

# Kovalente Inhibitoren: Modellierung und Design



*Dissertation zur Erlangung des naturwissenschaftlichen Doktorgrades der  
Julius-Maximilians-Universität Würzburg*

vorgelegt von  
**Erik Endres**  
aus Würzburg

Würzburg 2024



Eingereicht bei der Fakultät für Chemie und Pharmazie am

\_\_\_\_\_

Gutachter der schriftlichen Arbeit

1. Gutachter: \_\_\_\_\_

2. Gutachter: \_\_\_\_\_

Prüfer des öffentlichen Promotionskolloquiums

1. Prüfer: \_\_\_\_\_

2. Prüfer: \_\_\_\_\_

3. Prüfer: \_\_\_\_\_

Datum des öffentlichen Promotionskolloquiums

\_\_\_\_\_

Sämtliche Arbeiten dieser Dissertation wurden im Zeitraum von März 2019 bis März 2024 unter der Aufsicht von Prof. Dr. Christoph Sotriffer am Institut für Pharmazie und Lebensmittelchemie der Julius-Maximilians-Universität Würzburg durchgeführt. Die hier vorgestellten Ergebnisse wurden teilweise in wissenschaftlichen Publikationen veröffentlicht:

1. E. Endres, N. Y. Chen, C. Sotriffer, MD-Based Assessment of Covalent Inhibitors in Noncovalent Association Complexes: Learning from Cathepsin K as a Test Case, *Journal of Chemical Information and Modeling*, **2023**, 63, 10, 3186-3197

2. M. Scheiner, A. Sink, M. Hoffmann, C. Vrigneau, E. Endres, A. Charles, C. Sotriffer, T. Maurice, M. Decker, Photoswitchable Pseudoirreversible Butyrylcholinesterase Inhibitors Allow Optical Control of Inhibition in Vitro and Enable Restoration of Cognition in an Alzheimer's Disease Mouse Model upon Irradiation, *Journal of the American Chemical Society*, **2022**, 144, 7, 3279-3284

Aus Ersterer wurde Abbildung S11 mit Genehmigung übernommen (Copyright 2023 American Chemical Society, <https://doi.org/10.1021/acs.jcim.3c00061>), aus Letzterer Abbildung 3 adaptiert (Copyright 2022 American Chemical Society, <https://doi.org/10.1021/jacs.1c13492>).

Folgende Arbeiten, denen andere Forschungsthemen zugrunde liegen, wurden im genannten Zeitraum ebenfalls publiziert:

3. P. Johe, S. Jung, E. Endres, C. Kersten, C. Zimmer, W. Ye, C. Sönnichsen, U. Hellmich, C. Sotriffer, T. Schirmeister, H. Neuweiler, *ACS Chemical Biology*, **2021**, 16, 4, 661-670

4. M. Scheiner, A. Sink, P. Spatz, E. Endres, M. Decker, Photopharmacology on Acetylcholinesterase: Novel Photoswitchable Inhibitors with Improved Pharmacological Profiles, *ChemPhotoChem*, **2021**, 5, 149-159

5. M. Hoffmann, C. Stiller, E. Endres, M. Scheiner, S. Gunesch, C. Sotriffer, T. Maurice, M. Decker, Highly Selective Butyrylcholinesterase Inhibitors with Tunable Duration of Action by Chemical Modification of Transferable Carbamate Units Exhibit Pronounced Neuroprotective Effect in an Alzheimer's Disease Mouse Model, *Journal of Medicinal Chemistry*, **2019**, 62, 20, 9116-9140

Aus Letzterer wurde Abbildung 2 übernommen und adaptiert (Copyright 2019 American Chemical Society, <https://doi.org/10.1021/acs.jmedchem.9b01012>)

Folgende Publikationen, an denen mitgewirkt wurde, befanden sich zum Zeitpunkt der Ab-

gabe unter Begutachtung:

J. V. Patzke, F. Sauer, R. K. Nair, E. Endres, E. Proschak, V. Hernandez-Olmos, C. Sotriffer, C. Kisker, Structural basis for the bi-specificity of USP25 and USP28 inhibitors, Manuskript eingereicht bei *EMBO Reports* im November 2023

Teile dieser Arbeit basieren auf den Ergebnissen der vorangegangenen Masterarbeit:

E. Endres, MD-Profilierung nichtkovalenter Anlagerungskomplexe von kovalenten Inhibitoren des Cathepsin K, Julius-Maximilians-Universität Würzburg, Fakultät für Chemie und Pharmazie, **2019**

*Now these points of data  
make a beautiful line.  
And we're out of beta.  
We're releasing on time.  
So I'm GLaD I got burned.  
Think of all the things we learned  
for the people who are  
still alive.*

GLaDOS, Portal

## Danksagung

Die Promotion war wohl das bisher größte Projekt meines Lebens, welches sich nun insgesamt über fünf Jahre erstreckt hat. Während dieser Zeit habe ich unzählige Male Unterstützung von allen möglichen Seiten bekommen, und ich kann mir nicht vorstellen, dass ich es ohne diese Hilfe geschafft hätte. Daher möchte ich die Gelegenheit nutzen, mich an dieser Stelle bei all den Menschen zu bedanken.

Allen voran ist hierbei Prof. Dr. Christoph Sotriffer zu nennen, mein Doktorvater, der mich in seinen Arbeitskreis aufgenommen hat und damit diesen Abschnitt meines Lebens ermöglicht hat. Sein fachlicher Rat und seine Hilfsbereitschaft waren beispiellos, und ich bin froh, meine Doktorarbeit bei ihm geschrieben zu haben.

Während den letzten Jahren hatte ich die Gelegenheit, an vielen Projekten mitzuarbeiten und mich mit Forscherinnen und Forschern aus den unterschiedlichsten Fachrichtungen auszutauschen. Ein großes Dankeschön geht daher an meine Kooperationspartner im Rudolf-Virchow-Zentrum (Prof. Dr. Caroline Kisker, Florian Sauer, Jonathan Patzke und Radhika Nair) und im Arbeitskreis von Prof. Dr. Michael Decker (Matthias Hoffmann, Matthias Scheiner, Thomas Zimmermann und Alexandra Sink, sowie natürlich Michael Decker selbst), an Victor Olmos-Hernandez und Ewgenji Proschak aus der Gruppe von Prof. Dr. Stefan Knapp der Goethe-Universität Frankfurt, an Prof. Dr. Tanja Schirmeister und Patrick Johe von der Johannes-Gutenberg-Universität Mainz und an Priv.-Doz. Dr. Hannes Neuweiler. Es war spannend und schön, mit euch zu forschen!

Außerdem möchte ich mich noch bei den Menschen bedanken, mit denen ich wahrscheinlich die meiste Lebenszeit in den letzten Jahren verbracht habe: den vielen Arbeitskreis-Mitgliedern des AK Sotriffer. Mathias „Hiasl“ Diebold und Gerald „Gary“ Keller waren als Systemadministratoren eine unersetzbare Institution unserer Gruppe, und ohne sie wäre mein PC wohl längst in Flammen aufgegangen. Es war mir eine Freude, mit Carina Herbst, Stephan Böhler und Sara Endres (nicht verwandt, verheiratet oder verschwägert mit mir) das Praktikum des zweiten Pharmazie-Semesters zu betreuen. Zusammen haben wir so manche kleine (oder große) Katastrophe gemeistert, und ein bisschen vermisse ich die Zeit mit euch im Labor (aber wahrscheinlich nur, weil ich euch bereits vermisse). Tina Gehrig war meine Leidensgenossin im Kampf mit der kovalenten Bindung; mögen deine Warheads ihr Target nie verfehlen und deine Lysine stets nukleophil genug sein! Auch will ich mich bei dem ehemaligen Trio Joseph Kehrein, Sebastian Bothe und Markus Zehe bedanken, die mich

damals im AK willkommen geheißen und eingearbeitet haben. Von euch habe ich nicht nur fachlich-technische Hilfestellung erhalten, sondern auch in Sachen Humor. Gracias también a Natalia Yuan Chen y a Yesid Ramirez por su apoyo. Madeleine Bathon wünsche ich alles Gute. Herzlichen Dank auch an meine Praktikantin Constanze Weber und an meinen Praktikanten Sebastian Kolle für ihre Unterstützung.

Innerhalb des Instituts gibt es viele fleißige Hände, die den Laden am Laufen halten. Ein herzliches Dankeschön daher auch an Dr. Jens Schmitz, an Dr. Ludwig Höllein, an Lieselotte Möhler-Reitzenstein und an Martina Wecklein-Weidinger.

Außerhalb des Instituts hatte ich viele Kommilitonen und Freunde, die mir geholfen haben, auf andere Gedanken zu kommen und abzuschalten, damit ich nicht wie unsere GPUs durchbrenne. Ein großes Dankeschön für die langen Abende und die kurzweilige Zeit!

Und zuletzt möchte ich mich bei meiner Familie dies- und jenseits bedanken. Mama - danke und tut mir leid, dass ich deine Nerven geraubt habe als meine ausgingen. Danke an meine Brüder, die ohne Fragen zu stellen helfen, wann immer nötig. Ich vermisse dich, Papa. Und ich hoffe wirklich, dass es euch nun besser geht, Oma und Opa.





