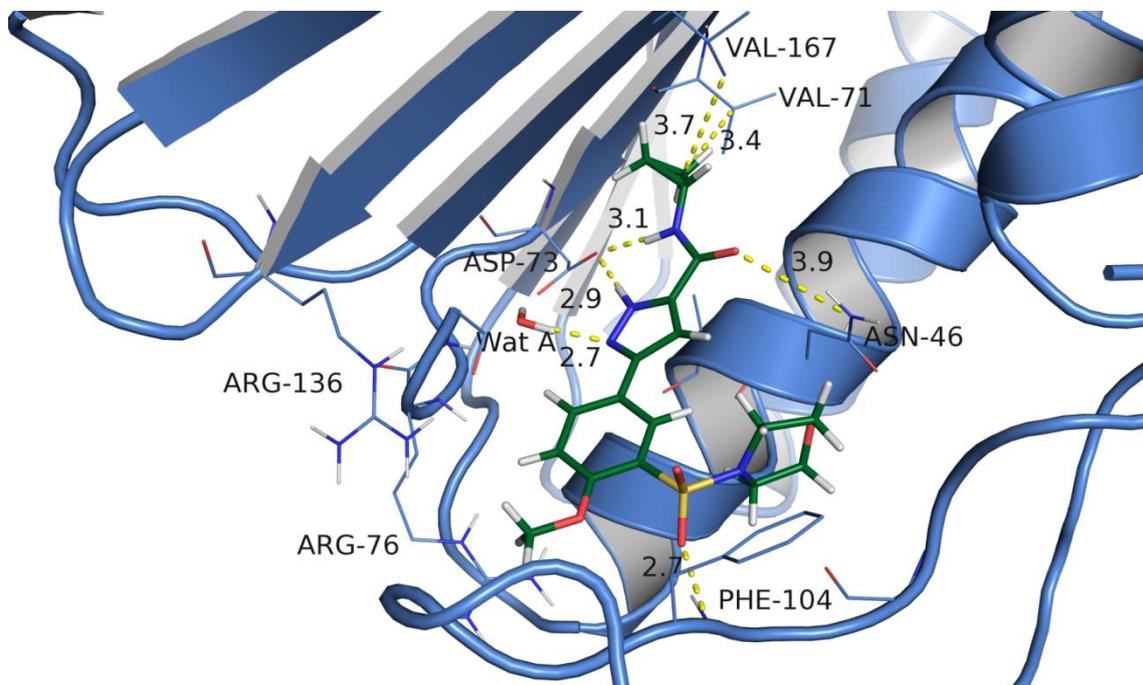


Abbildung 8.2 – Durch ein Glide-Docking erhaltene Pose von **126** in 3G7E_1. Zur Übersicht ist ein Abschnitt von Loop A ausgeblendet.

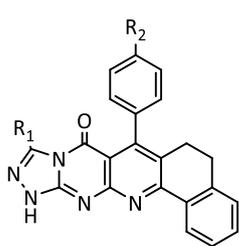
Die Interaktion mit dem Hauptpharmakophor fand über einen Pyrazol-Körper statt. Im W-Bereich befand sich der Stickstoff des Amids, der eine Wasserstoffbrücke mit Asp73 ausbildete. In Verlängerung trat noch ein Cyclopropyl-Rest auf, der den hydrophoben Charakter dieses Bereichs betonte. Der Sauerstoff des Amids befand sich knapp außerhalb der Reichweite für eine Wasserstoffbrücke für Asn46. Ein ähnliches Interaktionsmuster mit dem Hauptpharmakophor konnten RONKIN *et al.* bei N-Ethyl-Pyrazolamidderivaten identifizierten [88]. Jedoch unterschied sich die Struktur von den bisher beschriebenen Inhibitoren. Ein besonderes Merkmal stellte das Sulfonamid dar: Einerseits stabilisierte das Sulfonamid durch die Interaktion mit dem Backbone des Phe104 den Loop B, andererseits ermöglichte die Geometrie die Bindung des Morpholin-Ringes in der P-Tasche. Im A-Bereich der Tasche befand sich lediglich eine Methoxygruppe, die nicht die typischen Interaktionen mit den basischen Seitenketten im Ausgang der Bindetasche aufwies.

Für die weitere Optimierung zu einer Leitstruktur eignet sich der Cluster dieser Substanz nicht, da die Clustermitglieder zu große Reste für die E-Tasche besaßen. Die Optimierung für die Interaktion mit dem Hauptpharmakophor wurde schon durch weitergehende Untersuchungen abgedeckt (s. o.). Daher besteht der interessanteste Abschnitt im Sulfonamid und den Variationen des Morpholin-Ringes. Zusätzlich bietet die Substanz viele Veränderungsmöglichkeiten im A-Bereich, bei denen bei dieser Substanz nur unzureichend Interaktionen vorhanden waren.

– Substanz **149**

Bei der eigentlich bestellten Substanz **133** handelte es sich um **149** (s. Abschnitt 7.2 auf S. 188). Die fälschlicherweise erhaltene Substanz **149** weist jedoch eine sehr hohe Ähnlichkeit zu **133** auf. In Tabelle 8.1 ist der Cluster von **133** mit den entsprechenden Score-Werten und Platzierungen des PLP-Rescorings aus Abschnitt 0 auf S. 111 gezeigt. Obwohl die Substanz starr war und damit wenige Anpassungsmöglichkeiten an die Bindetasche besaß, konnte sie in alle Proteinstrukturen gedockt werden. Sogar die Clustermitglieder der Substanz wurden in fast jede Struktur, mit Ausnahme von 3G7E_2, erfolgreich gedockt und wiesen teilweise sehr gute PLP-Scores (< -115) auf. Dies verdeutlicht, dass das grundsätzliche Pharmakophor von **133** sehr erfolgsversprechend ist. **149** war sehr ähnlich zu **153**, das sich lediglich am R_2 -Substituenten Unterschied (Wasserstoff (**149**) zu Fluor (**153**)). **149** selbst war ebenfalls im ursprünglichen Datensatz enthalten, wurde aber nach dem ersten cRAISE-Docking herausgefiltert: Die beste Platzierung erreichte die Substanz im Docking in 3G7E_1 mit einem cRAISE-Score von -31.4 und einem Rank von 263,084.

Tabelle 8.1 – Übersicht des Clusters von **133**. Für jede Substanz ist der Unterschied t ($1 - T_k$) zu **133** angegeben und der Rank und Score vom Rescoring durch den PLP-Score aus dem Rescoring.

	Nr.	R ₁	R ₂	t	Struktur	PLP-Rank	PLP-Score
	133	NH ₂	H	–	1E11	7,904	-108.1
					1AJ6	250,434	-63.9
					3G7E_1	139	-122.9
					3G7E_2	93	-115.0
	151	NH ₂	Cl	0.01	1E11	3,041	-113.7
					1AJ6	67,795	-82.9
					3G7E_1	3,248	-108.7
					3G7E_2	180	-112.3
	152	H	H	0.05	1E11	1,395	-117.7
					1AJ6	42,391	-86.8
					3G7E_1	616	-116.2
					3G7E_2	–	–
	153	CH ₃	F	0.10	1E11	3,078	-113.7
					1AJ6	75,114	-81.9
					3G7E_1	522	-117.0
					–	–	–
	154	CH ₃	Cl	0.11	1E11	3,360	-113.2
					1AJ6	242,722	-64.7
					3G7E_1	4,076	-107.6
					3G7E_2	–	–

Zur Analyse der Pose wurde **149** im Nachhinein mit Glide entsprechend dem Vorgehen aus Abschnitt. Abschnitt 5.5 auf S. 120 erneut gedockt. Dabei wurde der gleiche Bindungsmodus erhalten, der auch bei **133** postuliert wurde (s. Abbildung 8.3).

Die Substanz **149** weist einige interessante Bindungseigenschaften auf. Das Hauptpharmakophor interagiert dabei mit einem anellierten Triazolopyrimidin. Diese Stoffklasse selbst ist zwar noch nicht beschrieben, das Interaktionsmuster mit Asp73 und Wat A, welches im weiteren Sinne von einem 7-Azaindol-Körper erfüllt wird, ist bekannt [73, 83]. Dennoch beinhaltet diese Struktur einige neue Eigenschaften. Das Stickstoffatom des Wasserstoffbrückenbindungsakzeptors von **149** befand sich einerseits in einer Guanidinpartialstruktur, deren Basizität durch den elektronenziehenden Effekt des Carbonyls leicht abgeschwächt war, andererseits war dieser Stickstoff auch Teil einer Amidinstruktur. Dadurch wurde der basische Charakter des Stickstoffs insgesamt verstärkt und begünstigte die Wasserstoffbrücke zum Wat A.

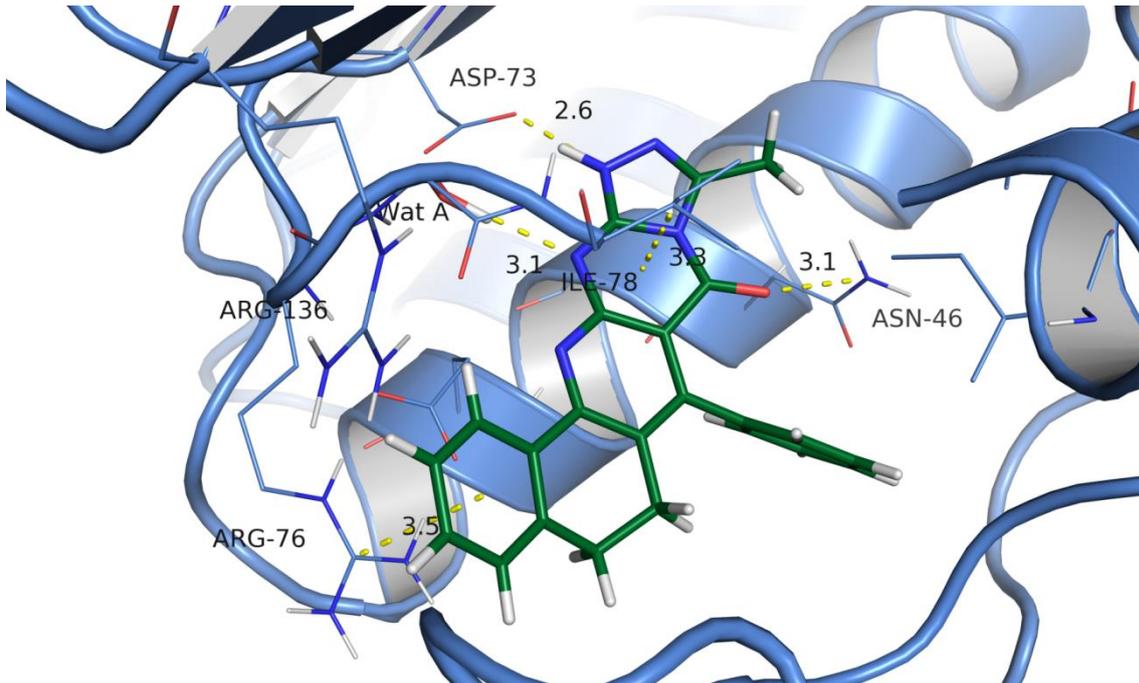


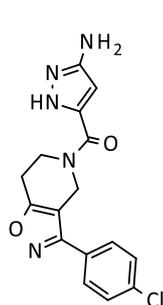
Abbildung 8.3 – Durch ein Glide-Docking erhaltene Pose von 149 in 3G7E_1. Zur Übersicht ist ein Abschnitt von Loop A ausgeblendet.

Der W-Bereich der Bindetasche wurde nicht besetzt. Daher ist unklar, ob das Wat B im Komplex vorhanden ist oder nicht. Mit dem Stickstoff der Triazol-Gruppe befindet sich zumindest ein Wasserstoffbrückenbindungsakzeptor in 2.7 Å Entfernung zur Position des Wat B. Ein weiterer bemerkenswerter Unterschied zu bekannten Komplexen bestand im Carbonyl-Sauerstoff am Ringsystem. Dieser befand sich in einer Entfernung von 3.1 Å zum Asn46. Bisherige Inhibitoren, wie bspw. der literaturbekannte Ligand in 3G7E, interagieren zwar mit Asn46, allerdings findet diese Interaktion in der Regel über die E-Tasche statt. Aus der Kristallstruktur von 1AJ6 ist bekannt, dass diese Aminosäure zumindest flexibel ist, da sich dort das Sauerstoffatom des Amids von der Seitenkette in Richtung der Tasche dreht. Möglicherweise ist dies auch für das Stickstoffatom des Amids möglich, so dass eine stabile Wasserstoffbrücke in Richtung der P-Tasche ausgebildet wird. Eine weitere entscheidende Interaktion findet mit Arg76 über ein π -Kationen-Stacking statt.

Damit bietet **149** hervorragende Optimierungsmöglichkeiten. Einerseits werden Schlüsselinteraktionen mit dem Protein ausgebildet: Wat A, Asp73, Asn46 und Arg76. Andererseits sind die weiteren Bereiche zum größten Teil von Kohlenwasserstoffen besetzt, die Möglichkeiten für Weiterentwicklungen bieten. Dabei sind vor allem der Phenylrest in der P-Tasche, die fehlende Interaktion zum Arg136 im A-Bereich oder die W-Tasche zu nennen. Weiterentwicklungen in diesen Bereichen würden auch Einfluss auf die bisher ungünstigen Eigenschaften, wie die Planarität und schlechten Lösungseigenschaften nehmen. Weiterhin begünstigte die gute synthetische Erschließbarkeit dieser Gruppe [267] und die Verfügbarkeit von den vorgestellten, sehr ähnlichen Derivaten (**151 – 154**) gute Ansatzpunkte für ein besseres Verständnis der Bindungspose und die Optimierung dieser Struktur.

– Substanz 140

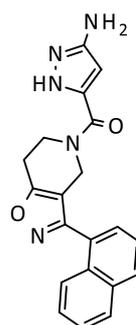
Bei **140** handelte es sich um einen Clusterleader. Der zugehörige Cluster bestand aus sechs weiteren Mitgliedern, die in ihren Bindungseigenschaften wahrscheinlich sehr heterogen waren. Die gemeinsame Komponente für die verschiedenen Strukturen war der Oxazolopiperidin-Körper (s. Abbildung 8.4). Für die Hauptinteraktion mit dem Asp73 und dem Wat A war bei **140** jedoch der Aminopyrazol-Rest verantwortlich. Dieser trat außer bei **140** nur bei **155** und **156** auf. **159** und **160** wiesen hingegen vollkommen andere Ringsysteme an dieser Stelle auf.



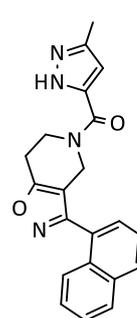
ZINC71415541
140
Leader



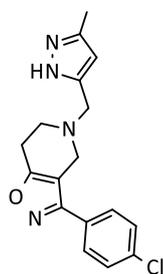
ZINC67869340
155
0.05



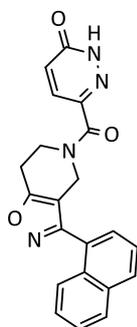
ZINC67654707
156
0.07



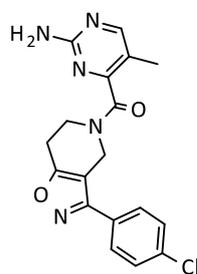
ZINC67642096
157
0.14



ZINC15064882
158
0.17



ZINC12708306
159
0.18



ZINC67947550
160
0.18

Abbildung 8.4 – Cluster von **140**. Für alle Substanzen sind der ZINC-Name und die Unterschiede (1 - Tanimotokoeffizient) zum Leader angegeben.

Für **140** war die Bestimmung der Pose am anspruchsvollsten. Die Substanz wurde aufgrund ihrer Pose in 3G7E_2 ausgewählt. Dabei war insbesondere die gute Bewertung durch das 3D-QSAR Modell und die interessante Interaktionsmuster entscheidend. Die Pose selbst muss in 3G7E_2 durchaus kritisch bewertet werden. Die Substanz wies Überschneidungen mit der Oberfläche des Enzyms auf und war gleichzeitig gestaucht (s. Abbildung 8.5).

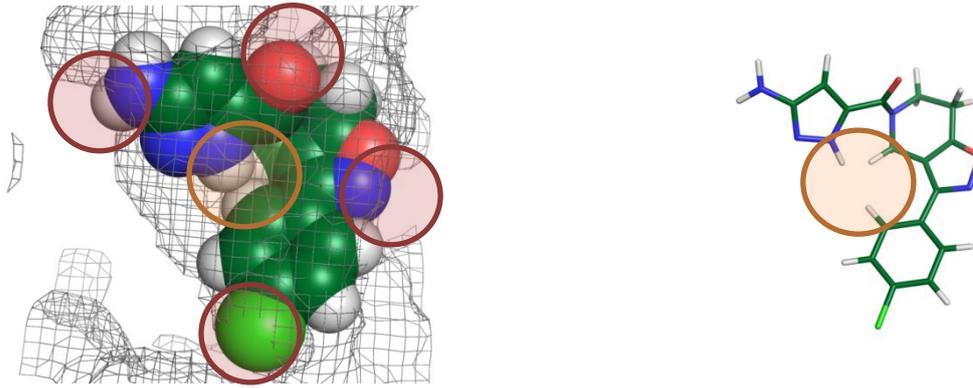


Abbildung 8.5 – Glide-Pose von **140** in 3G7E_2. Clashes mit dem Protein sind rot, interne Clashes orange eingekreist. **Links:** Die Proteinoberfläche ist als Gitter und **140** als Sphäre dargestellt. **Rechts:** Der Ligand ist als Stick in gleicher Orientierung und Perspektive wie bei der linken Darstellung gezeigt.

Nach der Verifizierung, dass es sich um einen Hit in der biologischen Testung handelte, wurden daher nochmal Glide-Dockings in den verschiedenen anderen Enzymstrukturen mit diesem Liganden durchgeführt. Die Prozedur war analog zum Postprocessing (s. Abschnitt 5.5 auf S. 120), mit dem Unterschied, dass pro Ligand bis zu zehn Posen herausgeschrieben wurden. Für **140** konnten dabei in den verschiedenen Strukturen sehr unterschiedliche Bindungsmodi gefunden werden. Generell traten in den Posen in den kleineren Taschen 3G7E_1 oder 3G7E_2 zumeist Clashes auf. Die von Glide bestbewertete Pose wurde im Docking mit 1E11 ermittelt: -8.314 kcal/mol (zum Vergleich, die ursprüngliche Pose in 3G7E_2 (s. Abbildung 8.5): -6.620 kcal/mol). Der Bindungsmodus war für ein Docking in 1E11 unüblich, da sowohl die A-Tasche als auch die P-Tasche komplett unbesetzt blieben (s. Abbildung 8.6). **140** band dafür im Bereich des Loop A, vergleichbar mit dem Isopropenyl-Rest vom Novobiocin in dieser Struktur. Dort befand sich der Fluorphenyl-Rest von **140** in einem sehr engen hydrophoben Interaktionsnetz. Auch für den Oxazol-Ring gab es mit Ile49 und Ser108 gute mögliche Interaktionspartner. Das Hauptpharmakophor interagiert mit dem Aminopyrazol-Körper von **140**. Aminopyrazole sind eine bisher noch nicht beschriebene Substanzklasse für diese Interaktion.

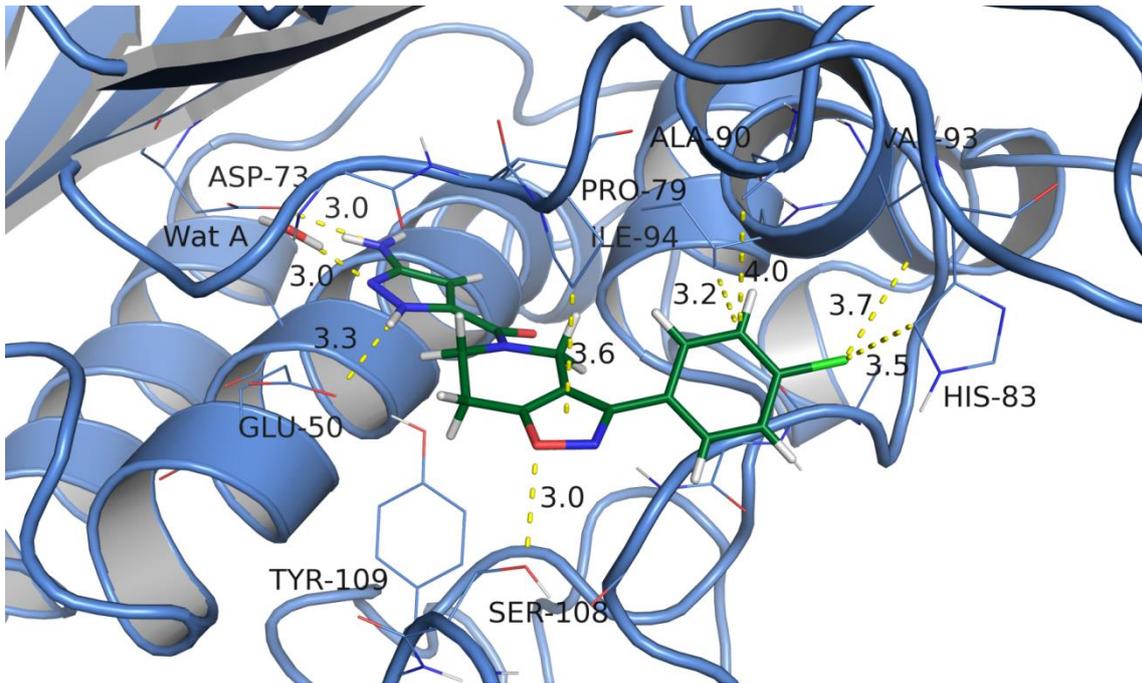


Abbildung 8.6 – Glide-Pose von **140** in 1EI1.

Beim Glide-Docking in 1AJ6 trat ebenfalls ein Bindungsmodus auf, der dem in Abbildung 8.6 ähnelte. Weiterhin gab es bei diesem Docking auch eine Pose, bei der **140** in Richtung des A-Bereichs band (s. Abbildung 8.7). Der Glide-Score war mit -6.315 kcal/mol jedoch deutlich schlechter.

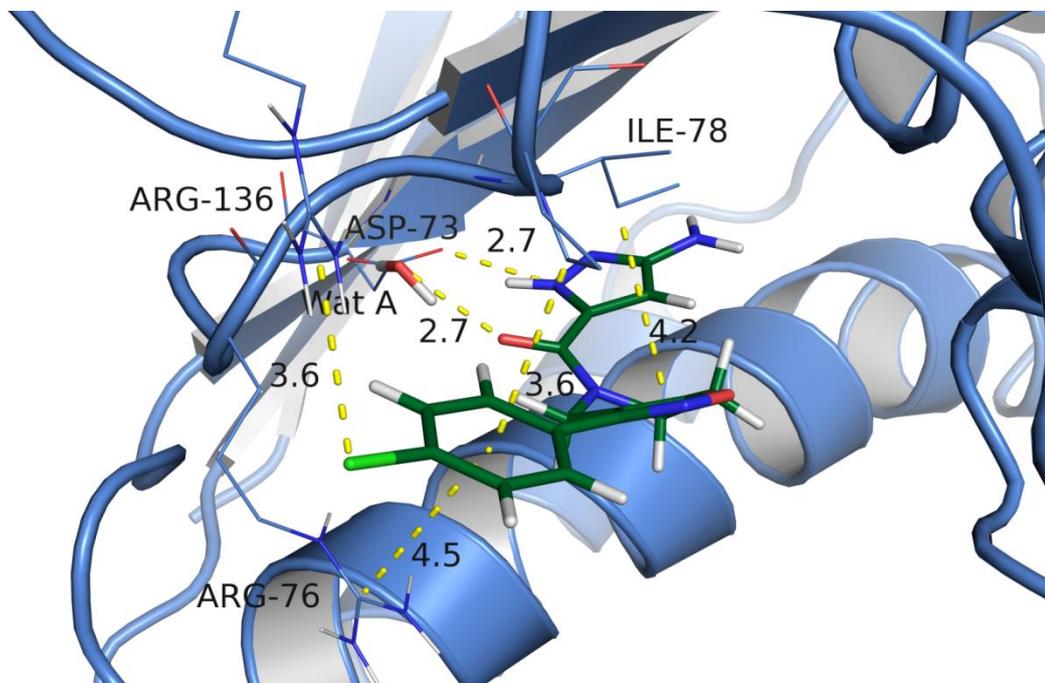


Abbildung 8.7 – Glide-Pose von **140** in 1AJ6.