# Inhaltsverzeichnis

<b>1. Einleitung</b>	<b>5</b>
1.1. Kovalente Inhibitoren . . . . .	5
1.1.1. Entwicklung der kovalenten Inhibitoren . . . . .	5
1.1.2. Kinetische Besonderheiten von kovalenten Inhibitoren . . . . .	8
1.1.3. Computergestützte Methoden zur Beschreibung kovalenter Inhibitoren . . . . .	9
1.2. Targets für kovalente Inhibition . . . . .	11
1.2.1. Papainähnliche Cysteinproteasen . . . . .	11
1.2.2. Butyrylcholin-Esterase (BChE) . . . . .	14
1.2.2.1. Photoswitches zur Untersuchung der BChE . . . . .	17
1.2.3. Deubiquitinasen . . . . .	17
1.2.3.1. USP25 und USP28 . . . . .	18
1.2.3.2. <i>Chlamydia</i> -Deubiquitinase 1 (ChlaDUB1) . . . . .	23
<b>2. Zielsetzung</b>	<b>24</b>
<b>3. Methoden</b>	<b>26</b>
3.1. Molekulares Docking . . . . .	26
3.1.1. Lamarckscher Genetischer Algorithmus und AutoDock . . . . .	28
3.1.2. WIDOCK-Protokoll . . . . .	31
3.2. MD-Simulationen . . . . .	32
3.2.1. Kraftfelder . . . . .	33
3.2.2. Bewegungsgleichungen . . . . .	34
3.2.3. Wassermodelle . . . . .	34
3.2.4. SHAKE-Algorithmus . . . . .	35
3.2.5. Beschreibung der Partialladungen . . . . .	35
3.2.6. Simulationsboxen und periodische Randbedingungen (PBC) . . . . .	36
3.2.7. Ewald-Summation . . . . .	37
3.2.8. Strukturaufbereitung und Equilibrierung . . . . .	37
3.2.9. Analyse von MD-Simulationen . . . . .	38
3.2.9.1. Root Mean Square Deviation (RMSD) . . . . .	38

3.2.9.2.	Principal Component Analysis (PCA)	39
3.3.	Quantenmechanische Methoden	41
3.3.1.	Hartree-Fock-Theorie und der Self-Consistent-Field-Ansatz	41
3.3.2.	Dichtefunktionaltheorie (DFT)	43
3.3.3.	Geometrieoptimierung und Frequenzberechnung	45
3.3.4.	Intrinsic-Reaction-Coordinate-Berechnungen	46

## 4. Ergebnisse 47

4.1.	CatK: Simulation von nichtkovalenten Modellen basierend auf kovalenten Protein-Ligand-Komplexen	47
4.1.1.	Auswahl der Kristallstrukturen	48
4.1.2.	MD-Protokoll	50
4.1.3.	Neu erfolgte Simulationen	53
4.1.3.1.	Verbesserung der 1ATK-Simulation	53
4.1.3.2.	1U9X, ein Arylnitril-Inhibitor	54
4.1.3.3.	Verbesserung der Simulation von 4X6H und 4X6I	55
4.1.3.4.	6QBS, ein Alkin-Inhibitor	60
4.1.3.5.	Zusammenfassung der gesamten Simulationsergebnisse	63
4.1.4.	Analyse des Wasserstoffbrückennetzwerks der katalytischen Einheit	64
4.1.4.1.	Gitter-Ansatz	65
4.1.4.2.	Zeitlich aufgelöster Ansatz	72
4.1.5.	Verlängerte Simulationen	77
4.1.6.	Vergleich zu anderen Cysteinproteasen	80
4.1.6.1.	CatS	80
4.1.6.2.	SARS-CoV2-Main Protease	85
4.1.7.	Simulationen von im Docking vorhergesagter Bindemodi	87
4.1.7.1.	Reproduktion der Ergebnisse der Kristallstrukturen durch Docking	87
4.1.7.2.	Simulation von Hits eines virtuellen Screenings	91
4.1.8.	Kernaussagen der CatK-Studie	96
4.2.	USP28: Strukturbasiertes Design eines kovalenten Inhibitors	101
4.2.1.	Analyse des möglichen Vorgehens anhand von Kristallstrukturen	101
4.2.2.	Strukturvorschläge für mögliche kovalente Inhibitoren	106
4.2.3.	Methodische Details der Docking-Protokolle	107

4.2.4.	Untersuchung der Möglichkeiten zur Entwicklung eines kovalenten Inhibitors basierend auf AZ1 . . . . .	111
4.2.4.1.	Docking-Ergebnisse der Nitril-Verbindungen . . . . .	111
4.2.4.2.	Docking-Ergebnisse der Alkin-Verbindung . . . . .	121
4.2.4.3.	Docking-Ergebnisse der Epoxide und Aziridine . . . . .	123
4.2.4.4.	Docking-Ergebnisse der sterisch anspruchsvollen Warheads .	127
4.2.4.5.	DFT-basierte Rechnungen zur Reaktivitätsvorhersage . . . .	130
4.2.4.6.	Experimentelle Ergebnisse und Ausblick auf mögliche weitere Vorgehensweisen . . . . .	138
4.2.5.	Untersuchung der Möglichkeiten zur Entwicklung eines verlinkten Inhibitors aus AZ1 und Vismodegib . . . . .	141
4.2.6.	MD-Simulation von USP25- und USP28-Konstrukten . . . . .	144
4.2.6.1.	Methodische Details der MD-Simulationen . . . . .	144
4.2.6.2.	MD-Simulationen der unmutierten USP-Konstrukte . . . . .	146
4.2.6.3.	MD-Simulationen der mutierten Varianten . . . . .	149
4.3.	ChlaDUB1: Abschätzung der Reaktivität verschiedener Nitril-Warheads . . .	153
4.4.	BChE: Nichtkovalentes Docking eines kovalent agierenden Photoswitches . .	163
4.4.1.	Stereochemische Analyse des Photoswitches . . . . .	164
4.4.2.	Methodische Details zum Docken eines Photoswitches . . . . .	167
4.4.3.	Ergebnisse des nichtkovalenten Dockings . . . . .	169
 <b>5. Zusammenfassung</b>		 <b>172</b>
 <b>6. Summary</b>		 <b>174</b>
 <b>Literaturverzeichnis</b>		 <b>176</b>
 <b>Abbildungsverzeichnis</b>		 <b>199</b>
 <b>Tabellenverzeichnis</b>		 <b>203</b>
 <b>Abkürzungsverzeichnis</b>		 <b>205</b>
 <b>A. Anhang</b>		 <b>207</b>
A.1.	CatK . . . . .	208
A.1.1.	Gitterbasierte Analysen zu CatK . . . . .	208
A.1.2.	Gitterbasierte Analysen zu CatS . . . . .	209

A.2. USP28 . . . . .	213
A.2.1. Thermodynamische Korrekturen zu den SCF-Energien . . . . .	213
A.3. ChlaDUB1 . . . . .	214
A.3.1. Thermodynamische Korrekturen zu den SCF-Energien . . . . .	214

---

# 1. Einleitung

## 1.1. Kovalente Inhibitoren

Kovalente Inhibitoren zeichnen sich dadurch aus, dass sie mit einer Aminosäure des zu adressierenden Proteins reagieren. Durch die Ausbildung einer chemischen Bindung weist diese Klasse an Wirkstoffen einige Merkmale auf, die sie substanziell von klassischen Wirkstoffen (gemeint sind solche, die lediglich über physikalische Wechselwirkungen mit dem Protein interagieren) unterscheidet. In den nachfolgenden Kapiteln werden diese speziellen Eigenschaften herausgearbeitet. Alle dreidimensionalen Abbildungen von Proteinen und Molekülen wurden, wie auch im Rest der Arbeit, mit PyMol erstellt.<sup>1</sup>

### 1.1.1. Entwicklung der kovalenten Inhibitoren

Die Entwicklung der kovalenten Inhibitoren ist gekennzeichnet durch ein stetiges Verschieben von allgemein akzeptierten Richtlinien, was beim Design von Arzneistoffen machbar und rational ist.<sup>2</sup> Vor wenigen Jahrzehnten wurde es als kritisch und abwegig betrachtet, Medikamente mit reaktiven Gruppen zu entwerfen, da befürchtet wurde, diese Gruppen würden promiskuitiv an Proteine binden, und durch diese ungewollten Interaktionen zu verschiedenen Nebenwirkungen wie immunogenen Reaktionen führen. Ironischerweise waren aber zu diesem Zeitpunkt bereits mehrere Blockbuster mit kovalentem Mechanismus auf dem Markt, bei denen erst nachträglich die Wirkweise nachvollzogen wurde. Acetylsalicylsäure, bekannt unter dem Handelsnamen Aspirin, überträgt seine Acetylgruppe auf ein Serin der Cyclooxygenase 1, verhindert dadurch die Umwandlung der Arachidonsäure in Prostaglandine und entfaltet so seine entzündungshemmende Wirkung.<sup>3-5</sup> Penicillin, das bereits 1928 von Alexander Fleming entdeckt wurde, verhindert die Synthese von Zellwänden in Bakterien mittels kovalenter Bindung durch Ringöffnung des  $\beta$ -Lactams.<sup>6</sup> Omeprazol, das breite Anwendung zur Behandlung von gastroösophagealen Reflux-Krankheiten findet, stellte sich als Prodrug für eine reaktive Spezies heraus, die eine Disulfidbrücke mit Protonenpumpen ausbildet und somit die Säureproduktion des Magens hemmt. Vor dem Hintergrund dieser unkritischen, aber massenhaft verwendeten Medikamente ist es nicht verwunderlich, dass sich mit der Zeit der Gedanke des *Targeted Covalent Inhibitor* (TCI) herausbildete.<sup>2</sup> Die zugrundeliegende

Idee für diese Klasse der Liganden ist, einen Warhead zu wählen, der nur ein Mindestmaß an Reaktivität aufweist, sodass Bindungen an off-Targets minimiert werden. Damit trotzdem eine Reaktion mit dem Zielprotein zustande kommt, ist es einerseits notwendig, dass der Ligand vor der Reaktion eine nichtkovalente Bindung eingeht, und andererseits die reaktive Gruppe möglichst optimal im Rahmen dieses Bindemodus auf die Ziel-Aminosäure ausgerichtet ist.<sup>7</sup> Ersteres legt nahe, das Design eines TCIs ausgehend von bekannten, nichtkovalenten Liganden zu starten und anschließend den Warhead an geeigneter Stelle in die Struktur miteinzubinden. Diese *Ligand-first*-Vorgehensweise<sup>2</sup> wurde beispielsweise bei der Entwicklung der *epidermal growth factor receptor* Inhibitoren, kurz EGFR-Inhibitoren angewendet.<sup>8-10</sup> Durch die Ausbildung einer irreversiblen kovalenten Bindung kann der Ligand sich nicht mehr von seinem Zielprotein lösen. Damit wird der biologische Effekt des Moleküls nicht mehr durch dessen Ausscheidung aus dem Organismus limitiert, sondern durch den Umsatz des Zielproteins, da dieses vollständig ab- und neu aufgebaut werden muss, um dessen Funktion uneingeschränkt wiederherzustellen. Dies ermöglicht es, im Vergleich zu traditionellen reversiblen Bindern, mit geringeren Dosen an Wirkstoff lange Wirkzeiten zu erreichen, sodass die Häufigkeit der Einnahme reduziert werden kann. Die Verlängerung der Verweildauer des Liganden durch die kovalente Bindung kann auch genutzt werden, um problematische Bindetaschen, die z. B. sehr solvensexponiert oder flach sind, zu adressieren oder stark bindende Substrate wie z. B. Adenosintriphosphat (ATP) zu verdrängen. Letzteres spielte eine zentrale Rolle bei den EGFR<sup>8-10</sup> - und *Bruton tyrosine kinase*-Inhibitoren, kurz BTK-Inhibitoren.<sup>11,12</sup> Im Gegensatz zu der ursprünglichen Befürchtung, Selektivität durch ungewollte Bindungen zu verlieren, kann diese tatsächlich gesteigert werden, indem innerhalb einer Proteinfamilie schlecht konservierte Aminosäuren als Reaktionspartner genutzt werden.<sup>2,8</sup>

Diese Vorteile veranlasste kürzlich viele Forschungsgruppen dazu, den klassischen *Ligand-first*-Ansatz um die exakt umgekehrte Herangehensweise zu ergänzen: Statt von einem reversiblen Binder zu starten und nachträglich einen Warhead zu inkorporieren, wird beim *Electrophile-first*-Ansatz<sup>2</sup> bewusst bereits beim Start des Designs nach kovalent reagierenden Verbindungen gesucht, z. B. durch das Screening von Moleküldatenbanken, die potenzielle Elektrophile beinhalten, um diese gefundenen Strukturen anschließend weiter zu optimieren. Diese Herangehensweise wird meist durch Assay-Methoden unterstützt, die die Selektivitäten der Elektrophile für unterschiedliche potenzielle Targets innerhalb des kompletten Proteoms wiedergeben können. *Activity-Based Protein Profiling*<sup>13</sup> (ABPP) ist ein typisches Beispiel für ein derartiges Protokoll. Hierbei werden prinzipiell die potenziellen Liganden mit einem Reporter über einen Linker verknüpft, sodass adressierte Proteine aus einer Probe (z.

B. Zell-Lysat) isoliert werden können und z. B. nach Verdauung über Massenspektrometrie quantifiziert werden kann, zu welchem Grad bestimmte Aminosäuren modifiziert wurden.

Die am häufigsten kovalent adressierte Aminosäure ist Cystein. Cystein besitzt eine niedrige natürliche Häufigkeit in Proteinen,<sup>14</sup> sodass off-Target-Bindungen unwahrscheinlich sind. Zeitgleich ist Cystein aufgrund seines elektronenreichen Schwefelatoms ein gutes Nukleophil. Darüber hinaus ist es leicht deprotonierbar oder liegt bereits deprotoniert vor (z. B. in katalytischen Einheiten), was zusätzlich die Reaktivität steigert. Erwähnenswert ist hierbei, dass unfunktionale Cysteine aktiviert werden können, indem durch den Liganden ein Protonenakzeptor in der Nähe des Thiols platziert wird.<sup>15-17</sup> Dies erlaubt selbst in allosterischen Bindetaschen die Nutzung eines kovalenten Inhibitors. Um die Einsetzbarkeit von kovalent agierenden Verbindungen aber zu erweitern, wird es notwendig sein, das Repertoire an standardmäßig adressierbaren Aminosäuren zu vergrößern. Von besonderem Interesse ist hierbei Lysin.<sup>14</sup> In vielen Proteinen ist es auf der Oberfläche zu finden und z. B. für die Ausbildung von Protein-Protein-Wechselwirkungen notwendig. In diesen Fällen liegt es aufgrund der Solvensexposition hauptsächlich protoniert vor und ist somit ein schlechtes Nukleophil. Bei vergrabenen Lysinen jedoch ist es durch die Proteinumgebung möglich, dass der  $pK_S$ -Wert der Lysinseitenkette deutlich gesenkt ist und diese als unprotoniertes und somit nukleophiles primäres Amin vorliegt.<sup>14</sup>

Auch wenn die Ergebnisse der letzten Jahre gezeigt haben, dass irreversibel reagierende Gruppen einen Liganden nicht *a priori* disqualifizieren, so bleiben doch einige Nachteile durch die dauerhafte kovalente Verknüpfung. Die modifizierte Aminosäure liegt bis zur ihrer Ausscheidung weiterhin in der Zelle vor und könnte prinzipiell in andere Proteine eingebaut werden. Derartig modifizierte Proteine können so potenziell nicht mehr die nötigen oder ungewollt neue Wechselwirkungen eingehen, was zu immunogenen Reaktionen führen könnte. Darüber hinaus sind durch die langfristige Wirkung von irreversiblen Inhibitoren leichter Wechselwirkungen mit anderen Medikamenten möglich. Um einen Kompromiss zwischen den Vorteilen der irreversiblen kovalenten und den klassischen nichtkovalenten reversiblen Wirkstoffen zu bilden, wird auf kovalent-reversibel bindende Warheads zurückgegriffen, z. B. Nitrile oder Aldehyde. Diese bieten immer noch viele Vorzüge der Ausbildung der chemischen Bindung, wie etwa die verlängerte Residenzzeit am Ziel, während zeitgleich die Risiken von reaktiven Gruppen minimiert werden. Das 2019 auf dem Markt erschienene Voxelotor, das als erstes Medikament die Polymerisation des Hämoglobins in Sichelzellen-Blutkörperchen verhindert, nutzt einen kovalent-reversiblen Mechanismus und kann so in hohen Dosen, 1500 mg täglich, gefahrlos verabreicht werden.<sup>18</sup>

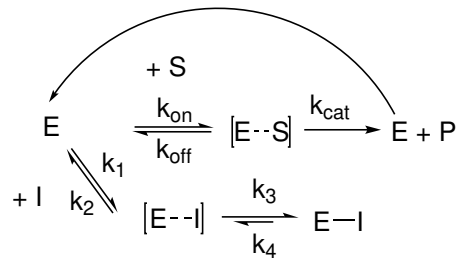


Abbildung 1.1.: Gleichgewichte zwischen Enzym E, kovalentem Inhibitor I und Substrat S. Jeder Schritt besitzt eine individuelle Geschwindigkeitskonstante.

### 1.1.2. Kinetische Besonderheiten von kovalenten Inhibitoren

Die besondere Kinetik von kovalenten Inhibitoren erschwert ihre experimentelle Auswertung, beschert ihnen aber zugleich interessante Anwendungsmöglichkeiten.