Diese Pose ähnelte mehr den bisherigen Liganden in bekannten Kristallstrukturen, bei denen der A-Bereich vor der Interaktion mit dem A-Loop stattfindet (vgl. auch Novobiocin s. Abschnitt 2.1 auf S. 27). Für diese Pose sprachen die sehr kurzen und stabilen Interaktionen zum Wat A und Asp73 mit jeweils 2.7 Å. Weiterhin waren noch Interaktionen im W- und E-Bereich möglich. Positiv zu bewerten war auch hier das schon zuvor häufig vorkommende π -Kationen-Stacking mit dem Arg76.

Damit wies die Substanz **140** ebenfalls gute Eigenschaften für eine weitere Entwicklung auf. Im Vergleich zu **149** war der Cluster von **140** nicht so ergiebig, da die verschiedenen Clustermitglieder sich in dem Hauptpharmakophor zum Teil erheblich unterschieden. Für die Weiterentwicklung dieser Substanz wäre es zunächst erforderlich, den genauen Bindungsmodus zu ermitteln, der hier nicht abschließend geklärt werden konnte. Dafür wies **140** einige interessante Merkmale auf: Ein neues Fragment für die Hauptinteraktion mit Wat A und Asp73, mit, unabhängig von der Pose, Optimierungsmöglichkeiten im W- und E-Bereich. Der Piperidin-Ring des Moleküls wird voraussichtlich keinen großen Einfluss auf die Bindung nehmen und wäre daher variabel. Viel Raum zur Optimierung wird zusätzlich durch den P- und den A-Bereich geboten. Des Weiteren bietet der Oxazol-Ring in Abhängigkeit vom Bindungsmodus interessante Ansätze für eine Weiterentwicklung.

9 a Zusammenfassung

Die zunehmende Resistenzentwicklung von Bakterien gegenüber Antibiotika erfordert die Suche nach innovativen Wirkstoffen mit neuen Wirkmechanismen. Ziel dieser Arbeit war die Suche nach neuen Substanzklassen für die Hemmung der ATPase-Untereinheit der Gyrase von *Escherichia coli* (*E. coli*), die eine bekannte, aber nicht in der klinischen Anwendung befindliche Zielstruktur zur Behandlung von bakteriellen Infektionserkrankungen darstellt. Dafür wurde eine strukturbasierte Leitstruktursuche durchgeführt, deren Ablauf in Abbildung 9a.1 schematisch dargestellt ist:

Bei der Analyse der Zielstrukturen konnten drei Röntgenkristallstrukturen der GyrB-Untereinheit von *E. coli* identifiziert werden, die in der ATP-Bindestelle einzigartige Charakteristika aufwiesen. In allen Komplexen konnte ein strukturell essentielles Wassermolekül identifiziert werden. Zusätzlich wies einer der Protein-Ligand-Komplexe ein weiteres, optionales Wassermolekül auf. Damit standen drei vorbereitete Zielstrukturen mit einem Wassermolekül und eine Zielstruktur mit zwei Wassermolekülen für das virtuelle Screening zur Verfügung.

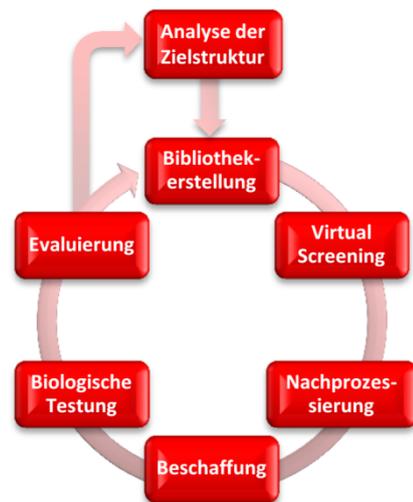


Abbildung 9a.1 – Schematisches Vorgehen der strukturbasierten Leitstruktursuche.

Die Substanzbibliothek wurde mithilfe der ZINC-Datenbank erstellt, einer vorgefilterten Datenbank mit kommerziell erhältlichen Substanzen. Für diese wurden weitere physikochemische Filter erarbeitet, die den Datensatz von ursprünglich 11 Mio. Substanzen auf 3.4 Mio. reduzierten. Die Bibliothek wurde mithilfe von cRAISE^{xv}, einem High-Throughput virtual Screening (HTvS)-Tool, in jedes der vier zuvor vorbereiteten Proteine gedockt. Dabei wurden zwei aufeinanderfolgenden Dockings mit steigender Komplexität durchgeführt. Insgesamt konnten von den 3.4 Mio. Substanzen ca. 360,000 Substanzen erfolgreich in den verschiedenen Proteinen positioniert werden.

Der Arbeitsablauf der darauf folgenden Nachprozessierung wurde zunächst mit einem Testdatensatz, der aus bekannten aktiven und automatisch generierten inaktiven Substanzen zusammengesetzt war, entworfen und evaluiert. Die endgültige Vorgehensweise bestand aus den folgenden Schritten: einem Rescoring, einem Leader-Clustering über Fingerprints und einem erneutem Docking mit Glide. Die final erhaltenen Posen von Glide wurden einer genauen visuellen Analyse unterzogen und führten zur Auswahl von 20 Substanzen, die nachfolgend erworben wurden.

^{xv} Das Screening mit cRAISE wurde im Arbeitskreis von Matthias Rarey aus dem Zentrum für Bioinformatik der Universität Hamburg durchgeführt.

Für die biologische Testung konnten zwei Assays etabliert werden: Für das primäre Screening wurde ein auf einer Fluoreszenzextinguierung basierender Assay entwickelt, der für die Gyrase noch nicht beschrieben wurde. Die Validierung der Hits erfolgte mithilfe eines literaturbekannten Supercoiling-Assays. Während der Testung zeigte jedoch keine der erworbenen Substanzen eine signifikante Inhibition der Gyrase.

Nach einer Evaluierung dieser Ergebnisse, der Analyse der verwendeten Bindungshypothesen und mithilfe von 3D-QSAR-Modellen^{xvi} wurden 23 weitere Substanzen aus dem Glide-Docking ausgewählt und erworben. Bei einer Konzentration von 500 μM wiesen zwei Substanzen eine vollständige und eine Substanz eine partielle Inhibition der Gyrase auf. Damit konnten drei neue kompetitive Inhibitoren der ATPase der Gyrase gefunden werden (s. Abbildung 9a.2). Diese Hit-Strukturen weisen alle neue Interaktionsmuster mit der Bindetasche der ATPase auf und können zu Leitstrukturen für neue Antibiotika weiterentwickelt werden.

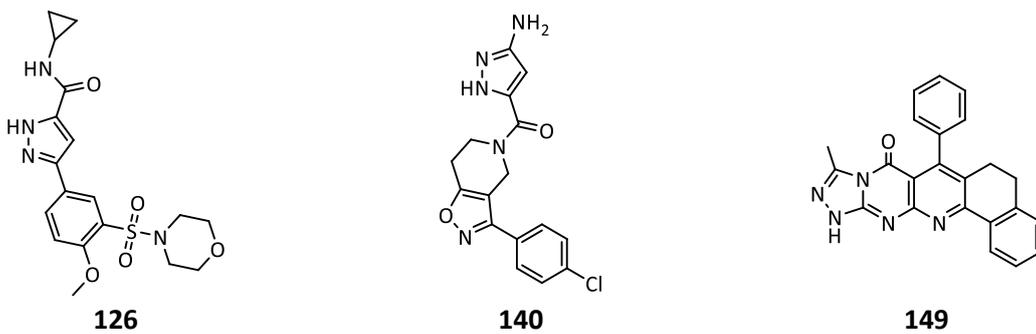


Abbildung 9a.2 –Hit-Strukturen der strukturbasierten Leitstruktursuche.

^{xvi} Die 3D-QSAR-Modelle wurde im Arbeitskreis von Thomas Lemcke im Institut für Pharmazie der Universität Hamburg entwickelt.

9 b Summary

Bacteria are a growing threat to mankind because acquired resistance to all known antibiotics is increasing permanently. Therefore, it is necessary to design novel and superior antibiotics with new modes of action. The aim of this work was to develop inhibitors for the ATPase of the gyrase of *Escherichia coli* (*E. coli*), which is a well-known target, but is currently not used in antibacterial therapy. The structure-based lead discovery procedure, which is shown schematically in figure 9b.1, was utilized for this task.

Analysis of the Protein Data Bank resulted in three different x-ray structures of *E. coli* with unique characteristics in the binding site. Within the binding pocket several conserved water molecules occurred. One water molecule could be identified as highly conserved and structurally important. Besides, there was an additional optional water molecule in one crystal structure. As a result, four different targets were chosen, which were used in the screening process: three targets with one water molecule and one additional target with two water molecules.

The compound library was created from the ZINC database, a prefiltered database for commercially-available compounds. This dataset was filtered using additional physicochemical properties to reduce the initial 11 million to 3.4 million compounds. The filtered compound library was subsequently docked utilizing cRAISE^{xvii}, a High-Throughput virtual Screening (HTvS) tool, into all four prepared targets. For this purpose, two consecutive steps with increasing complexity to place each compound within the binding site were used. Finally, 360,000 out of the 3.4 million substances were successfully docked.

The workflow for the following Postprocessing was developed and evaluated with a set of literature known actives and automatically created non-actives. The final workflow included the following steps: rescoring, leader clustering with fingerprints and redocking with Glide. Eventually, the poses from Glide were carefully visually inspected and 20 compounds were subsequently purchased.

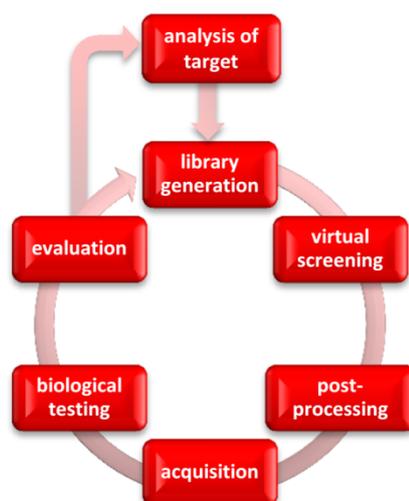


Figure 9b.1 – Schematic overview of the structure-based lead discovery.

^{xvii} The screening was performed by the workgroup of Matthias Rarey from the Center for Bioinformatics, University of Hamburg.

Two assays were established for biological testing: The primary screening was performed with a fluorescence quenching based assay, which has not been described for gyrase yet. A literature known supercoiling assay was additionally used to validate the initial hits. However, none of the 20 acquired compounds significantly inhibited the gyrase.

After an evaluation of these results, analysis of the used binding hypotheses and with the aid of 3D-QSAR models^{xviii} 23 additional substances were selected from the initial Glide docking and purchased. Two of these compounds completely inhibited the gyrase and one compound showed a partial inhibition at 500 μ M concentrations. Thus, three new competitive inhibitors of the ATPase of gyrase were identified (see figure 9b.2). These hits all exhibit new protein-ligand interaction patterns in the binding pocket of the ATPase and may be developed into lead compounds for new antibiotics.

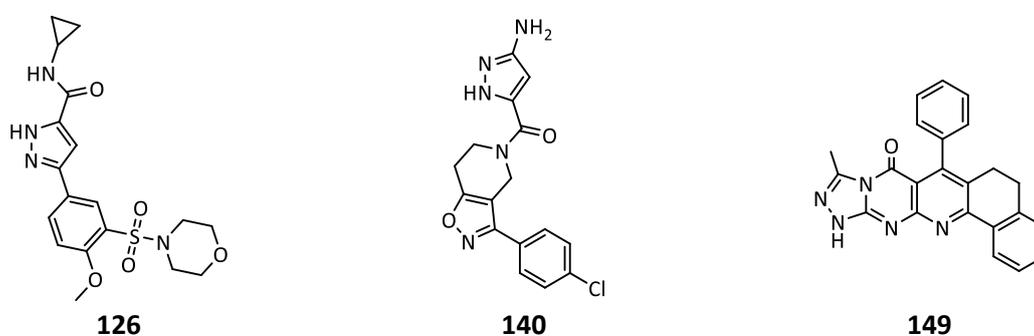


Figure 9b.2 – Hits of the structure based lead discovery.

^{xviii} The 3D-QSAR models were created by the workgroup of Thomas Lemcke from the Institute of Pharmacy, University of Hamburg.

10 Ausblick

Die Weiterentwicklung jeder Hit-Struktur erfordert zunächst die Synthese der jeweiligen Struktur und anschließend die Erstellung von Bibliotheken analoger (oder strukturell ähnlicher) Substanzen, um Struktur-Wirkungsbeziehungen (SAR) zu untersuchen. Dabei läuft eine Hit zur Leitstruktur-Optimierung (engl. *Hit to Lead*) ab. Für den Erfolg der Weiterentwicklung zu einem Antibiotikum liegen die Schwerpunkte in folgenden Bereichen: Bindungsaffinität, *in-vivo*-Aktivität und Selektivität.

– Bindungsaffinität

Die Bindungsaffinität beschreibt, wie stark eine Substanz an das Enzym bindet. In diesem Kontext beschreibt es damit die Stärke der Hemmung. Für eine antibiotisch wirkende Substanz ist eine sehr hohe Bindungsaffinität notwendig, da die Substanz nicht nur in den menschlichen Körper sondern auch noch in das Bakterium gelangen muss. Dementsprechend sind nur niedrige Konzentrationen in dem Bakterium selbst zu erwarten. Daher müssen Bindungsaffinitäten (IC_{50} -Werte) im nano- oder sogar subnanomolaren Bereich erreicht werden [268]. Die in dieser Arbeit gefundenen Substanzen sind im mikromolaren Bereich aktiv. Allerdings sind die Hit-Strukturen noch nicht optimal an die Bindetasche angepasst, da die Auswahl der Substanzen aus einer sehr heterogenen Datenbank erfolgte. Die Weiterentwicklung zu affineren Substanzen kann durch eine rationale strukturbasierte oder ligandenbasierte Optimierung erfolgen. Aufgrund der Docking-Posen gibt es Vorstellungen über den Bindungsmodus. Durch Erweiterung des Datensatzes und weitere *in vitro*-Testungen werden das Verständnis und die Optimierung weiter verbessert. Idealerweise kann dieser Prozess durch eine Strukturaufklärung, bspw. Röntgenkristallographie, unterstützt werden.

– *in-vivo* Aktivität

Neben der Untersuchung der Bindungsaffinität besteht ein weiterer wichtiger Schritt in dem Nachweis der *in vivo* Aktivität. Aufgrund der besonderen Barriereigenschaften für die Penetration in die Bakterienzelle – einerseits durch die Struktur der Zellwände und andererseits durch zahlreiche ein- und ausschleusende Transportproteine (z. B. Effluxpumpen) – ist die Vorhersage, ob die Substanz auch an den Wirkort gelangt sehr schwierig [95]. Aufgrund der Komplexität dieser Vorgänge ist die biologische Prüfung mit Bakterienzellen die sicherste und unkomplizierteste Methode zur Bestimmung der Aktivität. Erst mit diesem Test kann geklärt werden, ob und bei welchen Bakterienstämmen die Substanz an den Wirkort gelangen kann. Dies bestimmt letztendlich das Potential dieser Substanz als Antibiotikum.

– Selektivität

Die Selektivität gibt darüber Aufschluss, wieviel stärker ein Molekül an eine bestimmte Zielstruktur bindet als an eine andere. Ob eine selektive oder unselektive Bindung vorteilhaft ist, hängt von den zu untersuchenden Zielstrukturen ab. Die Betrachtung der Selektivität kann für kompetitive Inhibitoren der Gyrase auf drei Ebenen erfolgen: innerhalb des Bakteriums selbst, gegenüber verschiedenen Bakterien-Stämmen und gegenüber menschlichen Zellen.

Für die Selektivitätsbetrachtungen innerhalb des Bakteriums selbst ist vor allem die Hemmung von ähnlichen Zielstrukturen wichtig. Neben der Gyrase gibt es eine weitere, strukturell sehr ähnliche Topoisomerase, die DNA Topoisomerase IV (Topo IV). Die Hemmung zweier Zielstrukturen besitzt deutliche Vorteile gegenüber einer einfachen Hemmung. Dadurch nimmt nicht nur die Effizienz für die Therapie zu, sondern die Häufigkeit der Entwicklung von Resistenzen erheblich ab. Ein Nachteil von Novobiocin besteht darin, dass es die Gyrase selektiver als die Topo IV hemmt, wodurch eine schnellere Resistenzentwicklung begünstigt wird. Wird bei der Entwicklung von Gyrase-Hemmstoffen auf eine gleichmäßige Hemmung der Gyrase und der Topo IV geachtet, können im Vergleich zu Novobiocin bis über 100fach niedrigere Resistenzbildungsraten beobachtet werden [269]. In der *Protein Data Bank* (PDB) gibt es neben den Kristallstrukturen von der ATPase der GyrB-Untereinheit auch verschiedene Kristallstrukturen der ATPase der ParE-Untereinheit (die „B-Untereinheit“ der Topo IV). Mithilfe eines strukturbasierten Designs und der schon in der Literatur beschriebenen SAR-Eigenschaften kann eine Optimierung zu einem dualen Inhibitor stattfinden. Abschließend sind experimentelle Messungen der Bindungseigenschaften mit der Topo IV notwendig.

Bei der Betrachtung verschiedener Bakterienstämme ist ebenfalls eine unselektive Hemmung erwünscht. Nur wenn bei verschiedenen Stämmen die Gyrase (und ggf. Topo IV) gehemmt wird, besteht die Möglichkeit, einer breiten Anwendung. Diese Entwicklung wird dadurch begünstigt, dass die ATPase-Domäne in der Gyrase sehr stark konserviert ist. In der PDB-Datenbank sind ebenfalls verschiedene Kristallstrukturen von verschiedenen Stämmen zugänglich, die eine strukturbasierte Analyse ermöglichen. Entscheidend dafür sind jedoch ebenfalls die *in vivo*-Bestimmungen bei verschiedenen Stämmen (s. o.), da die Wirksamkeit insbesondere auch dadurch beeinflusst wird, ob die Substanzen die Zielstruktur erreichen. Dies ist insbesondere bei der Unterscheidung zwischen gram-positiven und gram-negativen Bakterien wichtig;

Bei humanen Enzymen wiederum wird eine sehr hohe Selektivität erfordert: Die gefundenen Substanzen sollen möglichst selektiv die bakterielle Gyrase hemmen, ohne humane Enzyme zu beeinflussen, um Nebenwirkungen zu vermeiden. Eventuelle Leitstrukturen sollten deswegen auf ihre Selektivität hin untersucht werden. Die größte Sequenzähnlichkeit weisen humane DNA Topoisomerasen vom Typ II auf. Weiterhin müssen aber auch Enzyme untersucht werden, die, wie die Gyrase, zur

GHKL-Klasse gehören. Von Novobiocin ist bspw. bekannt, dass es auch das *Heat Shock Protein 90* (HSP90) inhibieren kann [270]; erst kürzlich gelang mithilfe von Strukturvariationen von Novobiocin die Entwicklung neuer HSP90-Inhibitoren [271]. Dabei diente Novobiocin folglich als Leitstruktur für die Entwicklung von Inhibitoren des humanen Enzyms. Für diese Enzyme sind für erste Untersuchung als Anti-Target Röntgenkristallstrukturen in der PDB-Datenbank verfügbar. Letztlich muss aber auch hier die Selektivität experimentell bestimmt werden.

Der Fokus der Weiterentwicklung besteht zunächst darin, mehr Substanzderivate zu synthetisieren und weitere SAR-Informationen zu gewinnen. Die Erhöhung der Bindungsaffinität stellt das damit verbundene Ziel dar. Dennoch sollten die anderen hier diskutierten Parameter parallel untersucht und berücksichtigt werden. Insbesondere die *in-vivo*-Aktivität stellt bei der Entwicklung von Antibiotika eine hohe Hürde dar und sollte daher frühzeitig beachtet werden [268]. Methodisch können auch im weiteren Verlauf wieder strukturbasierte (z. B. Freien Bindungsenthalpie-Berechnungen) oder ligandenbasierte (z. B. 3D-QSAR-Modelle) Methoden eingesetzt werden. Ein kontinuierlicher Zyklus mit biologischen Experimenten ist aber dennoch notwendig.

11 Experimenteller Teil

11.1 Arbeitstechniken

11.1.1 Hochleistungs-Flüssigchromatografie (HPLC)

Die analytische Chromatographie erfolgte an einer 1220 Infinity LC der Firma Agilent, Waldbronn. Als Säule diente eine Poroshell 120 SB-C18 der Firma Agilent, Waldbronn (Poroshell 120, SB-C18, 4.6 x 150 mm, Partikelgröße 2.7 µm).

Es wurde Acetonitril des Reinheitsgrads HPLC (99.8 %) der Firma VWR Chemicals, Darmstadt, verwendet. Demineralisiertes Wasser wurde durch eine Millipore Mili-Q Filteranlage gereinigt.

Für die Reinheitsbestimmungen wurde der in Tabelle 11.1 aufgeführte Gradient verwendet. Alle untersuchten Substanzen waren basisch, so dass die Zugabe von Ameisensäure in die wässrige Phase notwendig war, um alle Moleküle vollständig zu protonieren und damit scharfe Peaks zu erhalten. Es wurde ein starker Gradient von 2 % auf 85 % organische Phase gewählt, um auch mögliche hydrophobe Verunreinigungen zu detektieren.

Tabelle 11.1 – Gradient-Protokoll für die Reinheitsbestimmung. A: Acetonitril; B: 0.5 % Ameisensäure in Wasser.

Zeit [min]	Anteil von A [%]	Anteil von B [%]
0 – 2	2	98
2 – 12	85	15
12 – 19	85	15
19 – 22	2	98
22 – 24	2	98

11.1.2 Elementaranalyse

Die Elementaranalysen für CHNS-O-Messungen wurden an einem C/H/N-Analysator vario EL III der Firma elementar, Hanau, und an einem EuroEA der Firma EuroVektor, Milan (Italien), durchgeführt.

11.1.3 Massenspektrometrie

ESI-Massenspektren wurden mit einem MicrOTOF-Q-Instrument der Firma Bruker, Bremen, aufgenommen. Die Proben wurden in einem Acetonitril/Wasser (0.5 % Ameisensäure)-Gemisch vermessen. Für die Aufnahme von HRMS-Spektren (*High Resolution Mass Spectrometry*, engl. für hochauflösende Massenspektrometrie) wurde Natriumformiat als Standard eingesetzt.

11.1.4 Weitere Geräte und Verbrauchsmaterialien

Tabelle 11.2 – Übersicht aller verwendeten Geräte und Verbrauchsmaterialien mit deren Bezugsquellen.

Geräte/ Verbrauchsmaterialien	Hersteller
Gel-Dokumentation, BioDoc Analyze Systeme	Analytik Jena, Jena
Genesys 10S UV-Vis Spectrophotometer	Thermo Scientific, Dreieich
Heizbad	Memmert, Schwabach
Horizon 58 Gelelektrophoresekammer	Biometra GmbH, Göttingen
Mikroliterpipette, Eppendorf Research plus (100 – 1000 µL)	Eppendorf, Hamburg
Mikroliterpipette, Finn Labsystems 4500 (0.5 – 10 µL, 5 – 40 µL, 40 – 200 µL)	Thermo Scientific, Dreieich
Mikrotiterplatte, Brandplates pure Grade, farblos	Brand GmbH + Co. KG, Wertheim
Mikrotiterplatte, Nunc Delta Surface, weiß	Thermo Scientific, Dreieich
Mithras LB 940 Multitechnologie Mikroplattenleser	Berthold Technologies, Bad Wildbad
pH-Meter E512	Metrohm, Filderstadt
Probenröhrchen 0,5 ml, Multiply-Pro	Sarstedt, Nümbrecht
Probenröhrchen 0.2 ml, Multiply-Pro	Sarstedt, Nümbrecht
Probenröhrchen 15 ml	Sarstedt, Nümbrecht
Probenröhrchen 15 ml , SafeSeal Micro Tubes	Sarstedt, Nümbrecht
Sterilfilter, Minisart 0.2 µm, 26 mm	Sartorius AG, Göttingen
Ultraschallbad, Sonorex Super	Bandelin, Berlin
Vortex-Genie 2, 230V	Scientific Industries, New York (USA)
Waage, Acculab Atilon	Acculab, Göttingen
Zentrifuge, Minispin Eppendorf	Eppendorf, Hamburg

Tabelle 11.3 – Übersicht aller verwendeten Chemikalien mit deren Bezugsquellen.

Chemikalie, Qualität	Hersteller
Acetonitril, HPLC Grade	VWR Chemicals, Darmstadt
Adenosintriphosphat (ATP)-Dinatriumsalz-Hydrat, BioXtra (≤ 0.1 % anorganisches Phosphat)	Sigma-Aldrich, Steinheim
Agarose, Bioreagenz	Sigma-Aldrich, Steinheim
Ameisensäure, Rectapur	VWR Chemicals, Darmstadt
Ammoniummolybdat-Tetrahydrat, zur Analyse	AppliChem, Darmstadt
Bovin Serum Albumin, für die Molekularbiologie	Carl Roth, Karlsruhe
Chinaldinrot, 95 %	Sigma-Aldrich, Steinheim
Citronensäure-Monohydrat, p. a.	Sigma-Aldrich, Steinheim
Dichlormethan, zur Analyse	Acros, Schwerte
Dimethylsulfoxid (DMSO), zur Analyse	VWR Chemicals, Darmstadt
Dithiothreitol, für die Molekularbiologie	VWR Chemicals, Darmstadt
DNA Gyrase (<i>E. coli</i>) + DNA Gyrase (<i>E. coli</i>) Reaction Buffer (Reaktionspuffer)	New England BioLabs, Frankfurt
DNA Gyrase (<i>E. coli</i>) Substrate (relaxiertes Plasmid)	New England BioLabs, Frankfurt
Eisessig, zur Analyse	VWR Chemicals, Darmstadt
Ethanol, zur Analyse	Sigma-Aldrich, Steinheim
Ethidiumbromid-Lösung, zur Fluoreszenz	Sigma-Aldrich, Steinheim
Ethylendiamintetraessigsäure (EDTA) –Dinatriumsalz-Dihydrat, for electrophoresis	Sigma-Aldrich, Steinheim
GeneRuler 1 kb DNA Ladder + 6x DNA Loading Dye (Ladungspuffer)	Thermo Fischer Scientific, Karlsruhe
Glycerol, Ph. Eur.	Merck, Darmstadt
Kaliumchlorid, zur Analyse	Merck, Darmstadt
Kaliumglutamat-Monohydrat, puriss. zur Analyse	Sigma-Aldrich, Steinheim
Magnesiumchlorid-Hexahydrat, Ph. Eur.	Merck, Darmstadt
Malachitgrün-Oxalat, für die Mikroskopie	VWR Chemicals, Darmstadt
Methanol, HPLC Grade	VWR Chemicals, Darmstadt
Natriumhydroxid, puriss.	Sigma-Aldrich, Steinheim
Novobiocin Natriumsalz, BC	AppliChem, Darmstadt
Polyvinylalkohol, zur argentometrischen Titration	Merck, Darmstadt
Salzsäure, rauchend	Th. Geyer, Renningen
Spermidin, BC	AppliChem, Darmstadt
Trikaliumphosphat, Reagenz	Sigma-Aldrich, Steinheim
Tris(hydroxymethyl)-aminomethan (TRIS), zur Analyse	Merck, Darmstadt

11.2 Arbeitsvorschriften

11.2.1 10x-Tris(hydroxymethyl)aminomethan-Essigsäure-EDTA-(TAE)-Puffer

Für die Herstellung des Agarosegels und die Durchführung der Gelelektrophorese wurde der TAE-Puffer benötigt. Dafür wurde ein 10x-TAE-Puffer angesetzt: 5.7 mL Eisessig wurden mit ca. 40 mL demineralisiertem Wasser gemischt. Nach Zugabe von 1.85 g Dinatrium-EDTA-Dihydrat und 24.2 g Tris(hydroxymethyl)aminomethan (Tris) und anschließendem Lösen aller Bestandteile wurde die Lösung auf 50.0 mL aufgefüllt. Die Lösung wurde bei Raumtemperatur gelagert. Für einen 1x-TAE-Puffer wurde die Lösung 1:10 verdünnt.

11.2.2 5x-Reaktionspuffer für Phosphomolybdat-Malachitgrün-Komplex-Assay (RP)

Für die Durchführung der auf den Phosphomolybdat-Malachitgrün-Komplex basierenden Assays wurde folgender Reaktionspuffer angesetzt (5x-RP): 1.06 g Tris, 0.2033 g Magnesiumchlorid-Hexahydrat, 0.4473 g Kaliumchlorid, 0.025 g bovines Serumalbumin und 14.806 g Glycerol wurden mit bidestilliertem Wasser auf ca. 40 mL aufgefüllt und gelöst. Mit Natronlauge oder Salzsäure wurde der pH-Wert auf 7.5 (bei 25 °C) eingestellt. Anschließend wurde der Reaktionspuffer mit bidestilliertem Wasser auf 50.0 mL aufgefüllt. Die Lösung wurde bei -20 °C gelagert. Für die Verwendung musste der Reaktionspuffer 1:5 verdünnt werden.

11.2.3 Farbreagenz

Das Farbreagenz für die auf den Phosphomolybdat-Malachitgrün-Komplex basierenden Assays wurde angelehnt an den Anweisungen von ROWLANDS *et al.* hergestellt [254]. Für das Farbreagenz wurden drei Lösungen vorbereitet: A) 57.12 mg Malachitgrünoxalat wurden in 50 mL bidestilliertem Wasser ($\cong 0,025$ mol/L) gelöst. B) 1.169 g Polyvinylalkohol wurden in 50 mL bidestilliertem Wasser ($\cong 2.32$ % (m/V)) gelöst. C) 4.94 g Ammoniummolybdat-Tetrahydrat wurden in 100 mL 6 M Salzsäure ($\cong 0.040$ mol/L) gelöst. Diese Lösungen wurden im Kühlschrank gelagert.

Das Farbreagenz wurde für jede Messreihe ca. 30 Minuten vor der Messung frisch hergestellt. Dafür wurden die Bestandteile im folgenden Verhältnis gemischt A:B:C:bidestilliertem Wasser: 2:1:1:2. Anschließend wurde die Lösung über einen Sterilfilter (Porengröße: 0.2 μ M) filtriert und bei Raumtemperatur stehen gelassen.

11.2.4 Weitere Lösungen

Zusätzlich wurden noch folgende Lösungen separat hergestellt:

– 60x-Dithiothreitol (DTT)-Lösung

38.6 mg DTT wurden in 2.0 mL bidestilliertem Wasser gelöst. Die Lösung wurde bei -20 °C gelagert. Für die Anwendung musste die Lösung 1:60 verdünnt werden.

– 25x-Spermidin-Lösung

18.2 mg Spermidin wurden in 1.0 mL bidestilliertem Wasser gelöst. Die Lösung wurde bei -20 °C gelagert. Für die Anwendung musste die Lösung 1:25 verdünnt werden.

– 25x-Adenosintriphosphat (ATP)-Lösung

21.34 mg ATP wurden in 1.5 mL bidestilliertem Wasser gelöst. Die Lösung wurde bei -20 °C gelagert. Für die Anwendung musste die Lösung 1:25 verdünnt werden.

– 10x-Kaliumglutamat-Lösung

6.0969 g Kaliumglutamat wurden in 10 mL bidestilliertem Wasser gelöst. Die Lösung wurde im Kühlschrank gelagert. Für die Anwendung musste die Lösung 1:10 verdünnt werden.

– 30 %ige Citronensäure-Lösung

2.77 g Citronensäuremonohydrat wurden in 10.0 mL bidestilliertem Wasser gelöst. Die Lösung wurde im Kühlschrank gelagert.

– Phosphatstandard-Reihe

Für die Phosphatstandard-Reihe wurde zunächst eine 1 M Phosphat-Lösung hergestellt. Dafür wurden 2.1226 g Trikaliumphosphat in 10.0 mL bidestilliertem Wasser gelöst. Für eine 1 mM Phosphat-Stammlösung wurden 50 µL der 1 M Phosphatlösung mit bidestilliertem Wasser auf 50.0 mL aufgefüllt. Ausgehend von dieser Stammlösung wurden die Phosphatstandards wie folgt angefertigt: Das entsprechende Volumen der Stammlösung (s. u.) wurde mit bidestilliertem Wasser auf 10.0 mL aufgefüllt, um den jeweiligen Phosphatstandard herzustellen. Die Lösungen wurden im Kühlschrank gelagert.

Phosphatstandard [µM]	Stammlösung [µL]	Phosphatstandard [µM]	Stammlösung [µL]
500	5000	50	500
400	4000	40	400
300	3000	30	300
200	2000	20	200
100	1000	10	100
80	800	5	50
60	600	1	10

11.2.5 Supercoiling Assay

Zunächst wurde zur Herstellung des Agarosegels eine 1 %ige Agarosesuspension in 1x-TAE-Puffer in der Mikrowelle ca. 3 Minuten erhitzt, bis sich eine klare Lösung bildete. Diese wurde nach einer kurzen Abkühlphase in die Gelelektrophoresekammer gegossen und bis zum Erstarren des Gels abgekühlt.

Der Supercoiling-Assay wurde nach dem Herstellerprotokoll durchgeführt. Das Reaktionsvolumen betrug 30 μL . Der 1x-Reaktionspuffer des Herstellers war folgendermaßen zusammengesetzt: 35 mM Tris-Hydrochlorid, 24 mM Kaliumchlorid, 4 mM Magnesiumchlorid, 2 mM DTT, 1.75 mM ATP, 5 mM Spermidin, 0.1 mg/ml bovines Serumalbumin und 6.5 % Glycerol. Weiterhin wurden 1.2 μL DMSO (ggf. mit Inhibitor oder zu testender Substanz) hinzugegeben. Anschließend wurde 1 μL (\cong 5 U) Gyrase hinzupipettiert. Der Ansatz wurde nach Zugabe von 0.5 μL (\cong 500 ng) relaxierten DNA-Plasmid für 30 Minuten bei 37 °C inkubiert. Die Reaktion wurde durch Einfrieren der Probe gestoppt.

Für die Gelelektrophorese wurden 10 μL des Enzymansatzes mit 2 μL 6x-Ladungspuffer gemischt. 10 μL dieser Lösung wurden schließlich in dem zuvor vorbereiteten Agarosegel bei 100 V für ca. 50 Minuten aufgetrennt. Als Elektrolytlösung für die Elektrophorese wurde 1x-TAE-Puffer verwendet. Nach Beendigung der Elektrophorese wurde das Gel für 30 Minuten in einem Ethidiumbromidbad (10 ppm in 0.5x-TAE-Puffer) und für 10 Minuten in 0.5x-TAE-Puffer gewaschen und anschließend dokumentiert. Zur Zuordnung der Banden wurde der GeneRuler 1 kb DNA Ladder verwendet, wenn nicht anders aufgeführt.

11.2.6 Phosphomolybdat-Malachitgrün-Komplex Assay

Im Verlauf dieser Arbeit wurde die Methode kontinuierlich optimiert. Die hier beschriebene Vorgehensweise beschreibt die final erarbeitete Prozedur. Diese wurde auch für die biologische Testung (s. Abschnitt 7 auf S. 183) verwendet.

Das Gesamt-Reaktionsvolumen betrug 30 μL . Es wurde, in Abhängigkeit von der Größe der Messreihe, ein Mehrfachansatz von folgendem 1x-EnzymMix hergestellt: 6 μL 5x-RP, 0.5 μL 60x-DTT-Lösung, 1.2 μL 25x-Spermidin-Lösung, 3 μL 10x-Kaliumglutamat-Lösung, 13.9 μL bidestilliertes Wasser und 1 μL (\cong 5 U) Gyrase (vorsichtig mischen). Parallel wurde ein Mehrfachansatz von folgendem Starter-Mix hergestellt (1x): 1.2 μL 25x-ATP-Lösung und 2 μL DNA-Ladder. Beide Ansätze wurden auf 37 °C erwärmt und für ca. 10 Minuten equilibriert. 25.6 μL vom EnzymMix wurde in die Reaktionsgefäße verteilt und mit jeweils 1.2 μL DMSO, das gegebenenfalls die Inhibitoren oder zu testenden Substanzen enthielt, versetzt.

Anschließend wurden 3.2 μL des StarterMixes zu dieser Lösung hinzugegeben, um die Reaktion zu starten. Die Inkubation fand bei 37 °C im Wasserbad statt. Die Inkubationszeit richtete sich nach der

Aktivität des Enzyms, die abhängig von der Charge und der Lagerungszeit variieren konnte. Sie musste daher für jede Messung angepasst werden und betrug zwischen 30 und 180 Minuten.

Durch Zugabe von 96 µL Farbreagenz wurde die Reaktion gestoppt. Nach ca. 10 Minuten Reaktionszeit wurde die weitere Farbreaktion mit 12 µL einer 30 %igen Citronensäure-Lösung unterbunden.

Abhängig davon, ob die Absorption oder die Fluoreszenzauslöschung gemessen wurde, gab es zwei verschiedene Vorgehensweisen:

a. Absorptionsmessung

Je 115 µL wurden in einen Well einer durchsichtige Mikrotiterplatte pipettiert. Anschließend wurde die Absorption im Plate Reader bei einer Wellenlänge von 620 nm vermessen.

b. Fluoreszenzauslöschungsmessung

Zunächst wurde die Fluoreszenz, Anregungswellenlänge: 430 nm/Emissionswellenlänge: 568 nm, am Plate Reader für jedes Well einer leeren, weißen Mikrotiterplatte vermessen und, entsprechend ihrer Fluoreszenzeigenschaften, die zu verwendenden Wells ausgewählt (s. Abschnitt 6.4 auf S. 172). Um Schwankungen bei der Fluoreszenz vorzubeugen, fand die Messung unmittelbar vor der eigentlichen Bestimmung statt. Jeweils 115 µL der fertigen Reaktionslösung wurden blasenfrei in die ausgewählten Wells der Mikrotiterplatte übertragen. Anschließend wurde die Fluoreszenz erneut mit dem Plate Reader mit denselben Einstellungen, wie bei der leeren Platte, gemessen.

Für jede Messreihe wurde ebenfalls die Phosphatstandard-Reihe zur Bestimmung der Phosphatkonzentration vermessen. Die Zusammensetzung war weitestgehend identisch zum Enzymansatz. Aus Kostengründen waren in diesem jedoch kein DNA-Ladder und keine Gyrase enthalten. Er setzte sich daher folgendermaßen zusammen (1x): 15 µL Phosphatstandard, 6 µL 5x-RP, 0.5 µL 60x-DTT-Lösung, 1.2 µL 25x-Spermidin-Lösung, 3 µL 10x-Kaliumglutamat-Lösung, 1.2 µL 25x-ATP-Lösung, 1.2 µL DMSO und 1.9 µL bidestilliertes Wasser. Das weitere Vorgehen war identisch zum Enzymansatz.

11.2.7 Auswertung der Fluoreszenzauslöschung

Für die Auswertung der Fluoreszenzauslöschung wird für jeden Well die berechnete Absorption A_{calc} bestimmt (s. Abschnitt 6.2.3 auf S. 157). Dafür wird die Fluoreszenz des leeren Wells (E_{leer}) und die Fluoreszenz des Wells mit dem Farbreagenz (E_{mess}) verwendet. Sie berechnet sich damit wie folgt (Formel 11.1):

$$A_{calc} = -\lg\left(\frac{E_{mess}}{E_{leer}}\right) \quad \text{Formel 11.1}$$

Mithilfe der Regressionsgeraden, die aus den berechneten Absorptionswerten der Phosphatstandardreihe erstellt wurde, konnte die freigesetzte Menge an Phosphat für die Enzymreaktion be-

stimmt werden. Für die Bestimmung der IC_{50} -Werte wurde die berechnete Absorption der Enzymmessung mit dem Inhibitor bei verschiedenen Konzentrationen verwendet.

11.3 Computermethoden

11.3.1 Darstellung dreidimensionaler Molekülstrukturen

Zur Darstellung von dreidimensionalen Strukturen wurde in dieser Arbeit PyMOL [98] verwendet. Wenn nicht anders aufgeführt, wurden folgende Darstellungen gewählt: Das Enzym in seiner Cartoon-Darstellung und seine Kohlenstoffatome wurden blau gefärbt. Bei der Darstellung der Seitenketten als Sticks oder Lines wurden unpolare Wasserstoffe ausgeblendet. Die Liganden wurden mit allen Wasserstoffatomen als Sticks dargestellt, wobei die Kohlenstoffatome der Liganden grün eingefärbt wurden.

11.3.2 Proteinvorbereitung

Der Ablauf der Vorbereitung der Proteinstrukturen war für alle Enzyme identisch. Die Vorbereitung wurde mit dem *Protein Preparation Wizard* von Schrödinger [272] durchgeführt: Im ersten Schritt wurde die Kette A der entsprechenden Kristallstruktur geladen und Ionen und Kristallisationsfragmente (bspw. Glycerol), die keinen Einfluss auf die Bindung der Liganden besaßen, entfernt. Anschließend wurde die Bindungsordnung für die Liganden manuell angepasst. Der *Protein Preparation Wizard* bestimmte die Bindungsordnung für die Aminosäuren und fügte dem entsprechend Wasserstoffatome hinzu. Alle Wassermoleküle, die sich mehr als 10 Å von dem Liganden entfernt befanden, wurden entfernt. Der *Protein Preparation Wizard* optimierte anschließend das gesamte Wasserstoffbrückennetzwerk einschließlich der Wassermoleküle, um die richtigen Tautomere (z. B. bei Histidin), Rotationen (Asparagin) und Ladungen (Glutamat) zu bestimmen. Dabei wurde von einem neutralen pH-Wert ausgegangen. Die optimierte Struktur wurde im Bereich der Bindetasche visuell auf ihre Plausibilität hin überprüft. Zum Abschluss wurden die Wasserstoffatome mithilfe des OPLS_2005-Kraftfeld [103] minimiert.

11.3.3 Loop Refinement

Für das *Refinements* des Loops wurden alle Seitenketten optimiert, die sich innerhalb eines Radius von 7.5 Å um den Loop befanden. Für die externe Dielektrizitätskonstante wurde 80 gewählt. Als minimale Distanz zwischen zwei Atomen wurde der 0.7fache vdW-Radius zugelassen. Es wurden bis zu fünf verschiedene Loops generiert. Für Loop A wurde das *Default Sampling*, für Loop B das *Ultra Extended Sampling* verwendet.

Input-Datei für Loop B:

```

STRUCT_FILE      Input.mae
JOB_TYPE         REFINE
PRIME_TYPE       4
LOOP_0_RES_0    A:97
LOOP_0_RES_1    A:119
MAX_JOBS         4
NUM_OUTPUT_STRUCT 5
RES_SPHERE       7.50
MIN_OVERLAP     0.70
USE_CRYSTAL_SYMMETRY no
USE_RANDOM_SEED  no
SEED             0
EXT_DIEL         80.00
USE_MEMBRANE     no
HOST             localhost

```

11.3.4 MD-Simulationen

Die MD-Simulationen wurden mit AMBER11 durchgeführt [111]. Vorbereitet wurden die Strukturen mit Antechamber [273]. Für die ionischen Liganden (Novobiocin in 1AJ6 und ANP in 1E11) wurde der R.E.D.-Server verwendet, um die Ladungen zu berechnen [274, 275]. Dafür wurden RESP A1A-Ladungen unter Verwendung von Gaussian auf dem HF/6-31G* Level berechnet [276, 277]. Für den Liganden von 3G7E wurden AM1-BCC-Ladungen verwendet [278, 279]. Das „general Amber force field“ (GAFF) wurde auf die Liganden [113] und das ff03SB Kraftfeld auf das Protein angewendet [114]. Die Ladungen wurden durch das Hinzufügen von Natriumionen neutralisiert und ein abgeflachtes Oktaeder wurde in einem 8 Å-Radius um den Komplex mit TIP3P-Wasser gefüllt [280].

Für die folgenden Simulationen betrug die Höchstgrenze für nicht-bindende Interaktion 8 Å. Für weitreichende elektrostatische Interaktionen wurde die Particle Mesh Ewald (PME) Methode verwendet [281]. Nach der Vorbereitung wurde die Lösung in 1000 Schritten minimiert, wobei für 500 Schritte die *Deepest Decent*-Methode und für 500 Schritte die *Conjugate Gradient*-Methode verwendet wurde. Während der Minimierung wurden alle Atome von Protein und Ligand mit einer Kraftkonstante von 15 kcal/mol Å beschränkt. Die gleiche Prozedur wurde ohne jeden Restraint wiederholt. Im Folgenden wurde der Komplex mit dem Langevin-Thermostat bei einer Kollisionsfrequenz von 2 ps⁻¹ innerhalb von 50 ps von 0 K auf 300 K aufgeheizt. Dies passierte unter periodischen Randbedingungen (engl. *Periodic Boundary Conditions*, PBC), bei denen das Volumen konstant gehalten wurde, so genannte NVT-Bedingungen: die Anzahl der Atome (N), das Volumen (V) und die Temperatur (T) sind festgelegt, der Druck (P) variabel. Dabei wurden der Ligand und Protein mit leichten Restraints in Position gehalten (5 kcal/mol Å). Zusätzlich wurde ab diesem Zeitpunkt der SHAKE-Algorithmus angewendet, um die Valenzschwingung von Wasserstoffatomen einzuschränken [112, 282]. Anschließend fand eine kurze Dichteequilibrierung über 50 ps statt, während der nur der Ligand mit den glei-

chen Restraints (5 kcal/mol Å) belegt war. Die PBC wurden dabei auf einen konstanten Druck (NPT-Bedingung) von 1 bar festgelegt.

Im Anschluss wurden MD-Simulationen ohne jede weitere Beschränkung durchgeführt.

Bash-Skript für MD-Simulation von 1AJ6: *# Forschungscluster RRZ Universität Hamburg*

```
#!/bin/bash

#PBS -V
#PBS -l nodes=1:ppn=12

source /ddn/env/intel/compiler/12.0
source /ddn/env/openmpi/intel/1.4.4
source /ddn/env/mvaphich2/1.8alp1/gnu+cuda

export AMBERHOME=/ddn/sw/amber/11/multigpu/gnu+cuda+mvaphich1.8alp1/amber11

# Minimierung mit Restraints
mpirun -np 12 $AMBERHOME/exe/sander.MPI -O -inf min1.info -i min1.in \
-o min1.out -p 1aj6.prmtop -c 1aj6.inpcrd -r min1.rst -ref 1aj6.inpcrd

# Minimierung ohne Restraints
mpirun -np 12 $AMBERHOME/exe/sander.MPI -O -inf min2.info -i min2.in \
-o min2.out -p 1aj6.prmtop -c min1.rst -r min2.rst

# Aufheizen
mpirun -np 3 $AMBERHOME/exe/pmemd.cuda_SPDP.MPI -O -inf heat.info \
-i heat.in -o heat.out -p 1aj6.prmtop -c min2.rst -r heat.rst \
-ref min2.rst -x heat.mdcrd

# Equilibrierung mit Restraint auf Liganden
mpirun -np 3 $AMBERHOME/exe/pmemd.cuda_SPDP.MPI -O -inf equi.info \
-i equi.in -o equi.out -p 1aj6.prmtop -c heat.rst -r equi.rst \
-x equi.mdcrd ref heat.rst

# MD
mpirun -np 3 $AMBERHOME/exe/pmemd.cuda_SPDP.MPI -O -inf mdl.info \
-i mdl.in -o mdl.out -p 1aj6.prmtop -c equi.rst -r mdl.rst -x mdl.mdcrd
```

Mit folgenden Input-Dateien:

<p><i>min1.in:</i></p> <p>Minimierung der Lösung</p> <pre>&cntrl imin = 1, maxcyc = 1000, ncyc = 500, ntb = 1, ntr = 1, cut = 8.0 /</pre> <p>Restraints auf Protein und Ligand</p> <pre>15.0 RES 1 207 END END</pre>	<p><i>min2.in:</i></p> <p>Minimierung der gesamten Struktur</p> <pre>&cntrl imin = 1, maxcyc = 1000, ncyc = 500, ntb = 1, ntr = 0, cut = 8.0 END</pre>
--	---

*heat.in:***50 ps Aufheizen**

```

&cntrl
  ig      = -1
  imin    = 0,
  irest   = 0,
  ntx     = 1,
  ntb     = 1,
  cut     = 8.0,
  ntr     = 1,
  ntc     = 2,
  ntf     = 2,
  tempi    = 0.0,
  temp0   = 300.0,
  ntt     = 3, gamma_ln = 2.0,
  nstlim  = 25000, dt = 0.002,
  ntpr    = 1000, ntwx = 1000,
  ntwr    = 10000

```

/

Restraints auf Protein und Ligand

```

5.0
RES 1 207
END
END

```

*equi.in:***50 ps Equilibrierung**

```

&cntrl
  ig      = -1
  imin    = 0,
  irest   = 1,
  ntx     = 7, ntb = 2,
  pres0   = 1.0, ntp = 1, taup = 2.0,
  cut     = 10.0,
  ntr     = 1,
  ntc     = 2,
  ntf     = 2,
  tempi    = 300.0,
  temp0   = 300.0,
  ntt     = 3, gamma_ln = 2.0,
  nstlim  = 25000, dt = 0.002,
  ntpr    = 1000, ntwx = 1000,
  ntwr    = 10000

```

/

Restraints auf Ligand

```

5.0
RES 207
END
END

```

*md1.in:***1 ns MD**

```

&cntrl
  ig      = -1
  imin    = 0,
  irest   = 1,
  ntx     = 7,
  ntb     = 2, pres0 = 1.0, ntp = 1, taup = 2.0,
  cut     = 10.0,
  ntr     = 1,
  ntc     = 2,
  ntf     = 2,
  tempi    = 300.0,
  temp0   = 300.0,
  ntt     = 3, gamma_ln = 2.0,
  nstlim  = 25000, dt = 0.002,
  ntpr    = 1000, ntwx = 1000, ntwr = 10000

```

END

11.3.5 Minimierung

Für die Vorbereitung der Liganden und Proteine wurde analog zu den MD-Simulationen (s. Abschnitt 11.3.4) vorgegangen: Nach Berechnung der Ladung wurde auf die Ligandenatome das GAFF-Kraftfeld und auf die Proteinatome das ff03SB Kraftfeld angewendet. Bei allen durchgeführten Minimierungen war die Ladung bereits ausgeglichen und die Komplexe enthielten in einer definierten Box explizites Wasser, für welches wiederum das TIP3P-Modell verwendet wurde (s. o.). Die Komplexe wurden anschließend in 1000 Schritten minimiert, wobei für 500 Schritte die *Deepest Decent*-Methode und für 500 Schritte die *Conjugate Gradient*-Methode verwendet wurde. Die Minimierung wurde bereits vorher gestoppt, wenn das Konvergenzkriterium, eine kleinere Energiedifferenz als $1.0 \cdot 10^{-4}$ kcal/mol Å, erreicht wurde.