Abbildung 1.1 zeigt schematisch die unterschiedlichen Gleichgewichte, die zwischen dem Enzym E, dem Substrat S, welches umgesetzt wird zum Produkt P, und dem kovalenten Inhibitor I existieren können (unter der Annahme, dass S und I um dieselbe Bindestelle konkurrieren).  $k_{\text{on}}$  und  $k_1$  stellen hierbei die Geschwindigkeitskonstanten der nichtkovalenten Assoziation für S und I dar,  $k_{\text{off}}$  und  $k_2$  die entsprechenden Geschwindigkeitskonstanten für die Dissoziation. Während das Substrat S mit der Geschwindigkeit  $k_{\text{cat}}$  zum Produkt P umgesetzt wird, ist es für I nach der Bildung des nichtkovalenten Komplexes möglich, mit dem Enzym zu reagieren und eine kovalente Bindung auszubilden. Die Geschwindigkeitskonstante  $k_3$  ist zu diesem Schritt zugehörig. Falls es sich um eine reversible Reaktion handelt, ist die Auflösung der Bindung mit der Geschwindigkeitskonstanten  $k_4$  möglich.

Durch die kovalente Reaktion mit dem Inhibitor kann das Enzym stetig aus dem Gleichgewicht entfernt werden. Dadurch muss es erst in der Zelle resynthetisiert werden, bevor dessen biologische Funktion wieder erfüllt werden kann. Auch wenn für einen kovalent-reversiblen Inhibitor die Rückreaktion via  $k_4$  technisch möglich ist, so ist diese potenziell so langsam, dass das Protein effektiv für seine komplette Lebenszeit blockiert ist und abgebaut wird, bevor es wieder frei vorliegt. Durch diese langfristige Blockade ändert sich die Zeitskala, auf der ein kovalent agierender Wirkstoff arbeitet, dramatisch, da die effektive Wirkdauer nun nicht mehr von der Ausscheidung des Inhibitors aus dem Organismus abhängt, sondern von der Reproduktion des blockierten Enzyms. Zeitgleich kann eine deutlich geringere Menge an Wirkstoff verabreicht werden, da nun nicht mehr dauerhaft eine bestimmte Konzentration innerhalb der Zelle aufrecht erhalten werden muss, um einen biologischen Effekt zu erzielen. Allerdings ist die Quantifizierung der einzelnen Geschwindigkeitskonstanten auf-



grund der Komplexität des Systems deutlich erschwert. Typischerweise wird die mittlere inhibitorische Konzentration ( $IC_{50}$ ) herangezogen, um unterschiedliche Liganden miteinander zu vergleichen. Diese wird bei festgesetzten Konzentrationen an Enzym und Substrat sowie einer bestimmten Inkubationszeit bestimmt und dient als Näherung für die Inhibitionsgleichgewichtskonstante  $K_i$ . Faktisch liegt aber bei kovalent-irreversiblen Inhibitoren kein Gleichgewicht mehr vor, da keine Rückreaktion mehr möglich ist, bzw. stellt sich das Gleichgewicht bei kovalent-reversiblen Inhibitoren zu langsam ein, sodass, thermodynamisch betrachtet, die gemessenen  $IC_{50}$ -Werte eines kovalenten Inhibitors wenig sinnvoll sind. Besonders deutlich wird das, wenn die Zeitabhängigkeit der apparenten  $IC_{50}$ -Werte betrachtet wird. Bei längeren Inkubationszeiten verbessern sich typischerweise die gemessenen  $IC_{50}$ -Werte eines kovalenten Inhibitors, was aber bei Standard-Screenings nicht auffällt, da nur bei einer festgesetzten Inkubationszeit gemessen wird.

Auf eine vollständige Ableitung der hier gültigen Geschwindigkeitsgesetze wird an dieser Stelle verzichtet und auf die Literatur verwiesen.<sup>19-21</sup> Es sollte aber noch erwähnt werden, dass das langsame, aber fest bindende Verhalten von kovalenten Inhibitoren kinetisch ähnlich ist zu Inhibitoren mit starkem Induced-Fit.<sup>19</sup>

### 1.1.3. Computergestützte Methoden zur Beschreibung kovalenter Inhibitoren

Das momentane Modell für den Prozess der kovalenten Bindung besteht aus zwei separierten Schritten (siehe Abbildung 1.1). Zuerst wird eine nichtkovalente Bindung zwischen Ligand und Protein ausgebildet. Diese verläuft analog zur Bindung eines klassischen, reversiblen Liganden und wird hauptsächlich durch nichtkovalente Wechselwirkungen wie Wasserstoffbrückenbindungen und van-der-Waals-Kräfte stabilisiert. Während dieser Phase muss der Warhead bereits korrekt orientiert in der Bindetasche liegen, damit der zweite Schritt der kovalenten Bindung erfolgen kann, die chemische Reaktion mit einer Aminosäure. Abhängig vom Warhead ist diese chemische Bindung irreversibel oder reversibel.

Zur Beschreibung von kovalenten Inhibitoren wurden unterschiedliche Programme entworfen, die an verschiedenen Schritten dieses Modells ansetzen. Für die vollständige und exakte Beschreibung der Reaktion des Warheads mit seinem Target wäre eine quantenmechanische Berechnung des Inhibitors samt Proteinumgebung und Solvens notwendig, was mit einem enormen Rechenaufwand verbunden ist. Daher werden typischerweise für Protokolle zur Modellierung der Bindung von kovalenten Inhibitoren verschiedene Näherungen und Abstraktionen genutzt.<sup>7</sup> Im Folgenden wird sich auf einige ausgewählte Vorgehensweisen konzentriert,

wie sie in häufig genutzten Programmen verwendet werden.

Für die Posenvorhersage von nichtkovalenten Liganden innerhalb einer Bindetasche existieren bereits mehrere Docking-Programme (siehe auch Kapitel 3.1). Prinzipiell ist es daher denkbar, diese Programme zu nutzen, um auch kovalente Inhibitoren zu docken, um den Zustand vor der Reaktion zu beschreiben. So kann z. B. getestet werden, ob der Warhead korrekt in der Nähe des Aminosäuren-Reaktionspartners platziert wird. Einige Docking-Programme (z. B. GOLD<sup>22</sup>) wurden so erweitert, sodass es möglich ist, die bereits kovalent gebundene Form zu docken. Dies vereinfacht das Suchproblem des Docking-Algorithmus immens, da viele Freiheitsgrade des Inhibitors durch die kovalente Bindung an das Protein verloren gehen. Allerdings wird hier davon ausgegangen, dass die Reaktion sicher stattfindet, ohne in Erwägung zu ziehen, ob die Orientierung des Warheads dafür geeignet ist. Es gibt Hinweise darauf, dass eine derartige Docking-Methodik zu schlechteren Screening-Ergebnissen führt als Protokolle, die zuerst den nichtkovalenten Zustand beschreiben und dann eruieren, ob eine Bindung möglich ist,<sup>23</sup> wie es bei der Schrödinger-Software Glide durchgeführt wird.<sup>24</sup> Darüber hinaus muss beachtet werden, dass die Bindungsenergie, die aus der chemischen Reaktion kommt, nicht oder nur als Parameter berücksichtigt wird,<sup>7</sup> d. h. es sollten nur Liganden miteinander verglichen werden, die über eine ähnliche Reaktivität verfügen.

Die Dynamik eines kovalenten Inhibitors kann auch mittels Molekulardynamiksimulationen, kurz MD-Simulationen, vorhergesagt werden (siehe Kapitel 3.2). Soll hierbei die aller Wahrscheinlichkeit nach nicht standardmäßig parametrisierte kovalente Bindung zum Protein modelliert werden, ist es jedoch notwendig, entsprechende Kraftfeldparameter zu entwickeln, z. B. anhand von quantenmechanischen Berechnungen von Torsionspotenzialen.

Wie bereits eingangs erwähnt, sind für energetische Abschätzungen der Reaktionsbarriere quantenmechanische Berechnungen notwendig. Häufig wird hierbei der Einfluss des Proteins und von Wassermolekülen vernachlässigt, um die Anzahl der für die Berechnungen benötigten Atome gering zu halten, da die Rechenzeit von quantenmechanischen Methoden mit der Anzahl an Elektronen skaliert.<sup>25</sup> Die Berechnung von Aktivierungsenergien setzt die Identifizierung von Übergangszuständen voraus, was nicht trivial ist. Eine einfache Abschätzung der Elektrophilie des Warheads basierend auf dem Energieunterschied zwischen dem *Highest Occupied Molecular Orbital* (HOMO) und dem *Lowest Unoccupied Molecular Orbital* (LUMO).<sup>26</sup> Die Ergebnisse dieser Berechnungen können z. B. mit Reaktionsgeschwindigkeiten aus Glutathion-Assays korreliert werden.<sup>26</sup>

## 1.2. Targets für kovalente Inhibition

Im Rahmen dieser Arbeit wurden unterschiedliche, kovalent adressierbare Enzyme untersucht. Im Folgenden werden diese Proteine und Proteinklassen näher dargestellt.