Bash-Skript für Minimierung von 1AJ6: *# HPC-Cluster RRZ Universität Hamburg*

```
#!/bin/bash

#PBS -V
#PBS -l nodes=1:ppn=12

# Minimierung ohne Restraints
mpirun -np 12 $AMBERHOME/exe/sander.MPI -O -inf min3.info -i min3.in -o
min3.out -p 1aj6.prmtop -c md1.rst -r min3.rst
```

min3.in:

Minimierung der gesamten Struktur

```
&cntrl
  imin = 1,
  maxcyc = 1000,
  ncyc = 500,
  ntb = 1,
  ntr = 0,
  cut = 8.0,
  ntx = 5,
```

/

11.3.6 SuperStar

Es wurde die SuperStar Version 2.0.3 verwendet. Als Sonde wurde die „Water Oxygen“ und als Datenquelle die CSD-Datenbank gewählt. Für jedes Protein wurden alle Aminosäuren in 5 Å Radius um den Liganden analysiert.

Input-Datei für SuperStar-Berechnungen:

```
JOBNAME           Jobname
PROBENAME         WATER OXYGEN
GRID_SPACING      0.7
CALC_CONTOUR_SURFACE 1
SAVE_SUPER_SCATTERPLOT 1
RUN_TYPE          SUPERSTAR
DATA_SOURCE       CSD
SHELL_VALUES      0.5 0
SIGCHECK_METHOD   POISSON SHELL
PROPENSITY_CORRECTION LOGP DEFAULT DEFAULT
MIN_CAVITY_VOLUME 10
CAVITY_RADIUS     10
DATA_TYPE         ISOSTAR
PEAK_FITTING      1
PEAK_FITTING_NCYCLES 1
MIN_PEAK_HEIGHT   1.99
PEAK_FITTING_REFINE 0
MOLECULE_FILE     Pfad zu Protein
CAVITY_DETECTION  0
SUBSTRUCTURE      Aminosäuren in 5 Å Entfernung um den Liganden
```

11.3.7 WaterClassifier

Die Bewertung mit WaterClassifier erfordert eine vorbereitete Protein-Datei von AutoDock Vina (PDBQT-Format) [136]. Das Skript wurde in persönlicher Kommunikation mit Greg Ross aus dem Arbeitskreis von Philip Biggin^{xix} erhalten.

```
>$ ScoreWaters.py -p Protein -w Wasser # Protein: PDBQT-Datei, Wasser: PDB-Datei
```

11.3.8 HINT und Rank

Die Bewertung der Wassermoleküle fand mithilfe eines Webservers^{xx} statt. Für diesen wurden die entsprechenden PDB-Dateien der Strukturen benötigt.

^{xix} Structural Bioinformatics and Computational Biochemistry Unit – Dpt. of Biochemistry – University of Oxford.

^{xx} <http://hinttools.isbdd.vcu.edu/WR/index.html>; zuletzt online verfügbar: März 2012.

11.3.9 DSX-Score

Für die Bewertung der Dockingergebnisse wurde der DSX-Score der Version 0.89 mit den „*pdb_pot_pharmacophoric_0511*“-Potentials verwendet.

```
>$ dsx_rhel_linux_32.lnx -P Protein -W Wasser -L Posen -D pdb_pot_pharmacophoric_0511
# Protein/Wasser: PDB-Datei, Posen: mol2-Datei
```

11.3.10 PLP-Rescoring

Für die Nutzung des PLP-Scores in LeadIT (Version 2.1.5) mussten in den „*Global Preferences*“ für die „*Geometry*“-Eigenschaften die Parameter des PLP-Scores geladen werden. Diese sind im Installationspfad von LeadIT unter: *./example/static_data/geometry_plp_score.dat* zu finden.

Für das Rescoring wurde eine Projektdatei mit dem vorbereiteten Enzym benötigt. Das Rescoring fand im Anschluss über den Befehlszeilenmodus von LeadIT statt:

```
>$ leadit -b rescoring.bat
```

rescoring.bat:

```
PROJECT
  READ projektname.fxx      # Projekt mit vorbereitetem Enzym
  END
for_each $(lig) FROMTO 1 x  # x ist die Anzahl der Substanzen in der mol2
  LIGAND
    READ mol2 $(lig) no     # mol2 mit den zu bewertenden Liganden
  END
  DOCKING
    ANALYZE                 # Durchführung des Rescorings
    seloutp output.txt a   # Output-Datei
    OUTPUT "$ (lig) "
    PLP fix n
    seloutp screen
end_for
QUIT y
```

11.3.11 Prime MM-GBSA-Score

Zur Berechnung des Prime MM-GBSA-Scores wurde die Prime Version 1.41 (Schrödinger [102]) mit dem VSGB 2.0 Lösungsmodell genutzt. Um den Einfluss der Minimierung auf den Komplex zu bestimmen, wurde jeweils eine Berechnung ohne Minimierung des Proteins (GBSA freezed genannt) und eine weitere mit Minimierung in einem 8 Å Radius um den Liganden (GBSA 8 Å) durchgeführt. Als Sampling-Methode wurde die Minimierung gewählt.

11.3.12 Fingerprints

Mithilfe des KNIME-Tools wurden 2048 Bit lange Daylight-Fingerprints mit RDKit erstellt (s. Abbildung 11.1) [283, 284].

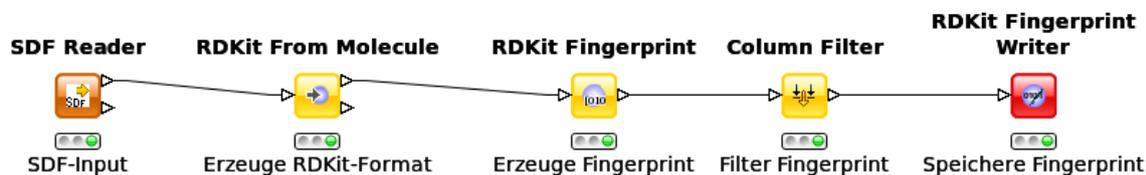


Abbildung 11.1 – Knime-Workflow zur Erstellung von Daylight-Fingerprints. Mithilfe einer SDF-Datei werden die Daylight-Fingerprints von RDKit erzeugt und abgespeichert.

11.3.13 Leader-Clustering

Das Clustering wurde mithilfe eines GPU-Skripts von LIAO *et al.* durchgeführt [228].

```
>$ ./gpu_leader -t Distance Fingerprints # Distance: Unterschied 1 - Tk,
                                     Fingerprints: Datei mit Fingerprints
```

11.3.14 Glide-Docking

Die Glide-Dockings wurden mit Version 6.2 durchgeführt [220]. Wenn nicht anders angegeben, wurde wie folgt vorgegangen:

Zunächst wurde ein Rezeptor-Grid erstellt. Dafür wurden die Bindetasche und deren Größe mit dem nativen Liganden der Röntgenkristallstruktur bestimmt. Weiterhin wurden entsprechend der Pharmakophore *Constraints* definiert.

Die Liganden wurden mit LigPrep [230] vorbereitet: Zunächst wurde auf die Ligandenatome das OPL_2005-Kraftfeld [103] angelegt. Mithilfe von Epik [220] wurde bei einem pH-Wert von 7.0 ± 2.0 mögliche Tautomere und Ionisierungen berechnet, die vorhanden jedoch beibehalten. Es wurden nur die vorgegebenen Chiralitätszentren behalten und keine neuen generiert. Für jeden Liganden wurden bis zu fünf Ringkonformationen erstellt.

Das Docking fand mit den vorbereiteten Liganden im *SP-Modus* statt. Die Liganden wurden flexibel behandelt, Stickstoff-Inversionen erlaubt und nicht-planare Amide bestraft. Epik-Strafen wurden nicht in den Docking-Score einberechnet. Während des Dockings mussten alle zuvor definierten *Constraints* erfüllt werden. Standardmäßig wurden weiterhin bis zu fünf Posen minimiert und maximal eine Pose pro Liganden abgespeichert.

12 Anhang

12.1 MD-Simulationen

Hier dargestellt sind die Endposen der MD-Simulationen mit verschiedenen Wassern von den Röntgenkristallstrukturen. Für jedes Bild ist der Ligand der Kristallstruktur grün dargestellt. Bei der Simulation ohne Wat A ist der Ligand rot, bei der Simulation nur mit Wat A gelb gezeigt. Alle Wasser sind zur Orientierung transparent dargestellt.

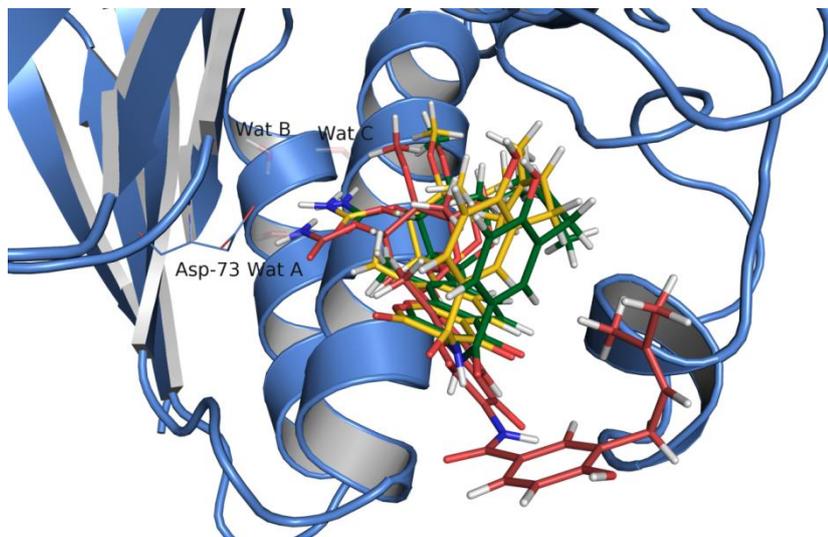


Abbildung 12.1 – MD-Simulation von 1AJ6 mit verschiedenen Wassern (Legende s. o.). Bei der Simulation ohne Wat A ist die Pose nicht stabil und das Stickstoffatom des Carbamats wandert an die Position des Wat A und der Bindungsmodus geht auch außerhalb der Bindetasche verloren.

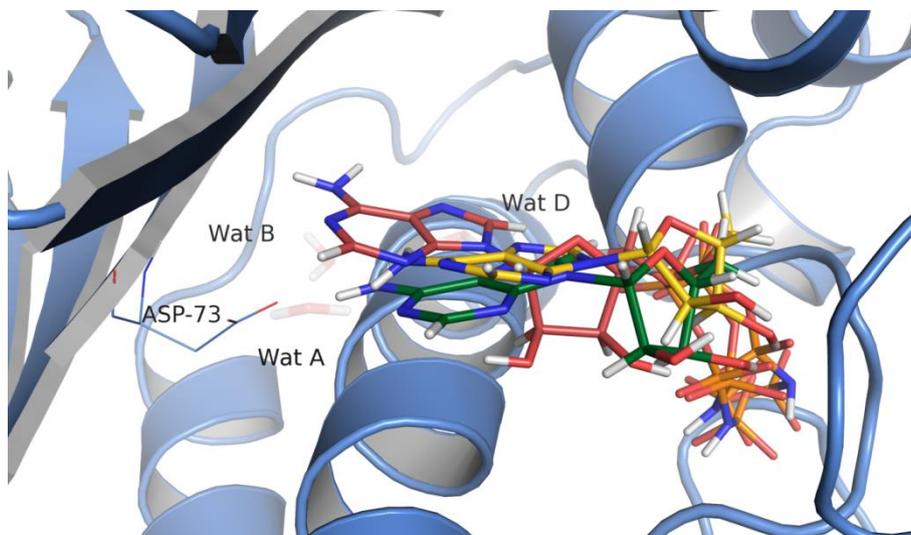


Abbildung 12.2 – MD-Simulation von 1EI1 mit verschiedenen Wassern (Legende s. o.). Ohne Wat A wandert der Adeninkörper weg und ein anderes Wasser übernimmt die Interaktion. Insgesamt ist dies die Struktur, in der Simulation nur mit Wat A die Pose unbeständiger ist, die Hauptinteraktionen bleiben dennoch erhalten.

12.2 SuperStar

Im Folgenden sind die Ergebnisse der SuperStar-Berechnungen gezeigt. Für die Darstellungen wurden alle Felder, die sich innerhalb der Liganden befanden entfernt und nur die frei zugänglichen Positionen berücksichtigt.

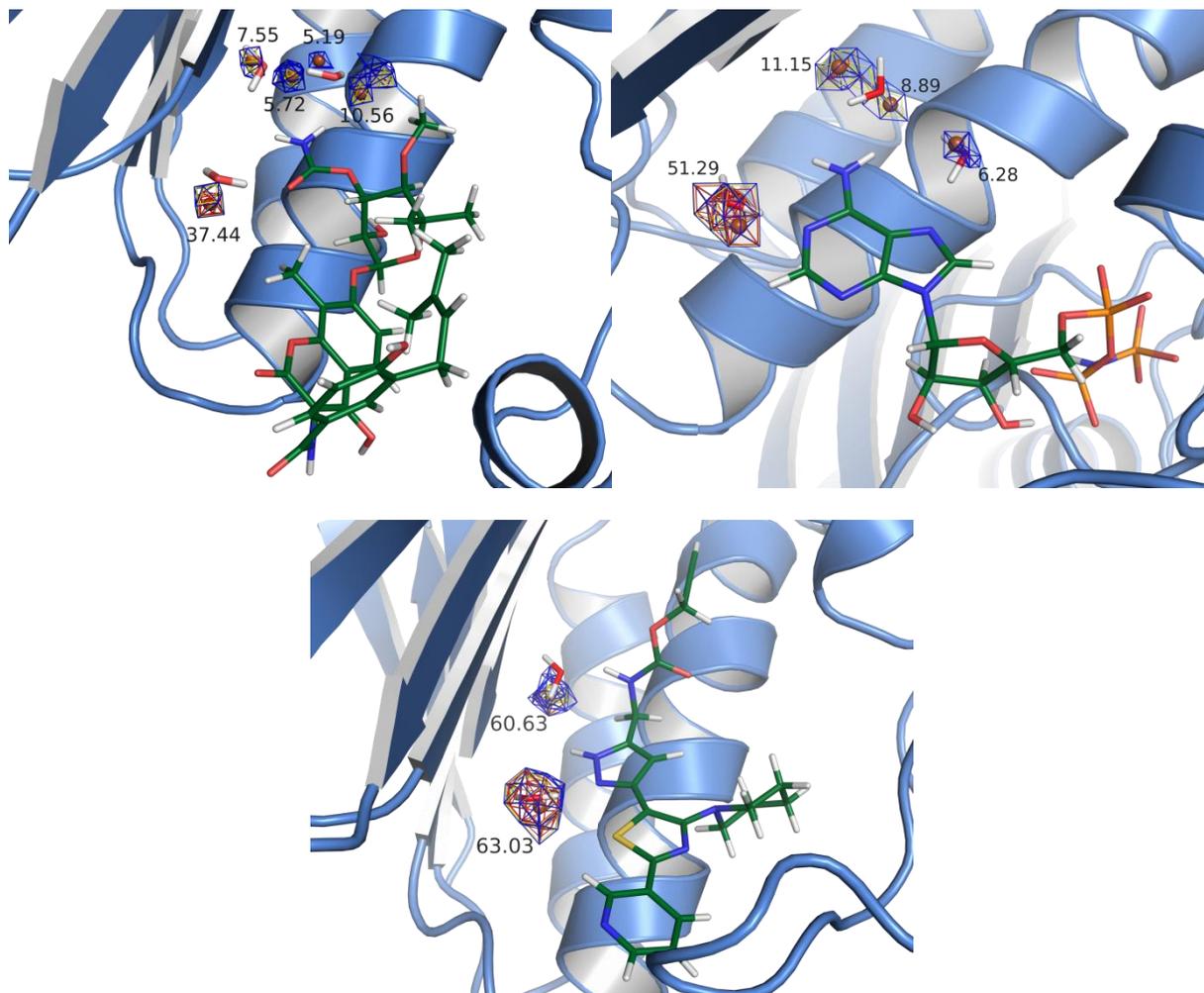


Abbildung 12.3 – Ergebnisse der SuperStar-Analyse. links oben: 1AJ6. rechts oben: 1E11. unten: 3G7E. Konturen: Wahrscheinlichkeit der Wasser-Sonde: blau > 2, gelb > 4, rot > 8. Sphären: Position der höchsten Wahrscheinlichkeit. Zahlenwert: Maximum der höchsten Wahrscheinlichkeit für die Sonde in diesem Bereich.

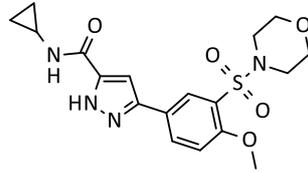
12.3 Substanzliste für DUD-E-Datensatz

Tabelle 12.1 – Substanzliste der für die Erstellung des DUD-E -Datensatzes verwendeten aktiven Substanzen. Für die Auswahl wurden nur Substanzen berücksichtigt, die eine gute Aktivität in der Gyrase von *E. coli* besaßen (Stand Sommer 2013). Aus Platzgründen sind nur die Quellen und die in den Quellen verwendeten Identifikationsnummer aufgeführt.

Literaturquelle	Identifizierungsnummer	Anzahl
[191]	5g, 5i, 5m, 5n, 5o	5
[73]	22, 23	2
[35]	1	1
[74]	1a, 1b, 7a, 7b, 8a, 8b	6
[85]	2 – 14, 16 – 21	19
[192]	13,24,27,31,33,37	6
[86]	Quorex, Dainippon, Vertex, Prolysis, AstraZeneca 1, AstraZeneca 2	6
[88]	1, 1g, 1h, 1i, 2a-k, 3b ,4a-e	21
[92]	1 – 5	5
[78]	1, 2, 26	3
[93]	12, 13	2
[94]	18	1
[193]	A, B, D	3
[194]	14, 23 –36	15
Summe		95

12.4 Spektren

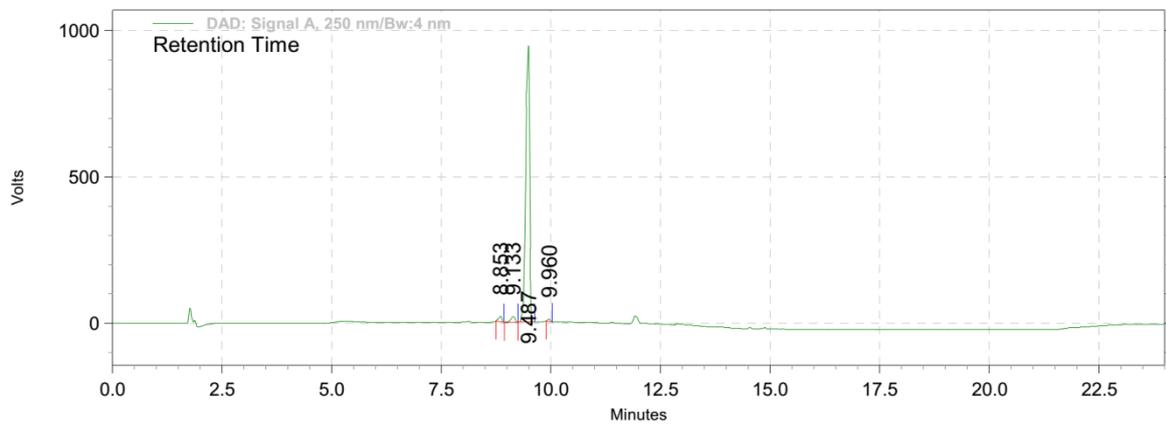
ZINC1586058 (126)



Die Substanz wurde für die Analytik in Dimethylsulfoxid gelöst.

HRMS (ESI): m/z ber. für C₁₈H₂₃N₄O₅S: 407.1384, gef.: 407.1384 [M+H]⁺

Gehalt: 95.62 ± 0.02 % (n = 3, α = 0.05); s_{rel} = 0.01 %

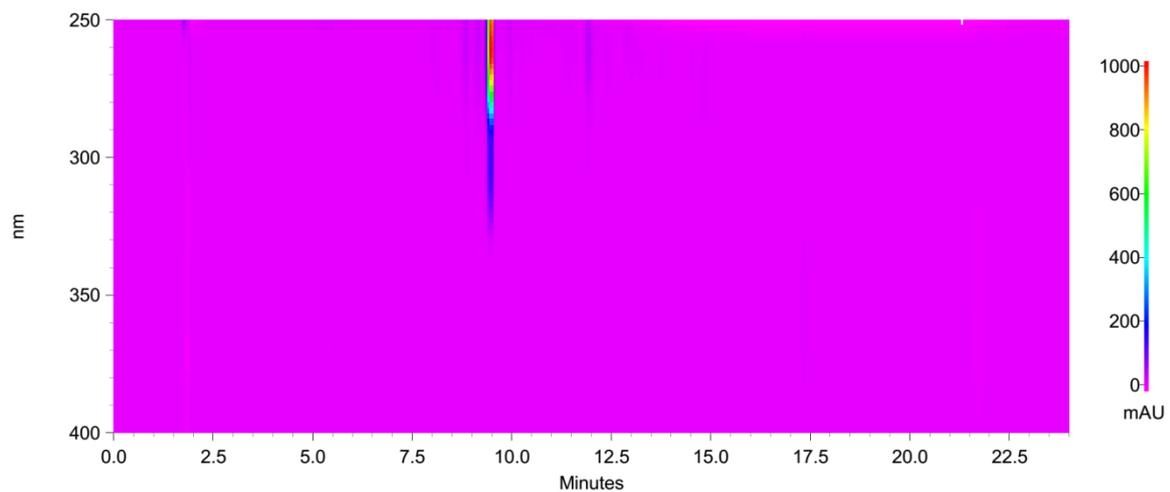


**DAD: Signal A,
250 nm/Bw:4 nm**

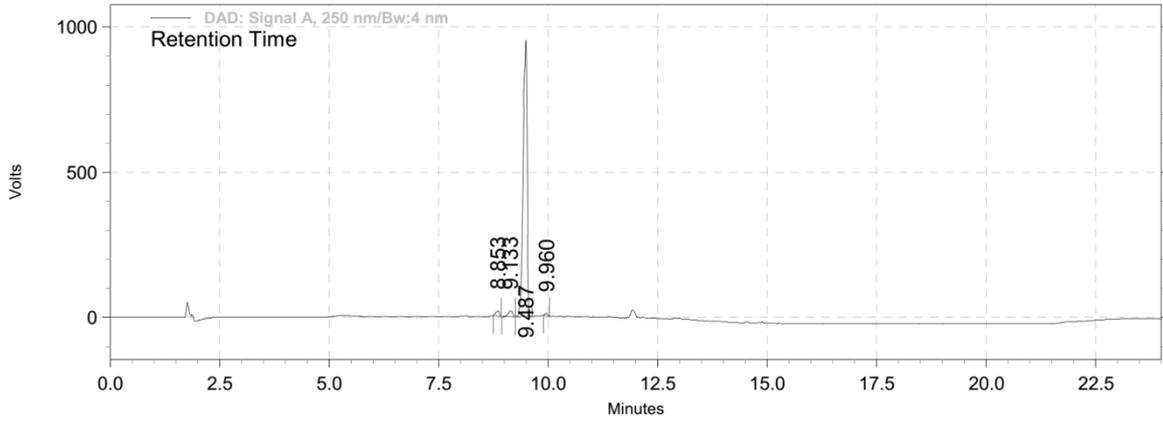
Results

Retention Time	Area	Area %	Height	Height %
8.853	192356	1.52	37771	1.82
9.133	291740	2.30	41408	1.99
9.487	12134512	95.61	1980693	95.36
9.960	72515	0.57	17250	0.83

Totals	Area	Area %	Height	Height %
	12691123	100.00	2077122	100.00



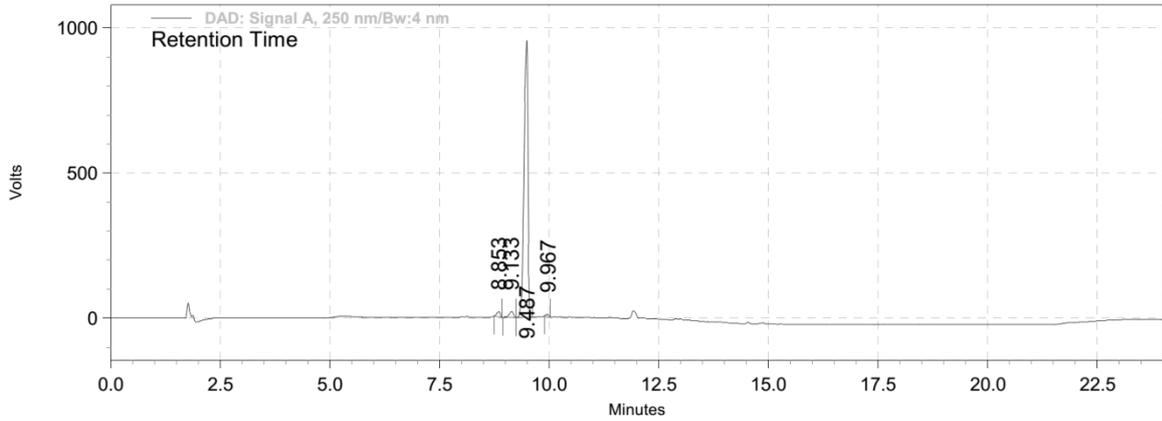
12. Anhang



**DAD: Signal A,
250 nm/Bw:4 nm
Results**

Retention Time	Area	Area %	Height	Height %
8.853	193146	1.51	37883	1.81
9.133	293431	2.30	41662	1.99
9.487	12213496	95.63	1995011	95.37
9.960	71906	0.56	17398	0.83

Totals	Area	Area %	Height	Height %
	12771979	100.00	2091954	100.00

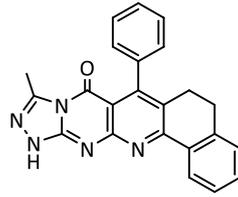


**DAD: Signal A,
250 nm/Bw:4 nm
Results**

Retention Time	Area	Area %	Height	Height %
8.853	195184	1.53	38039	1.81
9.133	291483	2.28	41494	1.98
9.487	12233350	95.62	1999177	95.37
9.967	73128	0.57	17565	0.84

Totals	Area	Area %	Height	Height %
	12793145	100.00	2096275	100.00

ZINC1586058 (161)

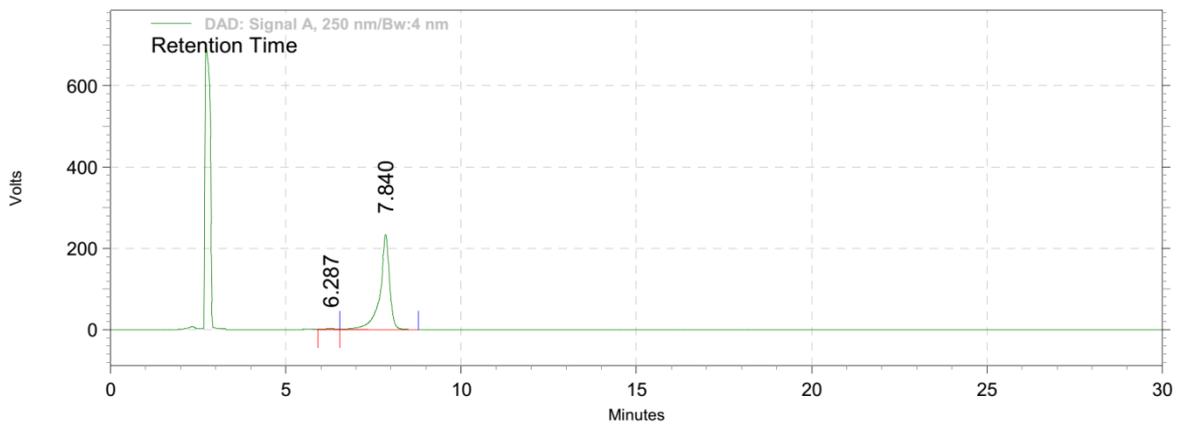


Die Substanz wurde für die Analytik in Dichlormethan gelöst.

HRMS (ESI): m/z ber. für C₂₃H₁₈N₅O: 380,1506, gef.: 380.1504 [M+H]⁺

CHNS-O-Analyse: gef. (ber.) [%]: C 72.26 (72.81); H 4.56 (4.52); N 18.67 (18.46); O 4,56 (4.22)

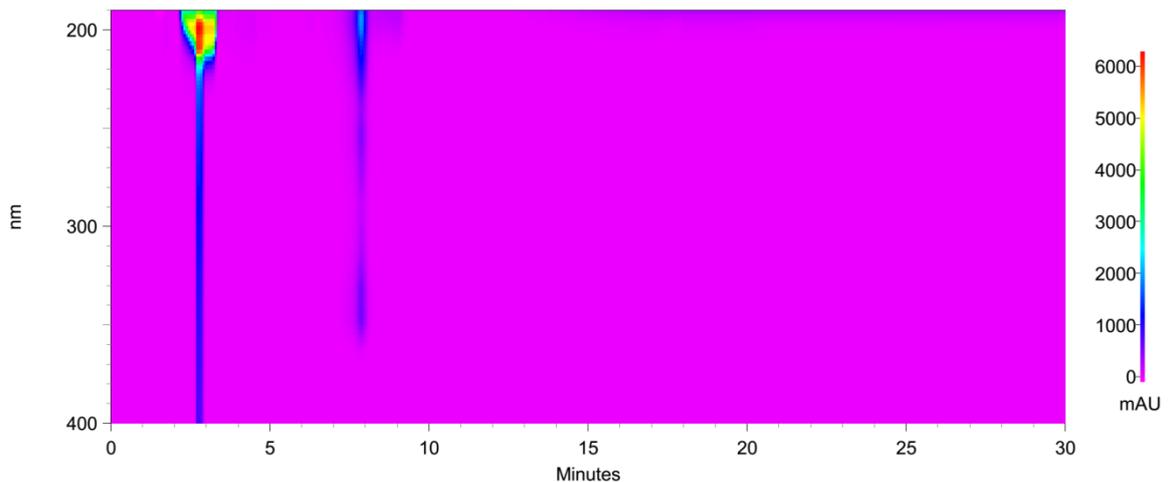
Gehalt: 99.33 ± 0.01 % (n = 3, α = 0.05); s_{rel} = 0.00 %



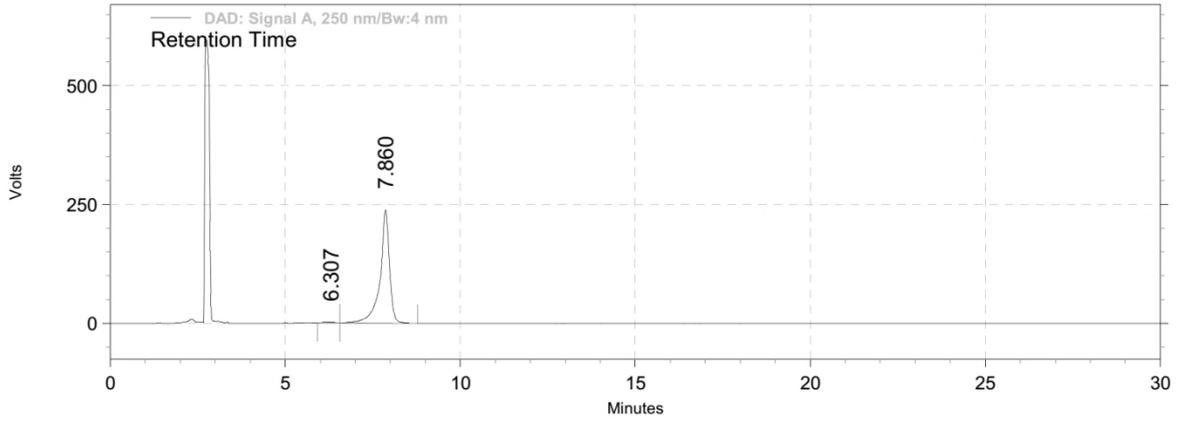
**DAD: Signal A,
250 nm/Bw:4 nm
Results**

Retention Time	Area	Area %	Height	Height %
6.287	63132	0.67	3806	0.77
7.840	9296594	99.33	490064	99.23

Totals	Area	Area %	Height	Height %
	9359726	100.00	493870	100.00



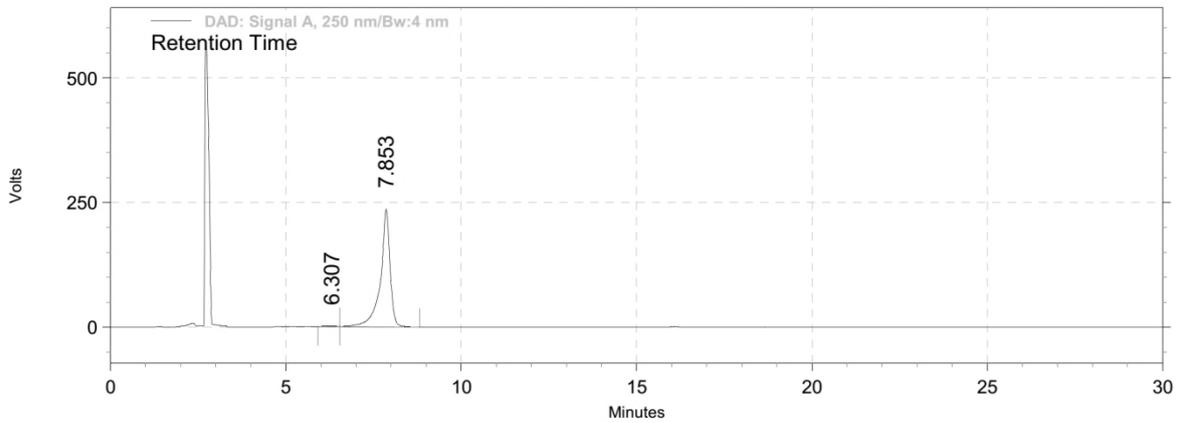
12. Anhang



**DAD: Signal A,
250 nm/Bw:4 nm
Results**

Retention Time	Area	Area %	Height	Height %
6.307	63629	0.67	3796	0.75
7.860	9380046	99.33	500487	99.25

Totals	9443675	100.00	504283	100.00
--------	---------	--------	--------	--------

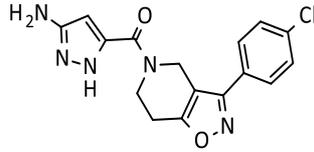


**DAD: Signal A,
250 nm/Bw:4 nm
Results**

Retention Time	Area	Area %	Height	Height %
6.307	63119	0.66	3720	0.75
7.853	9530412	99.34	495219	99.25

Totals	9593531	100.00	498939	100.00
--------	---------	--------	--------	--------

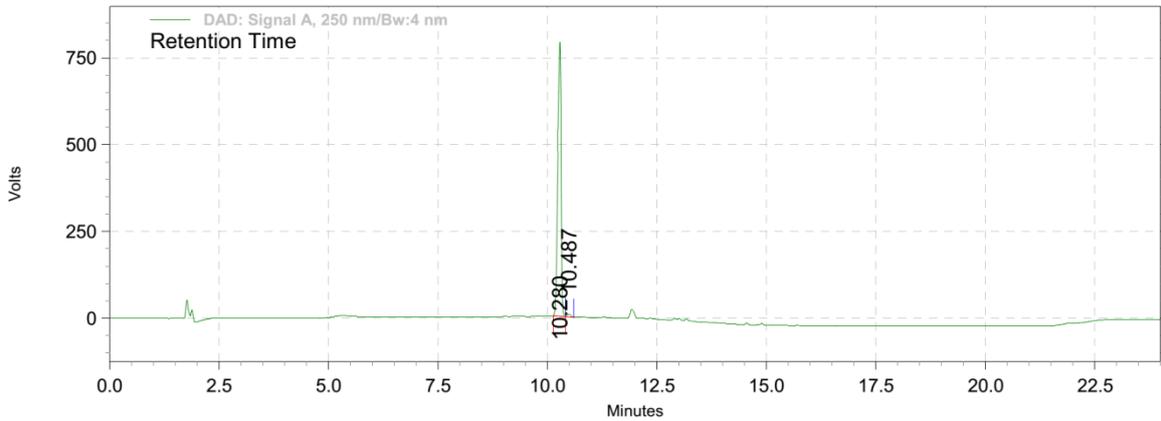
ZINC71415541 (140)



Die Substanz wurde für die Analytik in Dimethylsulfoxid gelöst.

HRMS (ESI): m/z ber. für C₁₆H₁₅ClN₅O₂: 344.0909, gef.: 344.0909 [M+H]⁺

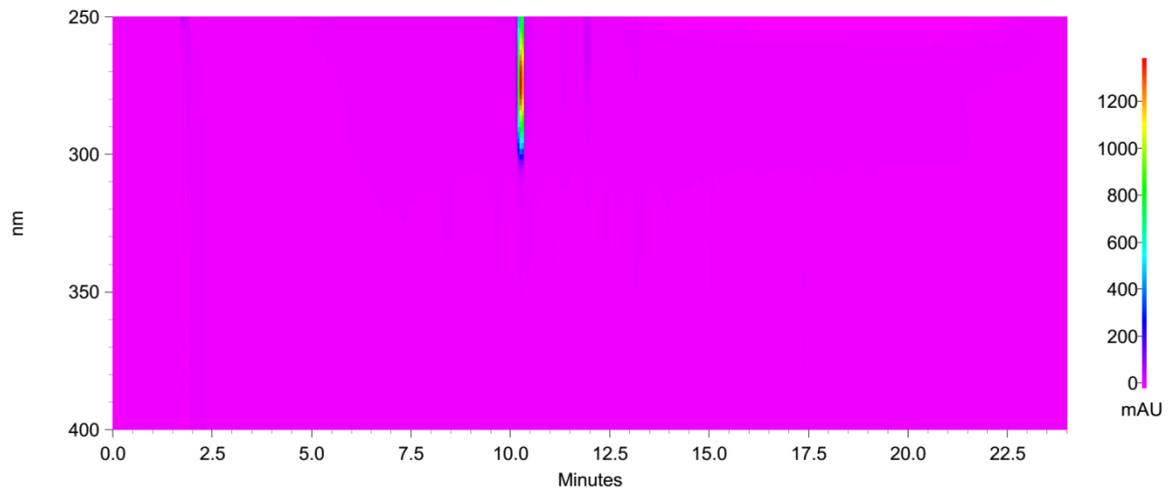
Gehalt: 98.92 ± 0.04 % (n = 3, α = 0.05); s_{rel} = 0.01 %



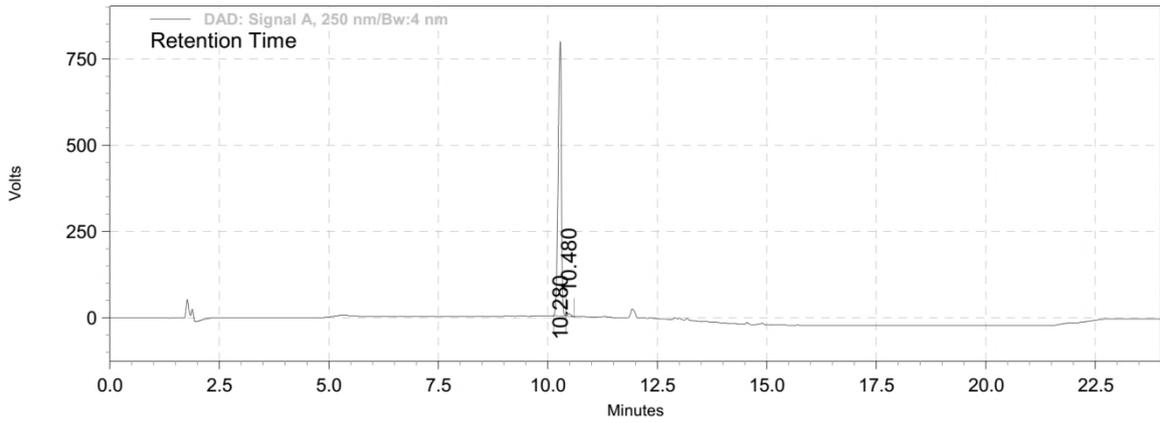
**DAD: Signal A,
250 nm/Bw:4 nm
Results**

Retention Time	Area	Area %	Height	Height %
10.280	8672823	98.91	1656533	99.15
10.487	95890	1.09	14163	0.85

Totals	Area	Area %	Height	Height %
	8768713	100.00	1670696	100.00

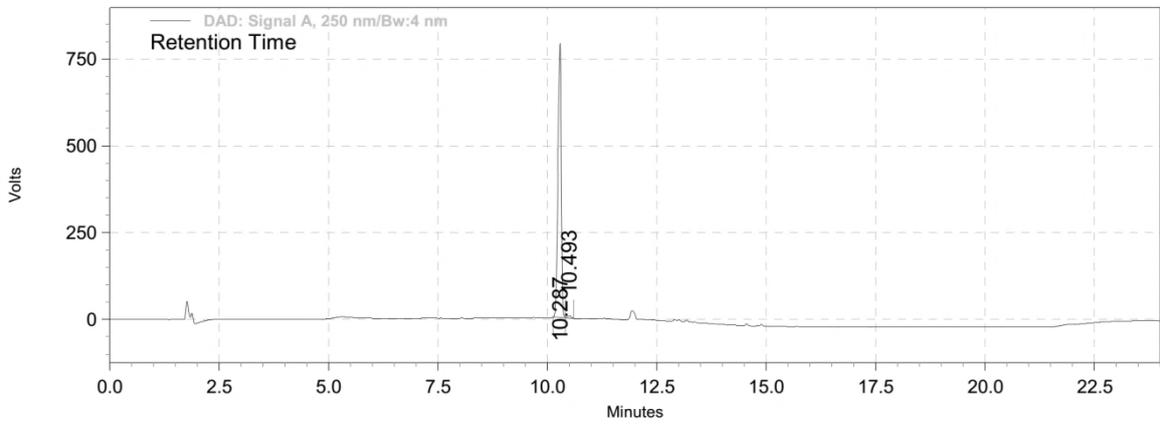


12. Anhang



**DAD: Signal A,
250 nm/Bw:4 nm
Results**

Retention Time	Area	Area %	Height	Height %
10.280	8718285	98.91	1668132	99.16
10.480	95660	1.09	14174	0.84
Totals	8813945	100.00	1682306	100.00



**DAD: Signal A,
250 nm/Bw:4 nm
Results**

Retention Time	Area	Area %	Height	Height %
10.287	8748647	98.94	1657698	99.17
10.493	94143	1.06	13897	0.83
Totals	8842790	100.00	1671595	100.00

13 Literatur

1. Perry R. D., Fetherston J. D.: *Yersinia pestis* - etiologic agent of plague. *Clin. Microbiol. Rev.* **1997**, 10 (1), 35–66.
2. Fleming A.: On the Antibacterial Action of Cultures of a *Penicillium*, with Special Reference to their Use in the Isolation of *B. influenzae*. *Br. J. Exp. Pathol.* **1929**, 10 (3), 226–236.
3. Fernandes P.: Antibacterial discovery and development - the failure of success? *Nat. Biotechnol.* **2006**, 24 (12), 1497–1503.
4. Lewis K.: Platforms for antibiotic discovery. *Nat. Rev. Drug Discov.* **2013**, 12 (5), 371–387.
5. Flemming: Penicillin. **1964**. *Nobel Lectures, Physiology or Medicine 1942-1962*. Elsevier Publishing Company, Amsterdam.
6. World Health Organization: Antimicrobial Resistance: Global Report on Surveillance. **2014**, Genf, Schweiz.
7. Magiorakos A.-P., Srinivasan A., Carey R. B., Carmeli Y., Falagas M. E., Giske C. G., Harbarth S., Hindler J. F., Kahlmeter G., Olsson-Liljequist B., Paterson D. L., Rice L. B., Stelling J., Struelens M. J., Vatopoulos A., Weber J. T., Monnet D. L.: Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. *Clin. Microbiol. Infect.* **2012**, 18 (3), 268–281.
8. Livermore D. M.: Fourteen years in resistance. *Int. J. Antimicrob. Agents.* **2012**, 39 (4), 283–294.
9. Livermore D. M.: Has the era of untreatable infections arrived? *J. Antimicrob. Chemother.* **2009**, 64 (suppl 1), i29.
10. Spellberg B., Guidos R., Gilbert D., Bradley J., Boucher H. W., Scheld W. M., Bartlett J. G., Edwards J., the Infectious Diseases Society of America: The Epidemic of Antibiotic-Resistant Infections: A Call to Action for the Medical Community from the Infectious Diseases Society of America. *Clin. Infect. Dis.* **2008**, 46 (2), 155–164.
11. Higgins P. G., Dammhayn C., Hackel M., Seifert H.: Global spread of carbapenem-resistant *Acinetobacter baumannii*. *J. Antimicrob. Chemother.* **2010**, 65 (2), 233–238.
12. Fernandes P.: The global challenge of new classes of antibacterial agents: an industry perspective. *Curr. Opin. Pharmacol.* **2015**, 24, 7–11.
13. Boucher H. W., Talbot G. H., Bradley J. S., Edwards J. E., Gilbert D., Rice L. B., Scheld M., Spellberg B., Bartlett J.: Bad Bugs, No Drugs. *Clin. Infect. Dis.* **2009**, 48 (1), 1–12.
14. Silver L. L.: Challenges of Antibacterial Discovery. *Clin. Microbiol. Rev.* **2011**, 24 (1), 71–109.

15. Stanton T. B.: A call for antibiotic alternatives research. *Trends Microbiol.* **2013**, 21 (3), 111–113.
16. So A. D., Shah T. A.: New business models for antibiotic innovation. *Upsala J. Med. Sci.* **2014**, 119 (2), 176–180.
17. Spellberg B.: The future of antibiotics. *Crit. Care Med.* **2014**, 18 (3), 228.
18. May M.: Drug development: Time for teamwork. *Nature.* **2014**, 509 (7498), S4.
19. Theuretzbacher U.: Accelerating resistance, inadequate antibacterial drug pipelines and international responses. *Int. J. Antimicrob. Agents.* **2012**, 39 (4), 295–299.
20. Misteli T.: Physiological importance of RNA and protein mobility in the cell nucleus. *Histochem. Cell Biol.* **2008**, 129 (1), 5–11.
21. Wang J. C.: Cellular roles of DNA topoisomerases: a molecular perspective. *Nat. Rev. Mol. Cell Biol.* **2002**, 3 (6), 430–440.
22. Wang J. C.: DNA Topoisomerases. *Annu. Rev. Biochem.* **1996**, 65 (1), 635–692.
23. Schoeffler A. J., Berger J. M.: DNA topoisomerases: harnessing and constraining energy to govern chromosome topology. *Q. Rev. Biophys.* **2008**, 41 (01), 41–101.
24. Higgins N. P., Peebles C. L., Sugino A., Cozzarelli N. R.: Purification of subunits of Escherichia coli DNA gyrase and reconstitution of enzymatic activity. *Proc. Natl. Acad. Sci.* **1978**, 75 (4), 1773–1777.
25. Kreuzer K. N., Cozzarelli N. R.: Formation and resolution of DNA catenanes by DNA gyrase. *Cell.* **1980**, 20 (1), 245–254.
26. Dutta R., Inouye M.: GHKL, an emergent ATPase/kinase superfamily. *Trends Biochem. Sci.* **2000**, 25 (1), 24–28.
27. Gubaev A., Klostermeier D.: The mechanism of negative DNA supercoiling: A cascade of DNA-induced conformational changes prepares gyrase for strand passage. *DNA Repair.* **2014**, 16, 23–34.
28. Papillon J., Ménétret J.-F., Batisse C., Hélye R., Schultz P., Potier N., Lamour V.: Structural insight into negative DNA supercoiling by DNA gyrase, a bacterial type 2A DNA topoisomerase. *Nucl. Acids Res.* **2013**, 41 (16), 7815–7827.
29. Rudolph M. G., Klostermeier D.: Mapping the Spectrum of Conformational States of the DNA- and C-Gates in Bacillus subtilis Gyrase. *J. Mol. Biol.* **2013**, 425 (15), 2632–2640.
30. Hearnshaw S. J., Chung T. T.-H., Stevenson, Clare Elizabeth Mary, Maxwell A., Lawson D. M.: The role of monovalent cations in the ATPase reaction of DNA gyrase. **2015**, 71 (4), 996–1005.
31. Lanz M. A., Klostermeier D.: Guiding strand passage. *Nucleic Acids Res.* **2011**, 39 (22), 9681–9694.