### 1.2.1. Papainähnliche Cysteinproteasen

Cysteinproteasen sind eine Klasse an Enzymen zur Spaltung von Peptidbindungen, die in verschiedenen Spezies und in den unterschiedlichsten Geweben gefunden werden können. Viele von ihnen sind allerdings strukturell sehr ähnlich aufgebaut. Eine wichtige Gruppe bilden hierbei die papainähnlichen Cysteinproteasen, die, wie der Name vermuten lässt, sehr dem in der Papaya gefundenen Papain ähneln. Viele von diesen Enzymen wurden und werden als mögliche Targets in der Wirkstoffentwicklung betrachtet:

- Cathepsin B (CatB), ein Vertreter aus der Familie der Cathepsine, gilt als mögliches Target für eine Krebstherapie. Innerhalb von Krebszellen kann ein gesteigertes Expressionsniveau an CatB festgestellt werden.<sup>27</sup>
- Cathepsin K (CatK), ein weiterer Vertreter aus der Familie der Cathepsine, ist hauptsächlich in den Osteoclasten zu finden und spaltet dort Collagen I, das für die Knochenstruktur benötigt wird. Inhibition von CatK verhindert den Abbau des Collagens, was bei Osteoporose genutzt werden könnte, um den Knochen neue Stabilität zu geben.<sup>28</sup> Der CatK-Inhibitor Odanacatib erreichte bereits klinische Studien der Phase III; dessen weitere Entwicklung wurde dann allerdings aufgrund von Nebenwirkungen eingestellt.<sup>29</sup> Nichtsdestotrotz werden noch heute Inhibitoren für CatK entwickelt.<sup>30</sup>
- Rhodesain und Cruzain sind Cysteinproteasen, die in den Parasiten *Trypanosoma brucei* beziehungsweise in *Trypanosoma cruzi* gefunden werden können. *Trypanosoma cruzi* ist hierbei für die Chagas-Krankheit, *Trypanosoma brucei* für die Afrikanische Trypanosomiasis verantwortlich. Beide Krankheiten gelten als vernachlässigte tropische Krankheiten.<sup>31</sup>
- *Severe Acute Respiratory Syndrome Corona Virus Type 2* (SARS-CoV-2) besitzt neben seiner Hauptprotease eine papainähnliche Protease, die ebenfalls als Target in Erwägung gezogen wurde.<sup>32</sup>

Abbildung 1.2 zeigt die Bindetasche von CatK als beispielhaften Vertreter der papainähnlichen Cysteinproteasen. Cysteinproteasen spalten Peptidbindungen über eine katalytische

Diade, die aus Cystein und Histidin besteht (in Abbildung 1.2 Cys25 und His162). Histidin dient hierbei als Base, die die Abspaltung des Protons des Thiols erleichtern soll. Bei CatK wird die Orientierung des Histidins durch eine Wasserstoffbrückenbindung zu einem Asparagin-Rest fixiert (Asn182 in Abbildung 1.2). Eine Tryptophan- und eine Glutamin-Seitenkette deuten zusätzlich in Richtung der beiden katalytischen Reste. Für letztere wird diskutiert, ob sie, analog zum Oxyanionenloch einer Serinprotease, dabei hilft, den Übergangszustand der katalysierten Reaktion durch Wasserstoffbrückenbindungen zu stabilisieren.<sup>33,34</sup>

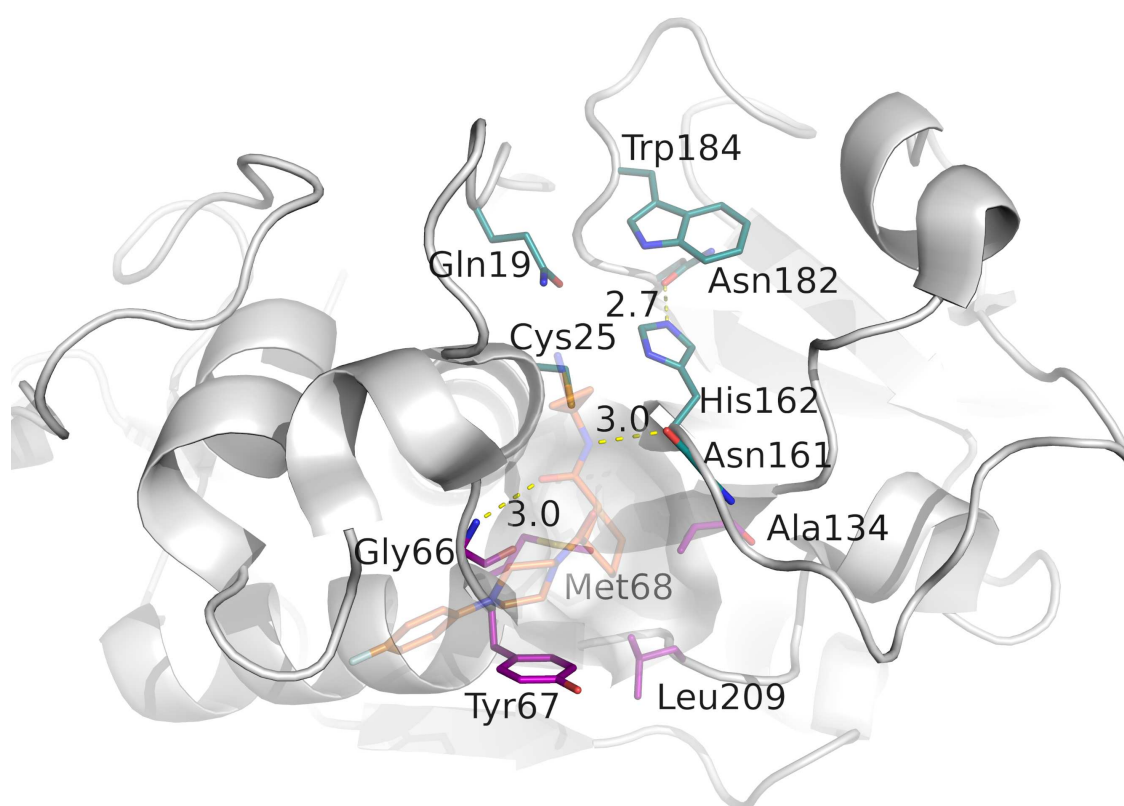


Abbildung 1.2.: CatK und dessen Bindetasche, anhand der PDB 4DMX dargestellt.<sup>35</sup> Der Ligand ist zur Orientierung transparent und orange dargestellt. Aminosäuren in der Nähe der aktiven Einheit sind dunkeltürkis gezeigt. Cys25 und His162 bilden die katalytische Diade. Asn182 stabilisiert His162 über eine Wasserstoffbrückenbindung in dessen Orientierung. Die Seitenketten von Gln19 und Trp184 sind ebenfalls auf die katalytische Einheit ausgerichtet. Tyr67, Met68, Ala134 und Leu209, in der Abbildung lila dargestellt, grenzen eine lipophile Vertiefung ab (die dazugehörige Oberfläche ist hervorgehoben), die für die Selektivität von CatK verantwortlich ist. Gly66 und Asn161 können Wasserstoffbrückenbindungen zu Peptidbindungen ausbilden. Alle Abstände sind in Å angegeben. Für bessere Sichtbarkeit wurde die Seitenkette von Asn161 nicht dargestellt.

Die bisher genannten Aminosäuren sind stark konserviert im Clan der papainähnlichen Cysteinproteasen (für eine Überlagerung verschiedener Strukturen siehe Kapitel 4.1.6.1, Abbildung 4.30). Substrat-Selektivität erreichen diese Cysteinproteasen durch eine benachbarte Tasche, die unterschiedlich große Aminosäuren toleriert. In CatK wird diese Vertiefung durch die Aminosäuren Tyr67, Met68, Ala134 und Leu209 gebildet. Diese lipophile Tasche wird durch Leucin-Seitenketten optimal gefüllt, sodass CatK präferiert nach dieser Aminosäure schneidet.<sup>36</sup>

Häufig wird die intrinsisch gesteigerte Reaktivität des katalytischen Cysteins im Wirkstoffdesign ausgenutzt, um kovalent wirkende Inhibitoren zu verwenden. Hierfür kann ein breites Spektrum an Warheads genutzt werden (siehe, exemplarisch für CatK, Kapitel 4.1.2). Als Design-Grundlage für die Erkennungseinheit des kovalenten Liganden wird sich oft an Substrat-analogen Strukturen orientiert. Das Backbone von Gly66 und Asn161 kann z. B. zeitgleich durch eine Amidbindung eines Liganden adressiert werden, analog zu einer Peptidbindung eines Proteins (siehe Abbildung 1.2).

Der exakte Mechanismus der Cysteinproteasen wird immer wieder in der Literatur diskutiert. Kritisch ist hierbei insbesondere der Protonierungszustand. Cystein- und Histidinseitenketten in Proteinen haben  $pK_S$ -Werte von 6.8 bzw. 6.6,<sup>37</sup> was nahe dem physiologischen pH-Wert liegt. Zusätzlich ist die Proteinumgebung in der Lage,  $pK_S$ -Werte deutlich zu verschieben,<sup>38</sup> sodass die Zuweisung der korrekten (sprich: dominanten) Protonierungszustände oft nicht trivial ist. So sind z. B. für die Cysteinseitenkette in bestimmten Einzelfällen auch  $pK_S$ -Werte von 2.5 bis 11.1 literaturbekannt.<sup>37</sup> Zu diesem Umstand trägt vor allem die hohe Polarisierbarkeit des Cysteins bei.