32. Lanz M. A., Farhat M., Klostermeier D.: The Acidic C-terminal Tail of the GyrA Subunit Moderates the DNA Supercoiling Activity of *Bacillus subtilis* Gyrase. *J. Biol. Chem.* **2014**, 289 (18), 12275–12285.
33. Costenaro L., Grossmann J. G., Ebel C., Maxwell A.: Modular Structure of the Full-Length DNA Gyrase B Subunit Revealed by Small-Angle X-Ray Scattering. *Structure.* **2007**, 15 (3), 329–339.
34. Bates A. D., Berger J. M., Maxwell A.: The ancestral role of ATP hydrolysis in type II topoisomerases: prevention of DNA double-strand breaks. *Nucleic Acids Res.* **2011**, 39 (15), 6327–6339.
35. Brino L., Urzhumtsev A., Mousli M., Bronner C., Mitschler A., Oudet P., Moras D.: Dimerization of *Escherichia coli* DNA-gyrase B Provides a Structural Mechanism for Activating the ATPase Catalytic Center. *J. Biol. Chem.* **2000**, 275 (13), 9468–9475.
36. Collin F., Karkare S., Maxwell A.: Exploiting bacterial DNA gyrase as a drug target. *Appl. Microbiol. Biotechnol.* **2011**, 92 (3), 479–497.
37. Hooper D. C.: Mechanisms of Action of Antimicrobials: Focus on Fluoroquinolones. *Clin. Infect. Dis.* **2001**, 32 (Supplement 1), S9.
38. Potron A., Poirel L., Nordmann P.: Emerging broad-spectrum resistance in *Pseudomonas aeruginosa* and *Acinetobacter baumannii*: Mechanisms and epidemiology. *Int. J. Antimicrob. Agents.* **2015**, 45 (6), 568–585.
39. Bax B. D., Chan P. F., Eggleston D. S., Fosberry A., Gentry D. R., Gorrec F., Giordano I., Hann M. M., Hennessy A., Hibbs M., Huang J., Jones E., Jones J., Brown K. K., Lewis C. J., May E. W., Saunders M. R., Singh O., Spitzfaden C. E., Shen C., Shillings A., Theobald A. J., Wohlkonig A., Pearson N. D., Gwynn M. N.: Type IIA topoisomerase inhibition by a new class of antibacterial agents. *Nature.* **2010**, 466 (7309), 935–940.
40. Huang Z., Lin K., You Q.: De novo design of novel DNA-gyrase inhibitors based on 2D molecular fingerprints. *Bioorg. Med. Chem. Lett.* **2013**, 23 (14), 4166–4171.
41. Surivet J.-P., Zumbunn C., Ruedi G., Hubschwerlen C., Bur D., Bruyère T., Locher H., Ritz D., Keck W., Seiler P., Kohl C., Gauvin J.-C., Mirre A., Kaegi V., Dos Santos M., Gaertner M., Delers J., Enderlin-Paput M., Boehme M.: Design, Synthesis, and Characterization of Novel Tetrahydropyran-Based Bacterial Topoisomerase Inhibitors with Potent Anti-Gram-Positive Activity. *J. Med. Chem.* **2013**, 56 (18), 7396–7415.
42. Hameed S. P., Raichurkar A., Madhavapeddi P., Menasinakai S., Sharma S., Kaur P., Nandishai-ah R., Panduga V., Reddy J., Sambandamurthy V. K., Sriram D.: Benzimidazoles: Novel Mycobacterial Gyrase Inhibitors from Scaffold Morphing. *ACS Med. Chem. Lett.* **2014**, 5 (7), 820–825.

43. So W., Crandon J. L., Nicolau D. P.: Pharmacodynamic Profile of GSK2140944 against Methicillin-Resistant *Staphylococcus aureus* in a Murine Lung Infection Model. *Antimicrob. Agents Chemother.* **2015**, 59 (8), 4956–4961.
44. Klebe G.: *Wirkstoffdesign: Entwurf und Wirkung von Arzneistoffen.* **2009**, 2. Edition. Spektrum Akademischer Verlag, Heidelberg.
45. Oprea T. I., Davis A. M., Teague S. J., Leeson P. D.: Is There a Difference between Leads and Drugs? A Historical Perspective. *J. Chem. Inf. Comput. Sci.* **2001**, 41 (5), 1308–1315.
46. Valler M. J., Green D.: Diversity screening versus focussed screening in drug discovery. *Drug Discov. Today.* **2000**, 5 (7), 286–293.
47. Fox S., Farr-Jones S., Yund M. A.: High Throughput Screening for Drug Discovery: Continually Transitioning into New Technology. *J. Biomol. Screening.* **1999**, 4 (4), 183–186.
48. Entzeroth M.: Emerging trends in high-throughput screening. *Curr. Opin. Pharmacol.* **2003**, 3 (5), 522–529.
49. Oprea T.: Current trends in lead discovery. *J. Comput.-Aided Mol. Des.* **2002**, 16 (5), 325–334.
50. Anderson A. C.: The Process of Structure-Based Drug Design. *Chem. Biol.* **2003**, 10 (9), 787–797.
51. Jorgensen W. L.: The Many Roles of Computation in Drug Discovery. *Science.* **2004**, 303 (5665), 1813–1818.
52. Sliwoski G., Kothiwale S., Meiler J., Lowe E. W., Barker E. L.: Computational Methods in Drug Discovery. *Pharmacol. Rev.* **2014**, 66 (1), 334–395.
53. Klebe G.: Recent developments in structure-based drug design. *J. Mol. Med.* **2000**, 78 (5), 269–281.
54. Golebiowski A., Klopfenstein S. R., Portlock D. E.: Lead compounds discovered from libraries. *Curr. Opin. Chem. Biol.* **2001**, 5 (3), 273–284.
55. Lyne P. D.: Structure-based virtual screening: an overview. *Drug Discov. Today.* **2002**, 7 (20), 1047–1055.
56. Cheng T., Li Q., Zhou Z., Wang Y., Bryant S.: Structure-Based Virtual Screening for Drug Discovery: a Problem-Centric Review. *AAPS J.* **2012**, 14 (1), 133–141.
57. Ferreira L., dos Santos R., Oliva G., Andricopulo A.: Molecular Docking and Structure-Based Drug Design Strategies. *Molecules.* **2015**, 20 (7), 13384.
58. Berman H. M., Westbrook J., Feng Z., Gilliland G., Bhat T. N., Weissig H., Shindyalov I. N., Bourne P. E.: The Protein Data Bank. *Nucleic Acids Res.* **2000**, 28 (1), 235–242.
59. Rishton G. M.: Reactive compounds and in vitro false positives in HTS. *Drug Discov. Today.* **1997**, 2 (9), 382–384.

60. Malvezzi A., Rezende L. de, Izidoro M. A., Cezari M. H. S., Juliano L., Amaral A. T. d.: Uncovering false positives on a virtual screening search for cruzain inhibitors. *Bioorg. Med. Chem. Lett.* **2008**, 18 (1), 350–354.
61. Lewis R. J., Singh O. M., Smith C. V., Skarzynski T., Maxwell A., Wonacott A. J., Wigley D. B.: The nature of inhibition of DNA gyrase by the coumarins and the cyclothialidines revealed by X-ray crystallography. *EMBO J.* **1996**, 15 (6), 1412–1420.
62. Galm U., Heller S., Shapiro S., Page M., Li S.-M., Heide L.: Antimicrobial and DNA Gyrase-Inhibitory Activities of Novel Clorobiocin Derivatives Produced by Mutasynthesis. *Antimicrob. Agents Chemother.* **2004**, 48 (4), 1307–1312.
63. Anderle C., Stieger M., Burrell M., Reinelt S., Maxwell A., Page M., Heide L.: Biological Activities of Novel Gyrase Inhibitors of the Aminocoumarin Class. *Antimicrob. Agents Chemother.* **2008**, 52 (6), 1982–1990.
64. Alt S., Mitchenall L. A., Maxwell A., Heide L.: Inhibition of DNA gyrase and DNA topoisomerase IV of *Staphylococcus aureus* and *Escherichia coli* by aminocoumarin antibiotics. *J. Antimicrob. Chemother.* **2011**, 66 (9), 2061–2069.
65. Tambo-ong A., Chopra S., Glaser B. T., Matsuyama K., Tran T., Madrid P. B.: Mannich reaction derivatives of novobiocin with modulated physicochemical properties and their antibacterial activities. *Bioorg. Med. Chem. Lett.* **2011**, 21 (19), 5697–5700.
66. Mayer C., Janin Y. L.: Non-quinolone Inhibitors of Bacterial Type IIA Topoisomerases: A Feat of Bioisosterism. *Chem. Rev.* **2014**, 114 (4), 2313–2342.
67. Phillips J. W., Goetz M. A., Smith S. K., Zink D. L., Polishook J., Onishi R., Salowe S., Wiltsie J., Allocco J., Sigmund J., Dorso K., Lee S., Skwish S., de la Cruz, Mercedes, Martín J., Vicente F., Genilloud O., Lu J., Painter R. E., Young K., Overbye K., Donald, Robert G. K., Singh S. B.: Discovery of Kibdelomycin, A Potent New Class of Bacterial Type II Topoisomerase Inhibitor by Chemical-Genetic Profiling in *Staphylococcus aureus*. *Chem. Biol.* **2011**, 18 (8), 955–965.
68. Lu J., Patel S., Sharma N., Soisson S. M., Kishii R., Takei M., Fukuda Y., Lumb K. J., Singh S. B.: Structures of Kibdelomycin Bound to *Staphylococcus aureus* GyrB and ParE Showed a Novel U-Shaped Binding Mode. *ACS Chem. Biol.* **2014**, 9 (9), 2023–2031.
69. Holdgate G. A., Tunnicliffe A., Ward W. H. J., Weston S. A., Rosenbrock G., Barth P. T., Taylor I. W. F., Pauptit R. A., Timms D.: The Entropic Penalty of Ordered Water Accounts for Weaker Binding of the Antibiotic Novobiocin to a Resistant Mutant of DNA Gyrase: A Thermodynamic and Crystallographic Study. *Biochemistry.* **1997**, 36 (32), 9663–9673.
70. Tomasic T., Masic L. P.: Prospects for developing new antibacterials targeting bacterial type IIA topoisomerases. *Curr. Top. Med. Chem.* **2014**, 14 (1), 130–151.

71. Sheng C., Zhang W.: Fragment Informatics and Computational Fragment-Based Drug Design: An Overview and Update. *Med. Res. Rev.* **2013**, 33 (3), 554–598.
72. Chen Y., Pohlhaus D. T.: In silico docking and scoring of fragments. *Drug Discovery Today: Technol.* **2010**, 7 (3), e149.
73. Boehm H.-J., Boehringer M., Bur D., Gmuender H., Huber W., Klaus W., Kostrewa D., Kuehne H., Luebbers T., Meunier-Keller N., Mueller F.: Novel Inhibitors of DNA Gyrase: 3D Structure Based Biased Needle Screening, Hit Validation by Biophysical Methods, and 3D Guided Optimization. A Promising Alternative to Random Screening. *J. Med. Chem.* **2000**, 43 (14), 2664–2674.
74. Tanitame A., Oyamada Y., Ofuji K., Kyoya Y., Suzuki K., Ito H., Kawasaki M., Nagai K., Wachi M., Yamagishi J.-i.: Design, synthesis and structure–activity relationship studies of novel indazole analogues as DNA gyrase inhibitors with Gram-positive antibacterial activity. *Bioorg. Med. Chem. Lett.* **2004**, 14 (11), 2857–2862.
75. Lübbers T., Angehrn P., Gmünder H., Herzig S.: Design, synthesis, and structure–activity relationship studies of new phenolic DNA gyrase inhibitors. *Bioorg. Med. Chem. Lett.* **2007**, 17 (16), 4708–4714.
76. Manchester J. I., Dussault D. D., Rose J. A., Boriack-Sjodin P. A., Uria-Nickelsen M., Ioannidis G., Bist S., Fleming P., Hull K. G.: Discovery of a novel azaindole class of antibacterial agents targeting the ATPase domains of DNA gyrase and Topoisomerase IV. *Bioorg. Med. Chem. Lett.* **2012**, 22 (15), 5150–5156.
77. Eakin A. E., Green O., Hales N., Walkup G. K., Bist S., Singh A., Mullen G., Bryant J., Embrey K., Gao N., Breeze A., Timms D., Andrews B., Uria-Nickelsen M., Demeritt J., Loch J. T., Hull K., Blodgett A., Illingworth R. N., Prince B., Boriack-Sjodin P. A., Hauck S., MacPherson L. J., Ni H., Sherer B.: Pyrrolamide DNA Gyrase Inhibitors: Fragment-Based Nuclear Magnetic Resonance Screening To Identify Antibacterial Agents. *Antimicrob. Agents Chemother.* **2012**, 56 (3), 1240–1246.
78. Sherer B. A., Hull K., Green O., Basarab G., Hauck S., Hill P., Loch J. T., Mullen G., Bist S., Bryant J., Boriack-Sjodin A., Read J., DeGrace N., Uria-Nickelsen M., Illingworth R. N., Eakin A. E.: Pyrrolamide DNA gyrase inhibitors. *Bioorg. Med. Chem. Lett.* **2011**, 21 (24), 7416–7420.
79. Uria-Nickelsen M., Blodgett A., Kamp H., Eakin A., Sherer B., Green O.: Novel DNA gyrase inhibitors: Microbiological characterisation of pyrrolamides. *Int. J. Antimicrob. Agents.* **2013**, 41 (1), 28–35.
80. Hameed S. P., Solapure S., Mukherjee K., Nandi V., Waterson D., Shandil R., Balganes M., Sambandamurthy V. K., Raichurkar A. K., Deshpande A., Ghosh A., Awasthy D., Shanbhag G., Sheikh G., McMiken H., Puttur J., Reddy J., Werngren J., Read J., Kumar M., R M., Chinnapattu M., Madhavapeddi P., Manjrekar P., Basu R., Gaonkar S., Sharma S., Hoffner S., Humnabadkar

- V., Subbulakshmi V., Panduga V.: Optimization of Pyrrolamides as Mycobacterial GyrB ATPase Inhibitors: Structure-Activity Relationship and In Vivo Efficacy in a Mouse Model of Tuberculosis. *Antimicrob. Agents Chemother.* **2014**, 58 (1), 61–70.
81. Tomašič T., Katsamakos S., Hodnik Ž., Ilaš J., Brvar M., Solmajer T., Montalvão S., Tammela P., Banjanac M., Ergović G., Anderluh M., Mašič L. P., Kikelj D.: Discovery of 4,5,6,7-Tetrahydrobenzo[1,2-d]thiazoles as Novel DNA Gyrase Inhibitors Targeting the ATP-Binding Site. *J. Med. Chem.* **2015**, 58 (14), 5501–5521.
82. Zidar N., Macut H., Tomašič T., Brvar M., Montalvão S., Tammela P., Solmajer T., Peterlin Mašič L., Ilaš J., Kikelj D.: N-Phenyl-4,5-dibromopyrrolamides and N-Phenylindolamides as ATP Competitive DNA Gyrase B Inhibitors: Design, Synthesis, and Evaluation. *J. Med. Chem.* **2015**, 58 (15), 6179–6194.
83. Tari L. W., Trzoss M., Bensen D. C., Li X., Chen Z., Lam T., Zhang J., Creighton C. J., Cunningham M. L., Kwan B., Stidham M., Shaw K. J., Lightstone F. C., Wong S. E., Nguyen T. B., Nix J., Finn J.: Pyrrolopyrimidine inhibitors of DNA gyrase B (GyrB) and topoisomerase IV (ParE). Part I: Structure guided discovery and optimization of dual targeting agents with potent, broad-spectrum enzymatic activity. *Bioorg. Med. Chem. Lett.* **2013**, 23 (5), 1529–1536.
84. Tari L. W., Li X., Trzoss M., Bensen D. C., Chen Z., Lam T., Zhang J., Lee S. J., Hough G., Phillipson D., Akers-Rodriguez S., Cunningham M. L., Kwan B. P., Nelson K. J., Castellano A., Locke J. B., Brown-Driver V., Murphy T. M., Ong V. S., Pillar C. M., Shinabarger D. L., Nix J., Lightstone F. C., Wong S. E., Nguyen T. B., Shaw K. J., Finn J.: Tricyclic GyrB/ParE (TriBE) Inhibitors: A New Class of Broad-Spectrum Dual-Targeting Antibacterial Agents. *PLoS ONE.* **2013**, 8 (12), e84409.
85. Charifson P. S., Grillot A.-L., Grossman T. H., Parsons J. D., Badia M., Bellon S., Deininger D. D., Drumm J. E., Gross C. H., LeTiran A., Liao Y., Mani N., Nicolau D. P., Perola E., Ronkin S., Shannon D., Swenson L. L., Tang Q., Tessier P. R., Tian S.-K., Trudeau M., Wang T., Wei Y., Zhang H., Stamos D.: Novel Dual-Targeting Benzimidazole Urea Inhibitors of DNA Gyrase and Topoisomerase IV Possessing Potent Antibacterial Activity: Intelligent Design and Evolution through the Judicious Use of Structure-Guided Design and Structure-Activity Relationships. *J. Med. Chem.* **2008**, 51 (17), 5243–5263.
86. Starr J. T., Sciotti R. J., Hanna D. L., Huband M. D., Mullins L. M., Cai H., Gage J. W., Lockard M., Rauckhorst M. R., Owen R. M., Lall M. S., Tomilo M., Chen H., McCurdy S. P., Barbachyn M. R.: 5-(2-Pyrimidinyl)-imidazo[1,2-a]pyridines are antibacterial agents targeting the ATPase domains of DNA gyrase and topoisomerase IV. *Bioorg. Med. Chem. Lett.* **2009**, 19 (18), 5302–5306.
87. Basarab G. S., Manchester J. I., Bist S., Boriack-Sjodin P. A., Dangel B., Illingworth R., Sherer B. A., Sriram S., Uria-Nickelsen M., Eakin A. E.: Fragment-to-Hit-to-Lead Discovery of a Novel Pyri-

- dylurea Scaffold of ATP Competitive Dual Targeting Type II Topoisomerase Inhibiting Antibacterial Agents. *J. Med. Chem.* **2013**, 56 (21), 8712–8735.
88. Ronkin S. M., Badia M., Bellon S., Grillot A.-L., Gross C. H., Grossman T. H., Mani N., Parsons J. D., Stamos D., Trudeau M., Wei Y., Charifson P. S.: Discovery of pyrazolthiazoles as novel and potent inhibitors of bacterial gyrase. *Bioorg. Med. Chem. Lett.* **2010**, 20 (9), 2828–2831.
89. Sun J., Lv P.-C., Yin Y., Yuan R.-J., Ma J., Zhu H.-L., Taylor P.: Synthesis, Structure and Antibacterial Activity of Potent DNA Gyrase Inhibitors: N'-Benzoyl-3-(4-Bromophenyl)-1H-Pyrazole-5-Carbohydrazide Derivatives. *PLoS ONE*. **2013**, 8 (7), e69751.
90. Liu J.-J., Sun J., Fang Y.-B., Yang Y.-A., Jiao R.-H., Zhu H.-L.: Synthesis, and antibacterial activity of novel 4,5-dihydro-1H-pyrazole derivatives as DNA gyrase inhibitors. *Org. Biomol. Chem.* **2014**, 12 (6), 998–1008.
91. Shirude P. S., Madhavapeddi P., Tucker J. A., Murugan K., Patil V., Basavarajappa H., Raichurkar A. V., Humnabdkar V., Hussein S., Sharma S., Ramya V. K., Narayan C. B., Balganesht T. S., Sambandamurthy V. K.: Aminopyrazinamides: Novel and Specific GyrB Inhibitors that Kill Replicating and Nonreplicating Mycobacterium tuberculosis. *ACS Chem. Biol.* **2012**, 8 (3), 519–523.
92. Brvar M., Perdih A., Oblak M., Masic L. P., Solmajer T.: In silico discovery of 2-amino-4-(2,4-dihydroxyphenyl)thiazoles as novel inhibitors of DNA gyrase B. *Bioorg. Med. Chem. Lett.* **2010**, 20 (3), 958–962.
93. Brvar M., Perdih A., Hodnik V., Renko M., Anderluh G., Jerala R., Solmajer T.: In silico discovery and biophysical evaluation of novel 5-(2-hydroxybenzylidene) rhodanine inhibitors of DNA gyrase B. *Bioorg. Med. Chem.* **2012**, 20 (8), 2572–2580.
94. Brvar M., Perdih A., Renko M., Anderluh G., Turk D., Solmajer T.: Structure-Based Discovery of Substituted 4,5'-Bithiazoles as Novel DNA Gyrase Inhibitors. *J. Med. Chem.* **2012**, 55 (14), 6413–6426.
95. Singh S. B.: Confronting the challenges of discovery of novel antibacterial agents. *Bioorg. Med. Chem. Lett.* **2014**, 24 (16), 3683–3689.
96. Lafitte D., Lamour V., Tsvetkov P. O., Makarov A. A., Klich M., Deprez P., Moras D., Briand C., Gilli R.: DNA Gyrase Interaction with Coumarin-Based Inhibitors: The Role of the Hydroxybenzoate Isopentenyl Moiety and the 5'-Methyl Group of the Noviose. *Biochemistry*. **2002**, 41 (23), 7217–7223.
97. Lamour V., Hoermann L., Jeltsch J.-M., Oudet P., Moras D.: An Open Conformation of the *Thermus thermophilus* Gyrase B ATP-binding Domain. *J. Biol. Chem.* **2002**, 277 (21), 18947–18953.
98. PyMOL: The PyMOL Molecular Graphics System. 1.6.0.0. **2012**. Schrödinger, LLC, New York, NY.

99. Bissantz C., Kuhn B., Stahl M.: A Medicinal Chemist's Guide to Molecular Interactions. *J. Med. Chem.* **2010**, 53 (14), 5061–5084.
100. Stanger F. V., Dehio C., Schirmer T.: Structure of the N-Terminal Gyrase B Fragment in Complex with ADP-P_i Reveals Rigid-Body Motion Induced by ATP Hydrolysis. *PLoS ONE.* **2014**, 9 (9), e107289.
101. Saíz-Urra L., Cabrera M. A., Froeyen M.: Exploring the conformational changes of the ATP binding site of gyrase B from *Escherichia coli* complexed with different established inhibitors by using molecular dynamics simulation: Protein-ligand interactions in the light of the alanine scanning and free energy decomposition methods. *J. Mol. Graphics Modell.* **2010**, 29 (5), 726–739.
102. Schrödinger: Prime. 1.41. **2010**. Schrödinger, LLC, New York, NY.
103. Banks J. L., Beard H. S., Cao Y., Cho A. E., Damm W., Farid R., Felts A. K., Halgren T. A., Mainz D. T., Maple J. R., Murphy R., Philipp D. M., Repasky M. P., Zhang L. Y., Berne B. J., Friesner R. A., Gallicchio E., Levy R. M.: Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J. Comput. Chem.* **2005**, 26 (16), 1752–1780.
104. Ghosh A., Rapp C. S., Friesner R. A.: Generalized Born Model Based on a Surface Integral Formulation. *J. Phys. Chem. B.* **1998**, 102 (52), 10983–10990.
105. Rossi K. A., Weigelt C. A., Nayeem A., Krystek S. R.: Loopholes and missing links in protein modeling. *Protein Sci.* **2007**, 16 (9), 1999–2012.
106. Jacobson M. P., Pincus D. L., Rapp C. S., Day, Tyler J. F., Honig B., Shaw D. E., Friesner R. A.: A hierarchical approach to all-atom protein loop prediction. *Proteins: Struct., Funct., Bioinf.* **2004**, 55 (2), 351–367.
107. Sellers B. D., Zhu K., Zhao S., Friesner R. A., Jacobson M. P.: Toward better refinement of comparative models: Predicting loops in inexact environments. *Proteins: Struct., Funct., Bioinf.* **2008**, 72 (3), 959–971.
108. Jorgensen W. L., Maxwell D. S., Tirado-Rives J.: Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, 118 (45), 11225–11236.
109. Pearlman D. A., Case D. A., Caldwell J. W., Ross W. S., Cheatham III, Thomas E., DeBolt S., Ferguson D., Seibel G., Kollman P.: AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comput. Phys. Commun.* **1995**, 91 (1–3), 1–41.
110. Salomon-Ferrer R., Case D. A., Walker R. C.: An overview of the Amber biomolecular simulation package. *WIREs Comput. Mol. Sci.* **2013**, 3 (2), 198–210.

111. Case D., Darden T., Cheatham T. I., Simmerling C., Wang J., Duke R. L., Walker R., Zhang W., Merz K., Roberts B., Wang B., Hayik S., Roitberg A., Seabra G., Kolossváry I., Wong K., Paesani F., Vanicek J., Liu J., Wu X., Brozell S., Steinbrecher T., Gohlke H., Cai Q., Ye X., Hsieh M., Cui G., Roe D., Mathews D., Seetin M., Sagui C., Babin V., Luchko T., Gusarov S., Kovalenko A., Kollman P.: AMBER 11. **2010**. University of California, San Francisco.
112. Ryckaert J.-P., Ciccotti G., Berendsen H. J. C.: Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, 23 (3), 327–341.
113. Wang J., Wolf R. M., Caldwell J. W., Kollman P. A., Case D. A.: Development and testing of a general amber force field. *J. Comp. Chem.* **2004**, 25 (9), 1157–1174.
114. Hornak V., Abel R., Okur A., Strockbine B., Roitberg A., Simmerling C.: Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf.* **2006**, 65 (3), 712–725.
115. Adachi M., Ohhara T., Kurihara K., Tamada T., Honjo E., Okazaki N., Arai S., Shoyama Y., Kimura K., Matsumura H., Sugiyama S., Adachi H., Takano K., Mori Y., Hidaka K., Kimura T., Hayashi Y., Kiso Y., Kuroki R.: Structure of HIV-1 protease in complex with potent inhibitor KNI-272 determined by high-resolution X-ray and neutron crystallography. *Proc. Natl. Acad. Sci.* **2009**, 106 (12), 4641–4646.
116. Schio L., Chatreaux F., Loyau V., Murer M., Ferreira A., Mauvais P., Bonnefoy A., Klich M.: Fine Tuning of physico-chemical parameters to optimise a new series of novobiocin analogues. *Bioorg. Med. Chem. Lett.* **2001**, 11 (11), 1461–1464.
117. Holm L., Rosenström P.: Dali server: conservation mapping in 3D. *Nucl. Acids Res.* **2010**, 38 (suppl 2), W545-W549.
118. Barillari C., Taylor J., Viner R., Essex J. W.: Classification of Water Molecules in Protein Binding Sites. *J. Am. Chem. Soc.* **2007**, 129 (9), 2577–2587.
119. Ball P.: Water as an Active Constituent in Cell Biology. *Chem. Rev.* **2007**, 108 (1), 74–108.
120. Klebe G.: Applying thermodynamic profiling in lead finding and optimization. *Nat. Rev. Drug Discov.* **2015**, 14 (2), 95–110.
121. Pitt W. R., Goodfellow J. M.: Modelling of solvent positions around polar groups in proteins. *Protein Eng.* **1991**, 4 (5), 531–537.
122. Raymer M. L., Sanschagrin P. C., Punch W. F., Venkataraman S., Goodman E. D., Kuhn L. A.: Predicting conserved water-mediated and polar ligand interactions in proteins using a K-nearest-neighbors genetic algorithm. *J. Mol. Biol.* **1997**, 265 (4), 445–464.
123. Verdonk M. L., Cole J. C., Watson P., Gillet V., Willett P.: Superstar: improved knowledge-based interaction fields for protein binding sites. *J. Mol. Biol.* **2001**, 307 (3), 841–859.

124. Amadasi A., Surface J. A., Spyrakis F., Cozzini P., Mozzarelli A., Kellogg G. E.: Robust Classification of “Relevant” Water Molecules in Putative Protein Binding Sites. *J. Med. Chem.* **2008**, 51 (4), 1063–1067.
125. Ross G. A., Morris G. M., Biggin P. C.: Rapid and Accurate Prediction and Scoring of Water Molecules in Protein Binding Sites. *PLoS ONE.* **2012**, 7 (3), e32036.
126. Zheng M., Li Y., Xiong B., Jiang H., Shen J.: Water PMF for predicting the properties of water molecules in protein binding site. *J. Comp. Chem.* **2013**, 34 (7), 583–592.
127. Nittinger E., Schneider N., Lange G., Rarey M.: Evidence of Water Molecules—A Statistical Evaluation of Water Molecules Based on Electron Density. *J. Chem. Inf. Model.* **2015**, 55 (4), 771–783.
128. Cui G., Swails J. M., Manas E. S.: SPAM. *J. Chem. Theory Comput.* **2013**, 9 (12), 5539–5549.
129. Kohlmann A., Zhu X., Dalgarno D.: Application of MM-GB/SA and WaterMap to SRC Kinase Inhibitor Potency Prediction. *ACS Med. Chem. Lett.* **2012**, 3 (2), 94–99.
130. Abel R., Young T., Farid R., Berne B. J., Friesner R. A.: Role of the Active-Site Solvent in the Thermodynamics of Factor Xa Ligand Binding. *J. Am. Chem. Soc.* **2008**, 130 (9), 2817–2831.
131. Verdonk M. L., Cole J. C., Taylor R.: SuperStar. *J. Mol. Biol.* **1999**, 289 (4), 1093–1108.
132. Bruno I., Cole J., Lommerse J. M., Rowland R. S., Taylor R., Verdonk M.: IsoStar: A library of information about nonbonded interactions. *J. Comput. Aided Mol. Des.* **1997**, 11 (6), 525–537.
133. Boer D. R., Kroon J., Cole J. C., Smith B., Verdonk M. L.: Superstar: comparison of CSD and PDB-based interaction fields as a basis for the prediction of protein-ligand interactions. *J. Mol. Biol.* **2001**, 312 (1), 275–287.
134. Goodford P. J.: A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* **1985**, 1985, 849–857.
135. Fornabaio M., Spyrakis F., Mozzarelli A., Cozzini P., Abraham D. J., Kellogg G. E.: Simple, Intuitive Calculations of Free Energy of Binding for Protein–Ligand Complexes. 3. The Free Energy Contribution of Structural Water Molecules in HIV-1 Protease Complexes. *J. Med. Chem.* **2004**, 47 (18), 4507–4516.
136. Trott O., Olson A. J.: AutoDock Vina. *J. Comp. Chem.* **2010**, 31 (2), 455–461.
137. Kellogg G. E., Abraham D. J.: Hydrophobicity: is LogPo/w more than the sum of its parts? *Eur. J. Med. Chem.* **2000**, 35 (7–8), 651–661.
138. Kellogg G. E., Chen D. L.: The Importance of Being Exhaustive. Optimization of Bridging Structural Water Molecules and Water Networks in Models of Biological Systems. *Chem. Biodiversity.* **2004**, 1 (1), 98–105.
139. Yu H., Rick S. W.: Free Energies and Entropies of Water Molecules at the Inhibitor–Protein Interface of DNA Gyrase. *J. Am. Chem. Soc.* **2009**, 131 (18), 6608–6613.

140. Ostrov D. A., Hernandez Prada J. A., Corsino P. E., Finton K. A., Le N., Rowe T. C.: Discovery of Novel DNA Gyrase Inhibitors by High-Throughput Virtual Screening. *Antimicrob. Agents Chemother.* **2007**, 51 (10), 3688–3698.
141. Lipinski C., Hopkins A.: Navigating chemical space for biology and medicine. *Nature.* **2004**, 432 (7019), 855–861.
142. Dobson C. M.: Chemical space and biology. *Nature.* **2004**, 432 (7019), 824–828.
143. Reymond J.-L., Ruddigkeit L., Blum L., van Deursen R.: The enumeration of chemical space. *WIREs Comput. Mol. Sci.* **2012**, 2 (5), 717–733.
144. Bohacek R. S., McMartin C., Guida W. C.: The art and practice of structure-based drug design: A molecular modeling perspective. *Med. Res. Rev.* **1996**, 16 (1), 3–50.
145. Ertl P.: Cheminformatics Analysis of Organic Substituents: Identification of the Most Common Substituents, Calculation of Substituent Properties, and Automatic Identification of Drug-like Bioisosteric Groups. *J. Chem. Inf. Comput. Sci.* **2002**, 43 (2), 374–380.
146. Ruddigkeit L., van Deursen R., Blum L. C., Reymond J.-L.: Enumeration of 166 Billion Organic Small Molecules in the Chemical Universe Database GDB-17. *J. Chem. Inf. Model.* **2012**, 52 (11), 2864–2875.
147. Bolton E. E., Wang Y., Thiessen P. A., Bryant S. H.: Chapter 12 PubChem: Integrated Platform of Small Molecules and Biological Activities. In: Ralph A. Wheeler and David C. Spellmeyer Annual Reports in Computational Chemistry. **2008**, vol 4. Elsevier, S. 217–241
148. Irwin J. J., Sterling T., Mysinger M. M., Bolstad E. S., Coleman R. G.: ZINC: A Free Tool to Discover Chemistry for Biology. *J. Chem. Inf. Model.* **2012**, 52 (7), 1757–1768.
149. American Chemical Society: CAS Fact Sheet. <http://www.cas.org/about-cas/cas-fact-sheets>. **2014**. Zuletzt aufgerufen: 12 Nov 2014.
150. Schneider G., Fechner U.: Computer-based de novo design of drug-like molecules. *Nat. Rev. Drug Discov.* **2005**, 4 (8), 649–663.
151. Lipinski C. A., Lombardo F., Dominy B. W., Feeney P. J.: Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **1997**, 23 (1-3), 3–25.
152. Ghose A. K., Viswanadhan V. N., Wendoloski J. J.: 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. Comb. Chem.* **1998**, 1 (1), 55–68.
153. Veber D. F., Johnson S. R., Cheng H.-Y., Smith B. R., Ward K. W., Kopple K. D.: Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. *J. Med. Chem.* **2002**, 45 (12), 2615–2623.
154. Leeson P.: Drug discovery: Chemical beauty contest. *Nature.* **2012**, 481 (7382), 455–456.

155. Irwin J., Shoichet B. K.: ZINC - A free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **2005**, 45, 177–182.
156. Hilbig M., Urbaczek S., Groth I., Heuser S., Rarey M.: MONA - Interactive manipulation of molecule collections. *J. Cheminf.* **2013**, 5 (1), 38.
157. Urbaczek S., Kolodzik A., Fischer J. R., Lippert T., Heuser S., Groth I., Schulz-Gasch T., Rarey M.: NAOMI. *J. Chem. Inf. Model.* **2011**, 51 (12), 3199–3207.
158. Hernandez M. Z., Cavalcanti, Suellen Melo T., Moreira, Diogo Rodrigo M., de Azevedo Junior, Walter Figueira, Leite, Ana Cristina Lima: Halogen Atoms in the Modern Medicinal Chemistry: Hints for the Drug Design. *Curr. Drug Targets.* **2010**, 11 (3), 303–314.
159. Lu Y., Liu Y., Xu Z., Li H., Liu H., Zhu W.: Halogen bonding for rational drug design and new drug discovery. *Expert Opin. Drug Discov.* **2012**, 7 (5), 375–383.
160. Klebe G.: Virtual ligand screening. *Drug Discov. Today.* **2006**, 11, 580–594.
161. Guedes I., Magalhães C. de, Dardenne L.: Receptor–ligand molecular docking. *Biophys. Rev.* **2014**, 6 (1), 75–87.
162. Brooijmans N., Kuntz I. D.: Molecular recognition and docking algorithms. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, 32, 335–373.
163. Friesner R. A., Banks J. L., Murphy R. B., Halgren T. A., Klicic J. J., Mainz D. T., Repasky M. P., Knoll E. H., Shelley M., Perry J. K., Shaw D. E., Francis P., Shenkin P. S.: Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, 47 (7), 1739–1749.
164. Rarey M., Kramer B., Lengauer T.: Multiple automatic base selection. *J. Comput.-Aided Mol. Des.* **1997**, 11, 369–384.
165. McGann M.: FRED Pose Prediction and Virtual Screening Accuracy. *J. Chem. Inf. Model.* **2011**, 51 (3), 578–596.
166. Abagyan R., Totrov M., Kuznetsov D.: ICM—A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. *J. Comput. Chem.* **1994**, 15 (5), 488–506.
167. Jones G., Willett P., Glen R. C., Leach A. R., Taylor R.: Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, 267 (3), 727–748.
168. Bissantz C., Folkers G., Rognan D.: Protein-Based Virtual Screening of Chemical Databases. 1. Evaluation of Different Docking/Scoring Combinations. *J. Med. Chem.* **2000**, 43 (25), 4759–4767.
169. Cavasotto C. N., Orry A. J. W.: Ligand docking and structure-based virtual screening in drug discovery. *Curr. Top. Med. Chem.* **2007**, 7 (10), 1006–1014.

170. Kokh D. B., Wade R. C., Wenzel W.: Receptor flexibility in small-molecule docking calculations. *WIREs Comput. Mol. Sci.* **2011**, 1 (2), 298–314.
171. Alonso H., Bliznyuk A. A., Gready J. E.: Combining docking and molecular dynamic simulations in drug design. *Med. Res. Rev.* **2006**, 26 (5), 531–568.
172. Craig I. R., Essex J. W., Spiegel K.: Ensemble Docking into Multiple Crystallographically Derived Protein Structures. *J. Chem. Inf. Model.* **2010**, 50 (4), 511–524.
173. Huang S.-Y., Zou X.: Advances and Challenges in Protein-Ligand Docking. *Int. J. Mol. Sci.* **2010**, 11 (8), 3016–3034.
174. Park H., Lee J., Lee S.: Critical assessment of the automated AutoDock as a new docking tool for virtual screening. *Proteins: Struct., Funct., Bioinf.* **2006**, 65 (3), 549–554.
175. Schellhammer I., Rarey M.: TriXX: structure-based molecule indexing for large-scale virtual screening in sublinear time. *J. Comput. Aided Mol. Des.* **2007**, 21 (5), 223–238.
176. Schlosser J., Rarey M.: Beyond the Virtual Screening Paradigm. *J. Chem. Inf. Model.* **2009**, 49 (4), 800–809.
177. Henzler A., Urbaczek S., Hilbig M., Rarey M.: An integrated approach to knowledge-driven structure-based virtual screening. *J. Comput. Aided Mol. Des.* **2014**, 28 (9), 927–939.
178. Schärfer C., Schulz-Gasch T., Hert J., Heinzerling L., Schulz B., Inhester T., Stahl M., Rarey M.: CONFECT. *Chem. Med. Chem.* **2013**, 8 (10), 1690–1700.
179. Urbaczek S., Kolodzik A., Rarey M.: The Valence State Combination Model: A Generic Framework for Handling Tautomers and Protonation States. *J. Chem. Inf. Model.* **2014**, 54 (3), 756–766.
180. Bietz S., Urbaczek S., Schulz B., Rarey M.: Protoss: a holistic approach to predict tautomers and protonation states in protein-ligand complexes. *J. Cheminf.* **2014**, 6, 12.
181. Kirchmair J., Markt P., Distinto S., Wolber G., Langer T.: Evaluation of the performance of 3D virtual screening protocols. *J. Comput.-Aided Mol. Des.* **2008**, 22 (3), 213–228.
182. Truchon J.-F., Bayly C. I.: Evaluating Virtual Screening Methods: Good and Bad Metrics for the “Early Recognition” Problem. *J. Chem. Inf. Model.* **2007**, 47 (2), 488–508.
183. Cheng T., Li X., Li Y., Liu Z., Wang R.: Comparative Assessment of Scoring Functions on a Diverse Test Set. *J. Chem. Inf. Model.* **2009**, 49 (4), 1079–1093.
184. Hawkins P. C., Warren G. L., Skillman A. G., Nicholls A.: How to do an evaluation: pitfalls and traps. *J. Comput. Aided Mol. Des.* **2008**, 22 (3-4), 179–190.
185. Korff M. von, Freyss J., Sander T.: Comparison of Ligand- and Structure-Based Virtual Screening on the DUD Data Set. *J. Chem. Inf. Model.* **2009**, 49 (2), 209–231.

186. Venkatraman V., Pérez-Nueno V. I., Mavridis L., Ritchie D. W.: Comprehensive Comparison of Ligand-Based Virtual Screening Tools Against the DUD Data set Reveals Limitations of Current 3D Methods. *J. Chem. Inf. Model.* **2010**, 50 (12), 2079–2093.
187. Mysinger M. M., Carchia M., Irwin J. J., Shoichet B. K.: Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *J. Med. Chem.* **2012**, 55 (14), 6582–6594.
188. Huang N., Shoichet B. K., Irwin J. J.: Benchmarking Sets for Molecular Docking. *J. Med. Chem.* **2006**, 49 (23), 6789–6801.
189. Good A. C., Oprea T. I.: Optimization of CAMD techniques 3. Virtual screening enrichment studies: a help or hindrance in tool selection? *J. Comput. Aided Mol. Des.* **2008**, 22 (3-4), 169-178.
190. Vogel S. M., Bauer M. R., Boeckler F. M.: DEKOIS. *J. Chem. Inf. Model.* **2011**, 51 (10), 2650–2665.
191. Lübbers T., Angehrn P., Gmünder H., Herzig S., Kulhanek J.: Design, synthesis, and structure-activity relationship studies of ATP analogues as DNA gyrase inhibitors. *Bioorg. Med. Chem. Lett.* **2000**, 10 (8), 821–826.
192. East S. P., White C. B., Barker O., Barker S., Bennett J., Brown D., Boyd E. A., Brennan C., Chowdhury C., Collins I., Convers-Reignier E., Dymock B. W., Fletcher R., Haydon D. J., Gardiner M., Hatcher S., Ingram P., Lancett P., Mortenson P., Papadopoulos K., Smee C., Thomaidis-Brears H. B., Tye H., Workman J., Czaplowski L. G.: DNA gyrase (GyrB)/topoisomerase IV (ParE) inhibitors: Synthesis and antibacterial activity. *Bioorg. Med. Chem. Lett.* **2009**, 19 (3), 894–899.
193. Uria-Nickelsen M., Neckermann G., Sriram S., Andrews B., Manchester J. I., Carcanague D., Stokes S., Hull K. G.: Novel topoisomerase inhibitors: microbiological characterisation and in vivo efficacy of pyrimidines. *Int. J. Antimicrob. Agents.* **2013**, 41 (4), 363–371.
194. Trzoss M., Bensen D. C., Li X., Chen Z., Lam T., Zhang J., Creighton C. J., Cunningham M. L., Kwan B., Stidham M., Nelson K., Brown-Driver V., Castellano A., Shaw K. J., Lightstone F. C., Wong S. E., Nguyen T. B., Finn J., Tari L. W.: Pyrrolopyrimidine inhibitors of DNA gyrase B (GyrB) and topoisomerase IV (ParE), Part II: Development of inhibitors with broad spectrum, Gram-negative antibacterial activity. *Bioorg. Med. Chem. Lett.* **2013**, 23 (5), 1537–1543.
195. Xie Y., Chen X., Qin J., Kong X., Ye F., Jiang Y., Liu H., Jiang H., Marmorstein R., Luo C.: Identification and synthesis of N-(thiophen-2-yl) benzamide derivatives as BRAFV600E inhibitors. *Bioorg. Med. Chem. Lett.* **2013**, 23 (8), 2306–2312.
196. Venkataswamy S., Emily T. R., Bret S., Steven W. L., David B. J., Hariprasad V., Sunil S.: High-Throughput Virtual Screening Identifies Novel N'-(1-Phenylethylidene)-benzohydrazides as Potent, Specific, and Reversible LSD1 Inhibitors. *J. Med. Chem.* **2013**, 56 (23), 9496–9508.

197. Park H., Kim S. Y., Kyung A., Yoon T.-s., Ryu S. E., Jeong D. G.: Structure-based virtual screening approach to the discovery of novel PTPMT1 phosphatase inhibitors. *Bioorg. Med. Chem. Lett.* **2012**, 22 (2), 1271–1275.
198. Koch O., Jäger T., Heller K., Khandavalli P. C., Pretzel J., Becker K., Flohé L., Selzer P. M.: Identification of M. tuberculosis Thioredoxin Reductase Inhibitors Based on High-Throughput Docking Using Constraints. *J. Med. Chem.* **2013**, 56 (12), 4849–4859.
199. Axerio-Cilies P., Lack N. A., Nayana M. R. S., Chan K. H., Yeung A., Leblanc E., Guns E. S. T., Rennie P. S., Cherkasov A.: Inhibitors of Androgen Receptor Activation Function-2 (AF2) Site Identified through Virtual Screening. *J. Med. Chem.* **2011**, 54 (18), 6197–6205.
200. Vidler L. R., Filippakopoulos P., Fedorov O., Picaud S., Martin S., Tomsett M., Woodward H., Brown N., Knapp S., Hoelder S.: Discovery of Novel Small-Molecule Inhibitors of BRD4 Using Structure-Based Virtual Screening. *J. Med. Chem.* **2013**, 56 (20), 8073–8088.
201. Pauli I., dos Santos R. N., Rostirolla D. C., Martinelli L. K., Ducati R. G., Timmers L. F. S. M., Basso L. A., Santos D. S., Guido R. V. C., Andricopulo A. D., Norberto de Souza O.: Discovery of New Inhibitors of Mycobacterium tuberculosis InhA Enzyme Using Virtual Screening and a 3D-Pharmacophore-Based Approach. *J. Chem. Inf. Model.* **2013**, 53 (9), 2390–2401.
202. Hamza A., Zhao X., Tong M., Tai H.-H., Zhan C.-G.: Novel human mPGES-1 inhibitors identified through structure-based virtual screening. *Bioorg. Med. Chem.* **2011**, 19 (20), 6077–6086.
203. Ripphausen P., Freundlieb M., Brunschweiger A., Zimmermann H., Müller C. E., Bajorath J.: Virtual Screening Identifies Novel Sulfonamide Inhibitors of ecto-5'-Nucleotidase. *J. Med. Chem.* **2012**, 55 (14), 6576–6581.
204. Murgueitio M. S., Henneke P., Glossmann H., Santos-Sierra S., Wolber G.: Prospective Virtual Screening in a Sparse Data Scenario: Design of Small-Molecule TLR2 Antagonists. *Chem. Med. Chem.* **2014**, 9 (4), 813–822.
205. Huang S.-Y., Grinter S. Z., Zou X.: Scoring functions and their evaluation methods for protein-ligand docking. *Phys. Chem. Chem. Phys.* **2010**, 12 (40), 12899–12908.
206. Moitessier N., Englebienne P., Lee D., Lawandi J., Corbeil C. R.: Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go. *Br. J. Pharmacol.* **2007**, 153 (Suppl 1), S7.
207. Liu J., Wang R.: Classification of Current Scoring Functions. *J. Chem. Inf. Model.* **2015**, 55 (3), 475–482.
208. Huey R., Morris G. M., Olson A. J., Goodsell D. S.: A semiempirical free energy force field with charge-based desolvation. *J. Comp. Chem.* **2007**, 28 (6), 1145–1152.
209. Stahl M., Rarey M.: Detailed Analysis of scoring functions for virtual screening. *J. Med. Chem.* **2001**, 44, 1035–1042.