Die am weitesten verbreitete Charakterisierung des Protonierungszustandes von Papain geht auf Polgár<sup>39</sup> zurück und wird häufig auch auf verwandte Proteine übertragen. Die pH-Abhängigkeit der Ratenkonstanten von Papain deuteten daraufhin, dass zwei ionisierbare Gruppen mit  $pK_S$ -Werten von 4.0 und 8.4 an der Reaktion beteiligt sind. Darüber hinaus konnte keine Verlangsamung der Reaktionsgeschwindigkeit beobachtet werden, als das schwerere  $D_2O$  anstelle von Wasser für das Experiment benutzt wurde. Aus letzterem Ergebnis wurde geschlossen, dass der Mechanismus nicht basenkatalysiert abläuft, sondern die katalytische Diade als reaktives, zwitterionisches Paar vorliegt. Der niedrige  $pK_S$ -Wert von 4.0 wurde dem Histidin zugesprochen (oder genauer: der konjugierten Säure des Histidins) und auf mögliche Interaktionen mit dem benachbarten Tryptophan zurückgeführt, welches das katalytische Histidin vom Solvens schützt. Dieses Modell des reaktiven Ionenpaares wurde für strukturell verwandte Proteine übernommen und ist weit verbreitet in der Literatur.

Neuere kinetische Studien deuten darauf hin, dass das Modell des Ionenpaares unzureichend

ist, um die Katalysefähigkeit der Cysteinproteasen zu erklären,<sup>40,41</sup> und für einige Cysteinproteasen wird in theoretischen Berechnungen der ungeladene Zustand für die katalytische Einheit ermittelt.<sup>42,43</sup> Shokhen *et al.* fanden, basierend auf einem Dichtefunktionaltheorie-Ansatz für die katalytische Einheit in Kombination mit einem impliziten Solvens als Umgebung, dass in freiem Papain das Cystein neutral und das Histidin positiv geladen sein sollte.<sup>44</sup> Für den nichtkovalenten Aldehyd-Papain-Komplex fanden sie hingegen einen ungeladenen Zustand für die katalytische Diade.<sup>45</sup> Mladenovic *et al.* fanden für CatB erneut den zwitterionischen Zustand präferiert, stellten aber fest, dass bereits die Entfernung eines einzigen Wassermoleküls aus der ersten Solvatationshülle des Cysteins ausreichen kann, um den zwitterionischen Zustand deutlich zu destabilisieren.<sup>46</sup> Für *Constant pH MD*-Simulationen, die insbesondere für die Ermittlung von Protonierungszuständen in Abhängigkeit des pH-Wertes gedacht sind, erschwert die hohe Polarisierbarkeit des Schwefelatoms die korrekte Parameterwahl für die benötigten impliziten Solvensmodelle. Daher scheitert diese Methode leicht daran, die experimentell bestimmte pH-Abhängigkeit der Cysteinproteasen zu reproduzieren.<sup>47,48</sup>

Ebenfalls zu betonen ist, dass auf Quantenmechanik (QM) basierten Studien von Papain oder einer verwandten papainähnlichen Cysteinprotease oft ohne einen Liganden innerhalb der Bindetasche arbeiten. Dieser kann aber, durch sein eigenes elektrostatisches Potenzial, Protonierungszustände in der Bindetasche beeinflussen.<sup>38</sup> So zeigten z. B. Rindfleisch *et al.*, dass ein lange bestehendes Paradoxon in der Katalyseeffizienz der Orotidin-5'-Monophosphat-Decarboxylase durch eine Änderung des Protonierungszustandes erklärt werden kann.<sup>49</sup> Die Frage nach „dem korrekten“ Protonierungszustand von papainähnlichen Cysteinproteasen bleibt daher kompliziert. Es erscheint ratsam, in jedem individuellen Fall genau zu untersuchen, welche Auswirkungen die Wahl eines bestimmten Protonierungszustand hat.

### 1.2.2. Butyrylcholin-Esterase (BChE)

Alzheimer-Demenz (AD) ist eine neurodegenerative Erkrankung und stellt die häufigste Form der Demenz dar.<sup>50</sup> Im Verlauf der Erkrankung nimmt die kognitive Leistung des Betroffenen immer weiter ab, bis selbst alltägliche Handlungen nicht mehr möglich sind. Auch wenn der genaue Auslöser der AD noch nicht bekannt ist,<sup>51</sup> so sind mehrere mit ihr einhergehende, physiologische Auffälligkeiten bekannt, die zur neurodegenerativen Wirkung der Krankheit beitragen: die Aggregation von amyloiden Plaques,<sup>52</sup> oxidativer Stress,<sup>53</sup> und Hyperphosphorylierung von  $\tau$ -Proteinen.<sup>54</sup> Darüber hinaus nimmt im Verlauf der AD die cholinerge Aktivität im Gehirn ab.<sup>55,56</sup>

Diese Abnahme an cholinergem Aktivität bildet die Grundlage für den momentan etablierten Therapieansatz. Um die Reduzierung der cholinergen Aktivität zu verlangsamen, wird versucht, die Konzentration des Neurotransmitters Acetylcholin **1** (ACh, siehe Abbildung 1.3) zu erhöhen. Dies wird erreicht, indem die Acetylcholin-Esterase (AChE), die für den Abbau des Neurotransmitters verantwortlich ist, blockiert wird.

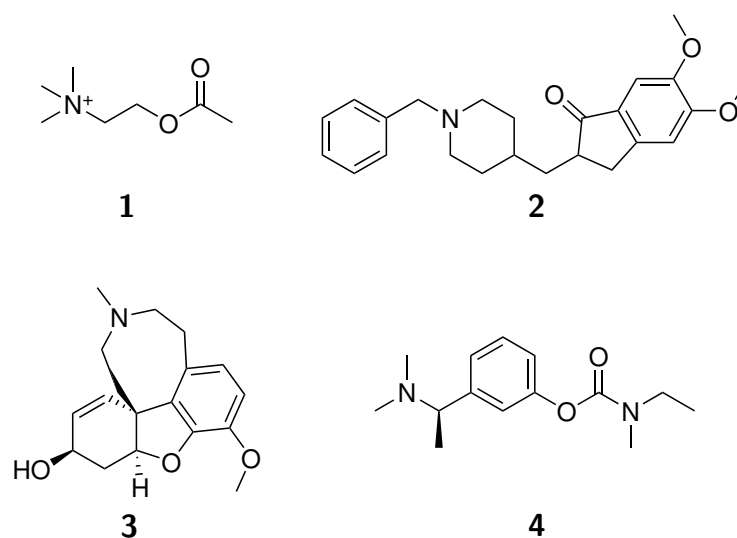


Abbildung 1.3.: Neurotransmitter Acetylcholin (**1**) und für Alzheimer verwendete Acetylcholin-Esterase-Inhibitoren Donepezil (**2**), Galantamin (**3**) und Rivastigmin (**4**).

AChE (siehe Abbildung 1.4a) ist eine Serinhydrolase, die über die klassische katalytische Triade aus Serin, Histidin (in Abbildung 1.4c orange eingefärbt) und Aspartat (in Abbildung 1.4c nicht gezeigt) verfügt. In AChE existiert neben dieser katalytischen Bindestelle am Grund der Tasche eine zweite Bindestelle am Eingangsbereich (in Abbildung 1.4c türkis eingezeichnet), die als periphere anionische Bindestelle bezeichnet wird.<sup>57</sup> Neben der AChE ist allerdings auch die verwandte BChE (siehe Abbildung 1.4b) in der Lage, ACh abzubauen und somit fehlende AChE-Funktion zu kompensieren.<sup>58,59</sup> BChE unterscheidet sich leicht im Aufbau von AChE, z. B. durch kleinere Aminosäuren am Eingang der Bindungstasche, und einer breiteren katalytischen Bindestelle (siehe Abbildung 1.4d).<sup>57</sup> Dadurch kann BChE ein breiteres Spektrum an Substraten hydrolysieren.<sup>50,57</sup>

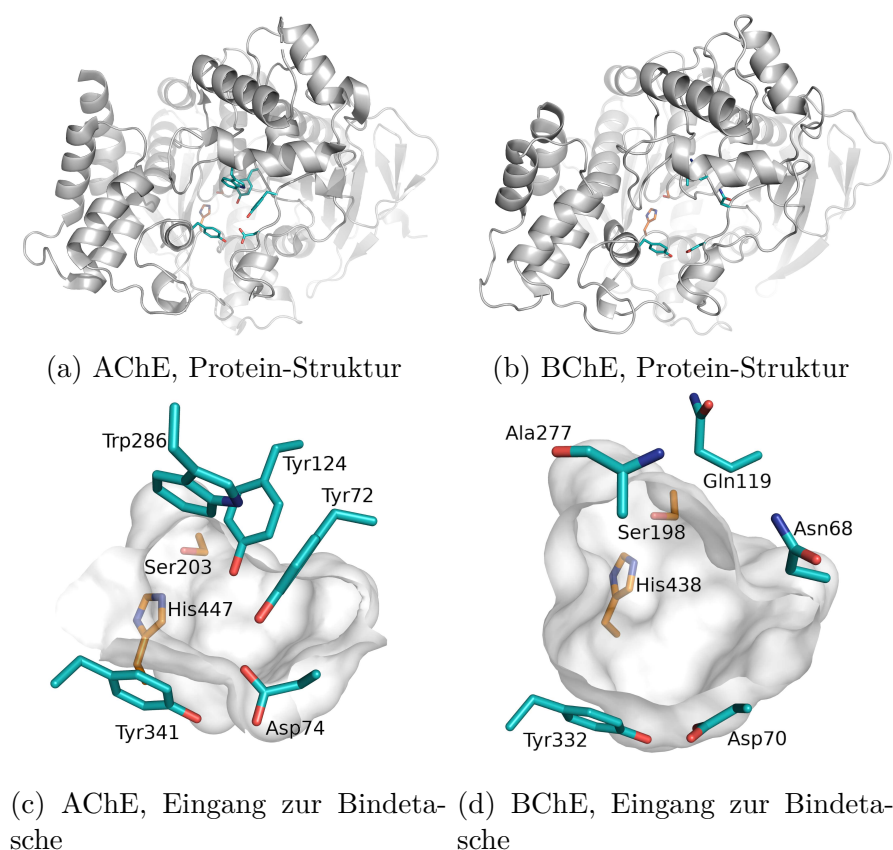


Abbildung 1.4.: Die beiden Proteine AChE und BChE im Vergleich. Die Abbildungen wurden basierend auf PDB 6QAA für BChE<sup>60</sup> und PDB 4EY4 für AChE erstellt.<sup>61</sup> Für AChE wird der Eingang der Bindetasche über sterisch anspruchsvolle Aminosäuren definiert, während in BChE hauptsächlich kleinere, flexiblere Seitenketten vorliegen (in petrol gezeigt). Das katalytisch relevante Serin und Histidin ist am Grund der Bindungstasche in orange gezeigt. Die Abbildungen für die Tascheingänge wurden in leicht abgewandelter Form an anderer Stelle bereits veröffentlicht und hier mit Genehmigung reproduziert.<sup>50</sup> Copyright 2019 American Chemical Society.

Momentan gibt es vier zugelassene AChE-Inhibitoren, die in frühen Stadien der AD Anwendung finden: Donepezil (**2**),<sup>62</sup> Galantamin (**3**)<sup>63</sup> und Rivastigmin (**4**)<sup>64</sup> (siehe Abbildung 1.3). Letzteres zeichnet sich durch einen kovalenten Wirkmechanismus aus, bei dem das Carbamat auf das Serin der AChE übertragen wird und so die Funktion der AChE blockiert wird (siehe Abbildung 1.5). Da die Hydrolyse des Carbamats sehr langsam erfolgt, wird hier von Pseudo-Irreversibilität gesprochen. Rivastigmin ist außerdem die einzige Verbindung, die ebenfalls BChE adressiert.<sup>64</sup> Zusätzliche Inhibitoren der BChE könnten dabei helfen, die cholinerge Aktivität im Verlauf der AD weiter aufrecht zu erhalten.



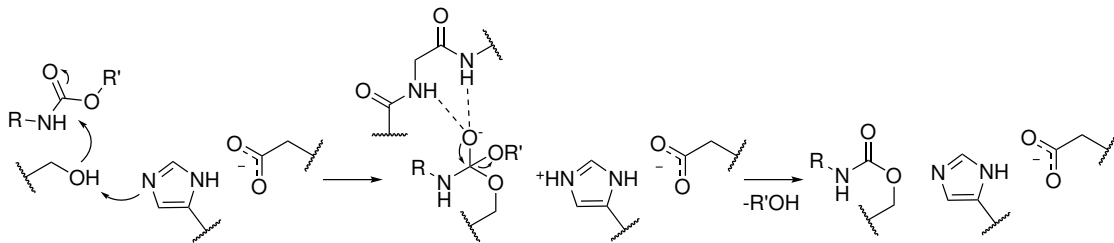


Abbildung 1.5.: Modellhafte Darstellung des Reaktionsmechanismus einer Serinhydrolase mit einem Carbamat.

### 1.2.2.1. Photoswitches zur Untersuchung der BChE

Photoswitches werden genutzt, um einen besseren Einblick in den Bindungsmodus und die Kinetik eines Liganden zu erhalten.<sup>65</sup> Photoswitches sind chemische Verbindungen, die bei Einstrahlung einer bestimmten Wellenlänge isomerisieren und somit ihre Geometrie ändern. Der wohl am häufigsten genutzte Photoswitch-Typ ist das Azobenzol.<sup>66</sup> Isomerisiert das Azobenzol von seiner thermodynamisch stabileren *trans*-Form durch Lichteinstrahlung in die *cis*-Form, nimmt die effektive Länge des Moleküls, und damit der sterische Anspruch, deutlich ab.

Da das neu gebildete Isomer andere physikalische und chemische Eigenschaften aufweist (unterschiedliche Größe, verändertes Dipolmoment etc.) als das ursprüngliche Molekül, kann dieses im Kontext der Enzym-Inhibition auch eine veränderte Inhibition aufweisen. So ist es möglich, lichtgesteuert die Aktivität eines Enzyms über einen Photoswitch zu kontrollieren. Typischerweise wird hierbei in der Anwendung bevorzugt, dass die *cis*-Verbindung die stärkere Inhibition aufweist, sodass mit Hilfe von Licht der Ligand „eingeschaltet“ werden kann.

Mit Hilfe von Substitutionen an den Benzolringen kann zusätzlich der Wellenlängenbereich für den Switch abgestimmt werden. Eine Verschiebung in den roten Bereich des Lichts erlaubt z. B., biologisches Gewebe leichter zu durchdringen.<sup>67</sup>

Photoswitches finden darüber hinaus Anwendung in der Erforschung der AD<sup>65</sup> und deren Zusammenhang zur BChE.<sup>68</sup>

### 1.2.3. Deubiquitinasen

Proteine werden durch die Verknüpfung mit Ubiquitin für den proteasomalen Abbau markiert.<sup>69</sup> Prinzipiell ist diese Verlinkung durch entsprechende Deubiquitinasen (DUBs) auflösbar. Durch das Zusammenspiel des Ubiquitinierungs- und Deubiquitinierungsprozesses kann

der Abbau eines Proteins, und somit auch dessen Prävalenz, sehr genau gesteuert werden. Insofern ist es wenig überraschend, dass das Ubiquitin-Proteasom-System in den Fokus für die Entwicklung von Wirkstoffen rückt, insbesondere in der Krebstherapie.<sup>70</sup> Durch die Deubiquitinierung ist es möglich, dass Onkoproteine nicht ausreichend abgebaut werden und so zur Tumorentstehung und zum Tumorwachstum beitragen. Umgekehrt sollte es daher auch möglich sein, durch Blockade entsprechender DUBs Onkoproteine herunterzuregulieren. Diese Strategie könnte es auch ermöglichen, Proteine abzubauen, die als „undruggable“ gelten und deren Wirkung nicht über klassische Liganden verhindert werden kann, da diese Targets ohne Protein-Interaktionspartner ungeordnet vorliegen.

### 1.2.3.1. USP25 und USP28

USP28 ist eine DUB aus der Familie der Ubiquitin-spezifischen Proteasen (USPs). Diese Cysteinprotease ist in Plattenepithelkarzinomen stark hochreguliert,<sup>71</sup> deubiquitiniert onkogene Substrate der E3-Ubiquitin Ligase SCF(Fbw7) wie c-Myc, Notch1, Cyclin E und HIF1 $\alpha$  in kolorektalen Karzinomen,<sup>72-75</sup> und stabilisiert die Histon-Demethylase LSD1 in Brust- und Magenkrebszellen, was die Tumorigenität steigert.<sup>76</sup> Spezifische USP28-Inhibitoren würden wesentlich dazu beitragen, die Rolle von USP28 in Krebszellen näher zu charakterisieren.

Selektivität wird allerdings durch die hohe strukturelle Ähnlichkeit zu anderen Mitgliedern der USP-Familie erschwert, insbesondere zu USP25. USP25 und USP28 werden in der Literatur<sup>77,78</sup> typischerweise in vier Bereiche unterteilt: Die sogenannten „Finger“, der „Daumen“ und die „Handfläche“ bilden die für die beiden USPs typische katalytische Struktur, an die Ubiquitin (UB) bindet. Letzteres passt aufgrund seiner Form und an der Oberfläche zahlreich vorhandener Lysine gut in die Mulde zwischen Finger, Daumen und Handfläche, in der sich mehrere Aspartat- und Glutamatreste befinden. Durch diese Positionierung deutet der C-Terminus des Ubiquitins auf das katalytische Zentrum dieser Cysteinproteasen.