210. Gohlke H., Klebe G.: Statistical potentials and scoring functions applied to protein–ligand binding. *Curr. Opin. Struct. Biol.* **2001**, 11 (2), 231–235.
211. Wang R., Lu Y., Wang S.: Comparative Evaluation of 11 Scoring Functions for Molecular Docking. *J. Med. Chem.* **2003**, 46 (12), 2287–2303.
212. Rogers D., Hahn M.: Extended-Connectivity Fingerprints. *J. Chem. Inf. Model.* **2010**, 50 (5), 742–754.
213. Daylight Chemical Information Systems, Inc. <http://www.daylight.com>. **2007**. Zuletzt aufgerufen: 14 Jan 2015.
214. Maldonado A. G., Doucet J. P., Petitjean M., Fan B.-T.: Molecular similarity and diversity in chemoinformatics: from theory to applications. *Mol. Divers.* **2006**, 10 (1), 39–79.
215. Xue L., Godden J. W., Bajorath J.: Database Searching for Compounds with Similar Biological Activity Using Short Binary Bit String Representations of Molecules. *J. Chem. Inf. Comput. Sci.* **1999**, 39 (5), 881–886.
216. Gehlhaar D. K., Verkhivker G. M., Rejto P. A., Sherman C. J., Fogel D. R., Fogel L. J., Freer S. T.: Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: conformationally flexible docking by evolutionary programming. *Chem. Biol.* **1995**, 2 (5), 317–324.
217. Gehlhaar D. K., Verkhivker G. M., Rejto P. A., Fogel D. B., Fogel L. J., Freer S. T.: Docking Conformationally Flexible Small Molecules into a Protein Binding Site Through Evolutionary Programming. In: McDonnell J. R., Reynolds R. G., Fogel D. B. Evolutionary Programming IV. Proceedings of the Fourth Annual Conference on Evolutionary Programming. **1995**. MIT Press, Cambridge, S. 615–627
218. BioSolveIT GmbH: LeadIT. 2.1.5. **2013**, St. Augustin.
219. Halgren T. A., Murphy R. B., Friesner R. A., Beard H. S., Frye L. L., Pollard W. T., Banks J. L.: Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J. Med. Chem.* **2004**, 47 (7), 1750–1759.
220. Schrödinger: Epik. 2.9. **2014**. Schrödinger, LLC, New York, NY.
221. Eldridge M. D., Murray C. W., Auton T. R., Paolini G. V., Mee R. P.: Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput. Aided Mol. Des.* **1997**, 11 (5), 425–445.
222. Gohlke H., Hendlich M., Klebe G.: Knowledge-based scoring function to predict protein-ligand interactions. *J. Mol. Biol.* **2000**, 295 (2), 337–356.
223. Neudert G., Klebe G.: DSX: A Knowledge-Based Scoring Function for the Assessment of Protein–Ligand Complexes. *J. Chem. Inf. Model.* **2011**, 51 (10), 2731–2745.

224. Li J., Abel R., Zhu K., Cao Y., Zhao S., Friesner R. A.: The VSGB 2.0 model: A next generation energy model for high resolution protein structure modeling. *Proteins: Struct., Funct., Bioinf.* **2011**, 79 (10), 2794–2812.
225. Willett P., Barnard J. M., Downs G. M.: Chemical Similarity Searching. *J. Chem. Inf. Comput. Sci.* **1998**, 38 (6), 983–996.
226. Butina D.: Unsupervised Data Base Clustering Based on Daylight's Fingerprint and Tanimoto Similarity: A Fast and Automated Way To Cluster Small and Large Data Sets. *J. Chem. Inf. Comput. Sci.* **1999**, 39 (4), 747–750.
227. Hodes L.: Clustering a large number of compounds. 1. Establishing the method on an initial sample. *J. Chem. Inf. Comput. Sci.* **1989**, 29 (2), 66–71.
228. Liao Q., Wang J., Watson I. A.: Accelerating Two Algorithms for Large-Scale Compound Selection on GPUs. *J. Chem. Inf. Model.* **2011**, 51 (5), 1017–1024.
229. Flower D. R.: On the Properties of Bit String-Based Measures of Chemical Similarity. *J. Chem. Inf. Comput. Sci.* **1998**, 38 (3), 379–386.
230. Schrödinger, LLC: LigPrep. 2.7. **2014**, New York, NY.
231. Fischer E.: Einfluss der Configuration auf die Wirkung der Enzyme. *Ber. Dtsch. Chem. Ges.* **1894**, 27 (3), 2985–2993.
232. Kunick C., Lauenroth K., Wieking K., Xie X., Schultz C., Gussio R., Zaharevitz D. W., Leost M., Meijer L., Weber A., Joergensen F. S., Lemcke T.: Evaluation and Comparison of 3D-QSAR CoMSIA Models for CDK1, CDK5, and GSK-3 Inhibition by Paullones. *J. Med. Chem.* **2004**, 47, 22–36.
233. Cherkasov A., Muratov E. N., Fourches D., Varnek A., Baskin I. I., Cronin M., Dearden J., Gramatica P., Martin Y. C., Todeschini R., Consonni V., Kuz'min V. E., Cramer R., Benigni R., Yang C., Rathman J., Terfloth L., Gasteiger J., Richard A., Tropsha A.: QSAR Modeling: Where Have You Been? Where Are You Going To? *J. Med. Chem.* **2014**, 57 (12), 4977–5010.
234. Modi G., Sharma H., Kharkar P. S., Dutta A. K.: Understanding the structural requirements of hybrid (S)-6-((2-(4-phenylpiperazin-1-yl)ethyl)(propyl)amino)-5,6,7,8-tetrahydronaphthalen-1-ol and its analogs as D2/D3 receptor ligands: a 3D QSAR investigation. *Med. Chem. Commun.* **2014**, 5 (9), 1384–1399.
235. Mizuuchi K., O'Dea M. H., Gellert M.: DNA gyrase. *Proc. Natl. Acad. Sci.* **1978**, 75 (12), 5960–5963.
236. Itaya K., Ui M.: A new micromethod for the colorimetric determination of inorganic phosphate. *Clin. Chim. Acta.* **1966**, 14 (3), 361–366.
237. Soyenkoff B.: A Micromethod of Phosphate Determination. *J. Biol. Chem.* **1947**, 168 (2), 447–457.

238. Pettersson L., Andersson I., Oehman L. O.: Multicomponent polyanions. 39. Speciation in the aqueous hydrogen ion-molybdate(MoO_4^{2-})-hydrogenphosphate(HPO_4^{2-}) system as deduced from a combined Emf-phosphorus-31 NMR study. *Inorg. Chem.* **1986**, 25 (26), 4726–4733.
239. Cogan E. B., Birrell G. B., Griffith O. H.: A robotics-based automated assay for inorganic and organic phosphates. *Anal. Biochem.* **1999**, 271 (1), 29–35.
240. Feng J., Chen Y., Pu J., Yang X., Zhang C., Zhu S., Zhao Y., Yuan Y., Yuan H., Liao F.: An improved malachite green assay of phosphate. *Anal. Biochem.* **2011**, 409 (1), 144–149.
241. Freel Meyers, Caren L., Oberthür M., Xu H., Heide L., Kahne D., Walsh C. T.: Characterization of NovP and NovN: Completion of Novobiocin Biosynthesis by Sequential Tailoring of the Noviosyl Ring. *Angew. Chem. Int. Ed.* **2004**, 43 (1), 67–70.
242. Tamura J. K., Gellert M.: Characterization of the ATP binding site on Escherichia coli DNA gyrase. Affinity labeling of Lys-103 and Lys-110 of the B subunit by pyridoxal 5'-diphospho-5'-adenosine. *J. Biol. Chem.* **1990**, 265 (34), 21342–21349.
243. Miller J. R., Herberg J. T., Tomilo M., McCroskey M. C., Feilmeier B. J.: Streptococcus pneumoniae gyrase ATPase: Development and validation of an assay for inhibitor discovery and characterization. *Anal. Biochem.* **2007**, 365 (1), 132–143.
244. Gellert M., Mizuuchi K., O'Dea M. H., Nash H. A.: DNA gyrase: an enzyme that introduces superhelical turns into DNA. *Proc. Natl. Acad. Sci.* **1976**, 73 (11), 3872–3876.
245. Maxwell A., Burton N. P., O'Hagan N.: High-throughput assays for DNA gyrase and other topoisomerases. *Nucl. Acids Res.* **2006**, 34 (15), e104.
246. Aaij C., Borst P.: The gel electrophoresis of DNA. *Biochim. Biophys. Acta, Nucleic Acids Protein Synth.* **1972**, 269 (2), 192–200.
247. Brody J. R., Kern S. E.: History and principles of conductive media for standard DNA electrophoresis. *Anal. Biochem.* **2004**, 333 (1), 1–13.
248. Felsenfeld G., Davies D. R., Rich A.: Formation of a three-stranded polynucleotide molecule. *J. Am. Chem. Soc.* **1957**, 79 (8), 2023–2024.
249. Rich A.: DNA comes in many forms. *Gene.* **1993**, 135 (1–2), 99–109.
250. Burrell M., Burton N., Maxwell A.: A High-Throughput Assay for DNA Topoisomerases and Other Enzymes, Based on DNA Triplex Formation. In: Fox K. R. Drug-DNA Interaction Protocols. **2010**, vol 613. Humana Press, S. 257-266
251. Taylor J. A., Mitchenall L. A., Rejzek M., Field R. A., Maxwell A.: Application of a Novel Microtitre Plate-Based Assay for the Discovery of New Inhibitors of DNA Gyrase and DNA Topoisomerase VI. *PLoS ONE.* **2013**, 8 (2), e58010.

252. Glaser B. T., Malerich J. P., Duellman S. J., Fong J., Hutson C., Fine R. M., Keblansky B., Tang M. J., Madrid P. B.: A High-Throughput Fluorescence Polarization Assay for Inhibitors of Gyrase B. *J. Biomol. Screening*. **2011**, 16 (2), 230–238.
253. Mraheil, M. Abu, Heisig A., Heisig P.: An improved assay for the detection of alterations in bacterial DNA supercoiling *in vivo*. *Pharmazie*. **2013**, 68 (7), 541–548.
254. Rowlands M. G., Newbatt Y. M., Prodromou C., Pearl L. H., Workman P., Aherne W.: High-throughput screening assay for inhibitors of heat-shock protein 90 ATPase activity. *Anal. Biochem*. **2004**, 327 (2), 176–183.
255. Zuck P., O'Donnell G. T., Cassaday J., Chase P., Hodder P., Strulovici B., Ferrer M.: Miniaturization of absorbance assays using the fluorescent properties of white microplates. *Anal. Biochem*. **2005**, 342 (2), 254–259.
256. Miyata Y., Chang L., Bainor A., Mcquade T. J., Walczak C. P., Zhang Y., Larsen M. J., Kirchhoff P., Gestwicki J. E.: High-Throughput Screen for Escherichia coli Heat Shock Protein 70 (Hsp70/DnaK): ATPase Assay in Low Volume by Exploiting Energy Transfer. *J. Biomol. Screening*. **2010**, 15 (10), 1211–1219.
257. Lambert J. H.: Photometria, sive de mensura et gradibus luminis, colorum et umbrae. **1760**. Sumptibus Vidae Eberhardi Klett, Augsburg.
258. Beer A.: Bestimmung der Absorption des rothen Lichts in farbigen Flüssigkeiten. *Ann. Phys*. **1852**, 86 (1), 78–88.
259. Gottschalk P. G., Dunn J. R.: The five-parameter logistic: A characterization and comparison with the four-parameter logistic. *Anal. Biochem*. **2005**, 343 (1), 54–65.
260. R Core Team: R: A Language and Environment for Statistical Computing. <http://www.R-project.org/>. **2013**, Vienna, Austria.
261. Ritz C., Streibig J. C.: Bioassay Analysis Using R. *J. Stat. Soft.* **2005**, 12 (5), 1–22.
262. Hu X., Machius M., Yang W.: Monovalent cation dependence and preference of GHKL ATPases and kinases1. *FEBS Lett*. **2003**, 544 (1), 268–273.
263. Bisswanger H.: Enzyme assays. *Perspect. Sci*. **2014**, 1 (1–6), 41–55.
264. Yuan P., Walt D. R.: Calculation for fluorescence modulation by absorbing species and its application to measurements using optical fibers. *Anal. Chem*. **1987**, 59 (19), 2391–2394.
265. Kromidas S.: Handbuch Validierung in der Analytik. **2000**. WILEY-VCH Verlag GmbH, Weinheim.
266. Hann M. M., Keserü G. M.: Finding the sweet spot: the role of nature and nurture in medicinal chemistry. *Nat. Rev. Drug Discov*. **2012**, 11 (5), 355–365.
267. Elgazzar, A. B. A., Gafaar, A. M., Hafez, H. N., Aly, A. S.: Novel Syntheses and Reactions of Polynuclear Heterocyclic Derivatives Derived From Thioxopyridopyrimidine, With a New Ring System. *Phosphorus, Sulfur, and Silicon and the Related Elements*. **2006**, 181 (8), 1859–1883.

268. Bisacchi G. S., Manchester J. I.: A New-Class Antibacterial—Almost. Lessons in Drug Discovery and Development: A Critical Analysis of More than 50 Years of Effort toward ATPase Inhibitors of DNA Gyrase and Topoisomerase IV. *ACS Infect. Dis.* **2014**, 1 (1), 4–41.
269. Stokes N. R., Thomaidis-Brears H. B., Barker S., Bennett J. M., Berry J., Collins I., Czaplewski L. G., Gamble V., Lancett P., Logan A., Lunniss C. J., Peasley H., Pommier S., Price D., Smees C., Haydon D. J.: Biological Evaluation of Benzothiazole Ethyl Urea Inhibitors of Bacterial Type II Topoisomerases. *Antimicrob. Agents Chemother.* **2013**, 57 (12), 5977–5986.
270. Langer T., Schlatter H., Fasold H.: Evidence that the Novobiocin-Sensitive ATP-Binding Site of the Heat Shock Protein 90 (hsp90) is Necessary for its Autophosphorylation. *Cell Biol. Int.* **2002**, 26 (7), 653–657.
271. Gunaherath, G. M. Kamal B., Marron M. T., Wijeratne, E. M. Kithsiri, Whitesell L., Gunatilaka, A. A. Leslie: Synthesis and biological evaluation of novobiocin analogues as potential heat shock protein 90 inhibitors. *Bioorg. Med. Chem.* **2013**, 21 (17), 5118–5129.
272. Schrödinger: Schrödinger Suite 2011 Protein Preparation Wizard: Epik 2.2; Impact 5.7; Prime 3.0. **2011**. Schrödinger, LLC, New York, NY.
273. Wang J., Wang W., Kollman P. A., Case D. A.: Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graphics Modell.* **2006**, 25 (2), 247–260.
274. Vanquelef E., Simon S., Marquant G., Garcia E., Klimerak G., Delepine J. C., Cieplak P., Dupradeau F.-Y.: R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. *Nucl. Acids Res.* **2011**, 39 (suppl 2), W511-W517.
275. Dupradeau F.-Y., Pigache A., Zaffran T., Savineau C., Lelong R., Grivel N., Lelong D., Rosanski W., Cieplak P.: The R.E.D. tools: advances in RESP and ESP charge derivation and force field library building. *Phys. Chem. Chem. Phys.* **2010**, 12 (28), 7821–7839.
276. Bayly C. I., Cieplak P., Cornell W., Kollman P. A.: A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993**, 97 (40), 10269–10280.
277. Frisch M. J., Trucks G. W., Schlegel H. B., Scuseria G. E., Robb M. A., Cheeseman J. R., Scalmani G., Barone V., Mennucci B., Petersson G. A., Nakatsuji H., Caricato M., Li X., Hratchian H. P., Izmaylov A. F., Bloino J., Zheng G., Sonnenberg J. L., Hada M., Ehara M., Toyota K., Fukuda R., Hasegawa J., Ishida M., Nakajima T., Honda Y., Kitao O., Nakai H., Vreven T., Montgomery J. J. A., Peralta J. E., Ogliaro F., Bearpark M., Heyd J. J., Brothers E., Kudin K. N., Staroverov V. N., Kobayashi R., Normand J., Raghavachari K., Rendell A., Burant J. C., Iyengar S. S., Tomasi J., Cossi M., Rega N., Millam J. M., Klene M., Knox J. E., Cross J. B., Bakken V., Adamo C., Jaramillo J., Gomperts R., Stratmann R. E., Yazyev O., Austin A. J., Cammi R., Pomelli C., Ochterski J. W.,

- Martin R. L., Morokuma K., Zakrzewski V. G., Voth G. A., Salvador P., Dannenberg J. J., Dapprich S., Daniels A. D., Farkas Ö., Foresman J. B., Ortiz J. V., Cioslowski J., Fox D. J.: Gaussian 09 Revision C.01.
278. Jakalian A., Bush B. L., Jack D. B., Bayly C. I.: Fast, efficient generation of high-quality atomic charges. AM1-BCC model. *J. Comp. Chem.* **2000**, 21 (2), 132–146.
279. Jakalian A., Jack D. B., Bayly C. I.: Fast, efficient generation of high-quality atomic charges. AM1-BCC model. *J. Comp. Chem.* **2002**, 23 (16), 1623–1641.
280. Jorgensen W. L., Chandrasekhar J., Madura J. D., Impey R. W., Klein M. L.: Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, 79 (2), 926–935.
281. Darden T., York D., Pedersen L.: Particle mesh Ewald: An N log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, 98 (12), 10089–10092.
282. Miyamoto S., Kollman P. A.: Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **1992**, 13 (8), 952–962.
283. Landrum G.: RDKit: Open-source cheminformatics. **2014**.
284. Berthold M. R., Cebon N., Dill F., Di Fatta G., Gabriel T., Georg F., Meinel T., Ohl P., Sieb C., Wiswedel B.: KNIME: The Konstanz Information Miner; www.knime.org. **2006**.

14 Gefahrstoffverzeichnis

Tabelle 14.1 – Übersicht über die in dieser Arbeit verwendeten Gefahrenstoffe, deren GHS-Kennzeichnung, *Hazard-Statements* (H-Sätze) und *Precautionary-Statements* (P-Sätze).

Gefahrstoff	GHS-Kennzeichnung	H-Sätze	P-Sätze
Acetonitril	Gefahr, 03, 07	H225 H302+H312+H332 H319	P210 P280 P305+P351+P338 P309+P310
Ameisensäure	Gefahr, 05	H314	P280 P301+P330+P331 P305+P351+P338 P309+P310
Citronensäure	Gefahr, 07	H319	P305 + P351 + P338
Dichlormethan	Achtung, 07, 08	H315 H319 H335 H336 H351 H373	P261-P281-P305 + P351 + P338
Dithiothreitol	Achtung, 07	H302 H319 H335 H315	P280 P302+P352 P305+P351+P338 P309+P311
Eisessig	Gefahr, 02, 05	H226 H314	P210 P243 P280 P301+P330+P331 P304+P340 P309+P310
Ethanol	Gefahr, 02, 07	H225 H319	P210-P280-P305 + P351 + P338-P337 + P313- P403 + P235
Ethidiumbromid- Lösung	Gefahr, 06, 08	H331 H341	P261-P281-P311
Ethylendiamin- tetraessigsäure	Achtung, 07	H315 H319 H335	P261 P305 + P351 + P338
Malachitgrün	Gefahr, 05, 07, 08, 09	H361d H302 H318 H410	P201 P281 P273 P305+P351+P338 P309+P310
Methanol	Gefahr, 02, 06, 08	H225 H301+H311+H331 H370	P210 P243 P280 P302+P352 P304+P340 P309+P310
Natriumhydroxid	Gefahr, 05	H290 H314	P280-P303 + P361 + P353-P304 + P340 + P310-P305 + P351 + P338
Novobiocin	Achtung, 07	H317 H319	P280 P305+P351+P338
Salzsäure	Gefahr, 05, 07	H290 H314 H335	P261 P280-P305 + P351 + P338 + P310
Spermidin	Gefahr, 05	H314	P280 P305+P351+P338 P313
Trikaliumphosphat	Gefahr, 05, 07	H318 H335	P280 P304 + P340 + P312 P305 + P351 + P338 + P310

Bei Dichlormethan handelt es sich um eine KMR-Substanz der Kategorie K2, bei Ethidiumbromid um M2. Es wurde mit keiner dokumentationspflichtigen KMR-Substanz gearbeitet.

Danksagung

Eine Arbeit dieses Umfangs kann man nicht alleine schaffen. Daher möchte ich mich bei allen Personen bedanken, die mich in dieser Zeit mit ihrer fachlichen, geistigen und ganz besonders moralischen Kompetenz unterstützt haben.

Herrn Prof. Dr. Duchstein möchte ich für die freundliche Aufnahme in den Arbeitskreis danken. Mein Dank gilt auch Herrn Prof. Torda für die Übernahme des Zweitgutachtens. Ein ganz besonderer Dank gilt Herrn Dr. Lemcke, der mich in diese spannende Thematik des *Molecular Mode(l)ings* einführte, und durch seine vielen fachlichen und privaten Ratschläge, sein Zuhören und das Teilen seiner Erfahrungen eine großartige Unterstützung war – nicht nur für die Arbeit oder der weitreichenden Film- und Kulturhistorie sondern auch für das ganz banale Leben. Ein großer Dank gilt auch allen Arbeitsgruppenmitgliedern: Sebastian, Jan, Börny, Daniel und Pia. Für das private Drumherum, das fachliche Drinhinein, die zahllosen Kartenabenden und die spannenden, erholsamen, spaßigen und natürlich „lehrreichen“ Forschungsreisen. Besonders möchte ich mich bei Bernie für die wenigen Hektoliter geteilten Kaffees, dem Dasein bei Problemen, dem hoffnungslosen Frustschweigen, dem Grillen im Garten, dem Kochen, der gemeinsamen Zeit... bedanken. Dankeschön! Für die Schufferei im Arbeitskreis möchte ich mich neben den Praktikanten auch bei unserem Bachelor Jan und bei Jaime bedanken, deren Arbeit eine große Unterstützung war.

Für die Einführung in die Welt des biochemischen Experimentierens möchte ich Iris danken. Die nicht nur entscheidend dabei half, die Versuche zu planen und durchzuführen, Geräte und Chemikalien anzuschaffen oder eine vernünftige Protokollführung durchzusetzen, sondern auch in der Zeit dazwischen eine lebenswerte und hilfsbereite Freundin war. Ebenso möchte ich der guten Seele des 4. Stockwerks Christine danken. Nicht nur, dass sie sich viel zu häufig um das lästige Drumherum kümmert, sondern sie hat auch für alle anderen Probleme ein offenes Ohr.

Für die Arbeiten rund um die Substanzbibliothek und das virtuelle Screening möchte ich Herrn Prof. Rarey und seiner Arbeitsgruppen danken. Dies gilt im besonderen Maße für Angela und Therese, die mit mir, und auch ohne mich, viele Stunden damit verbrachten, zu verstehen, was ich denke und zu erklären, wie es eigentlich richtig ist. Vielen Dank für die vielen, immer sehr interessanten Gespräche und die Hilfe bei der Durchführung des Screenings! Dies gilt auch für Hille, der für mich wöchentliche neue Mona-Versionen kompilierte und im Notfall einsprang.

Ebenso möchte ich mich bei Herrn Prof. Maison und seinem Arbeitskreis: Franzi, Heike, Ella, Carsten, Verena, Jasper, Fabian, Flo, Steffen, Marcel, Natalija, Serge, Malte, Dorith und Faiza bedanken. Die nicht nur außergewöhnlich guten Kuchen, Kaffee und Unterhaltung, sondern auch Ideen, Denksätze und Freundschaften anzubieten haben. Dies gilt besonders für Saskia, mit der ich zahlreiche Stunden auf Weiterbildungen, dem roten, meist stehenden, Partybus und etliche Tage im Praktikum verbrachte. Mit der man seinen Kummer und seine Freude (Ok, davon gibt es nicht so viel :->) wunderbar teilen kann. Danke.

Aber auch den anderen Arbeitskreisen des Instituts möchte ich danken: Nina, Tamara, Kerstin, Sönke, Mac, Ines, Marten, Andy, Claudia, Robert, Anne, Kym, Marco, HP, Katja, Alex, Phyllis, Marie; Jorge, Karoline und Eva. Mit jedem von euch verbrachte ich auch neben der Uni viele Stunden beim Sport, auf der Assistenten- oder Bierbank. Danke für diese gemeinsame Zeit. *Hakuna Matata!* Sie wird mir immer in Erinnerung bleiben!

Sollten sich in der Arbeit keine Rechtschreibfehler mehr befinden, bedanke ich mich bei meinen fleißigen Korrekturlesern: Börns(!), Peschi, Theres, Anna, Frakli, Sonja, Hasi, Heike, Iris und Phyllis. Befinden sich, wider Erwarten, dennoch Fehler in dieser Arbeit, so sei ihnen verziehen.

Auch abseits der Uni möchte ich meinen Freunden und, vor allem während der Zeit des Schreibens, der MACGKTM danken, die für die nötige Ablenkung und heiteren Gedanken neben der Uni gesorgt haben.

Ein ganz besonderer Dank gilt meiner Familie und ein ganz viel besonderer Dank meinen beiden Eltern, die mich nicht nur in jeder erdenklichen Art auf meinem Weg unterstützt haben, sondern auch in jeder unerdenklichen. Vielen Dank für eure bedingungslose Unterstützung, auch während der schwierigen Zeiten. Ohne euch hätte ich es niemals geschafft! Vielen lieben Dank!

Lebenslauf

entfällt aus datenschutzrechtlichen Gründen.

Eidesstattliche Erklärung

Hiermit versichere ich an Eides statt, die vorliegende Dissertation mit dem Titel:

Strukturbasierte Leitstruktursuche ATP-kompetitiver

Inhibitoren der Gyrase Untereinheit B

selbst verfasst und keine anderen als die angegebenen Hilfsmittel benutzt zu haben. Ich versichere, dass diese Dissertation nicht in einem früheren Promotionsverfahren eingereicht wurde.

Hamburg, den

Unterschrift