Die katalytischen Domänen von USP25 und USP28 weisen eine sehr hohe Sequenzidentität auf (57%),<sup>78</sup> insbesondere die Bindestelle zu Ubiquitin (siehe unten). Der vierte Bereich in den USPs ist eine Insertion in diese katalytische Domäne (*USP25/28 catalytic domain inserted domain*, UCID). Diese UCID besteht aus einem *Coiled-Coil*-Motiv, unterteilt in UCID-„Stäbe“, die eine Verbindung zu den „Fingern“ herstellen, und in die UCID-„Spitze“. Der obere Teil der UCID-Stäbe ist für die Oligomerisierung von USP25 und USP28 verantwortlich: USP28- und USP25-Monomere bilden ein Dimer aus, in dem sich zwei Spitzen zu einem Konstrukt anlagern, das vergleichbar zu einem Paar Kirschen geformt ist (siehe Abbildung 1.6a). Diese Dimere sind wie die Monomere biologisch aktiv.<sup>78</sup> Die UCID-Spitze ist allerdings unterschiedlich in den beiden Proteinen aufgebaut. In USP28 liegt die Spitze

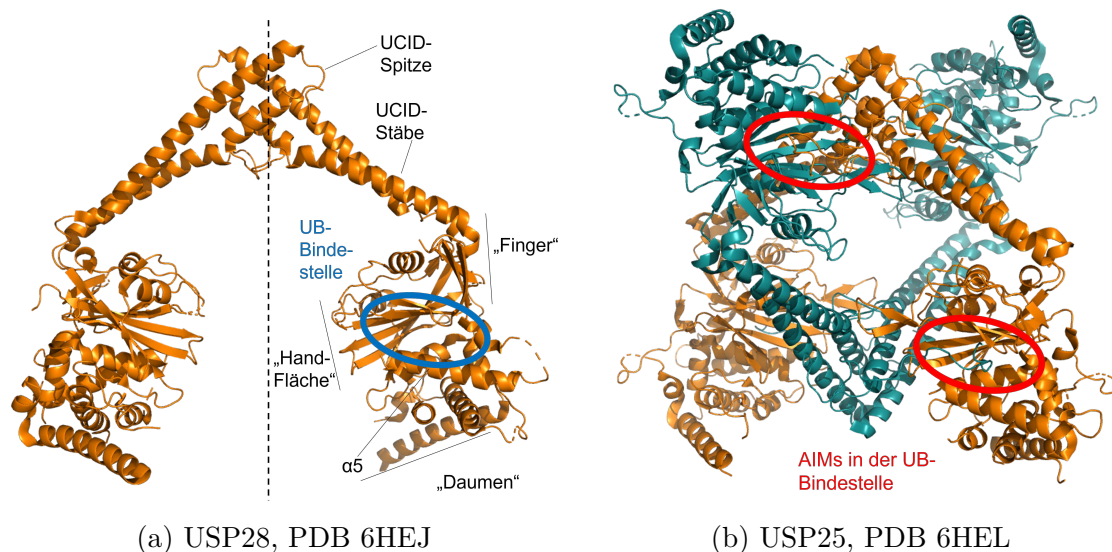


Abbildung 1.6.: Oligomerisierung von USP28 und USP25. Die Tetramerisierung in USP25 wird durch das autoinhibitorische Motiv (AIM) verursacht.

unstrukturiert vor und kann nicht vollständig in der Kristallstruktur beobachtet werden. In USP25 jedoch ragt die UCID-Spitze in eine Furche zwischen der zentralen  $\alpha$ -Helix 5 und einer Reihe an  $\beta$ -Faltblättern, die Teil der „Handfläche“ sind. Dies führt zur Verkeilung zweier USP25-Dimere ineinander und somit zur effektiven Tetramerisierung (siehe Abbildung 1.6b). Auf funktionaler Ebene bedeutet das die Inaktivität von USP25, da die zentrale Bindestelle des Ubiquitins durch die UCID-Spitze blockiert wird. Die entsprechende Aminosäuresequenz wird daher auch als autoinhibitorisches Motiv (AIM) bezeichnet. Mehrere Wasserstoffbrückenbindungen stabilisieren dieses Motiv in der Furche; außerdem wird ein Phenylalanin-Rest (Phe522) tief in eine lipophile Tasche der Bindestelle vergraben. Da USP25 also anscheinend autoinhibiert vorliegt, wird davon ausgegangen, dass die tetramere Form eine Art Speicherform darstellt, bis aktives, nicht-tetrameres USP25 benötigt wird.<sup>78</sup> Bisher ist allerdings nicht bekannt, was genau die Aktivierung von USP25 auslöst. Bisherige Screening-Kampagnen brachten einige Verbindungen hervor, die USP28 und USP25 inhibieren:

- AstraZeneca fand beim Screening einer Datenbank von 40 000 Strukturen vier Verbindungen, die alle dasselbe Grundgerüst eines benzylichen Aminoethanols aufweisen.<sup>70</sup> Abbildung 1.7 zeigt die aktivste dieser vier Verbindungen, AZ1 (5). Die Substanzen waren in der Lage, in Zellen die Konzentration und die Halbwertszeit des Onkogens c-Myc zu modulieren, was das Potenzial dieser Verbindungen unterstreicht. Es konnte allerdings keine Bindestelle identifiziert werden.

- Vismodegib (**6**), ein Inhibitor des Hedgehog-Signaling-Prozesses, der bereits von der *Food and Drug Administration* (FDA) als Arzneimittel gegen Basalzellen-Karzinome zugelassen wurde, kann ebenfalls USP25 und USP28 herunterregulieren.<sup>79</sup> Es wurde versucht, mit Hilfe einer Kombination von Wasserstoff-Deuterium-Austausch - Massenspektrometrie (HDX-MS) und molekularem Docking die Bindestelle von Vismodegib zu charakterisieren. Di Zhou *et al.*<sup>80</sup> gelang es kürzlich, eine Kristallstruktur für Vismodegib im Komplex mit USP28 zu lösen (siehe unten).
- Liu *et al.*<sup>81</sup> fanden eine Reihe an Triazolo-Pyrimidinen, die vergleichbare IC<sub>50</sub>-Werte erreichten wie AZ1. Exemplarisch ist hier **7** gezeigt. Auch hier wurden Docking-Studien durchgeführt, um, basierend auf einer bereits publizierten Kristallstruktur (PDB 6H4I), die Bindestelle zu charakterisieren. Einige Abbildungen legen allerdings nahe, dass nicht aufgelöste Loops innerhalb des Proteins nicht modelliert wurden, da Acetyl-Kappen, die für Protein-Enden genutzt werden, als Interaktionspartner sichtbar sind.<sup>81</sup> So besteht die Möglichkeit, dass hier in eine artifizielle Bindestelle gedockt wurde (wo in Realität sich eigentlich ein sehr flexibler Loop befinden würde).
- Ruiz *et al.*<sup>82</sup> zeigten, dass die Verbindung FT206 (**8**) USP28 und USP25 inhibiert, aber selektiv gegenüber anderen USPs ist. Darüber hinaus war die Verbindung in der Lage, die Protein-Level des Onkogens c-Myc zu senken (vergleichbar zu AZ1).

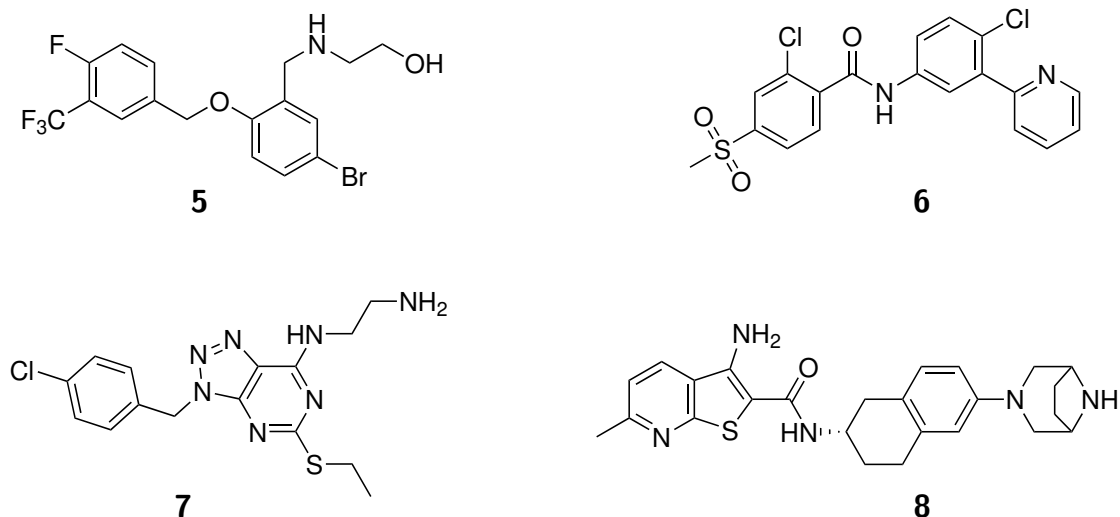


Abbildung 1.7.: Literaturbekannte Inhibitoren von USP28.

Der Arbeitsgruppe um Prof. Kisker ist es gelungen, zum Zeitpunkt der Entstehung dieser Arbeit unveröffentlichte Kristallstrukturen für AZ1 (**5**), Vismodegib (**6**) und FT206 (**8**) im

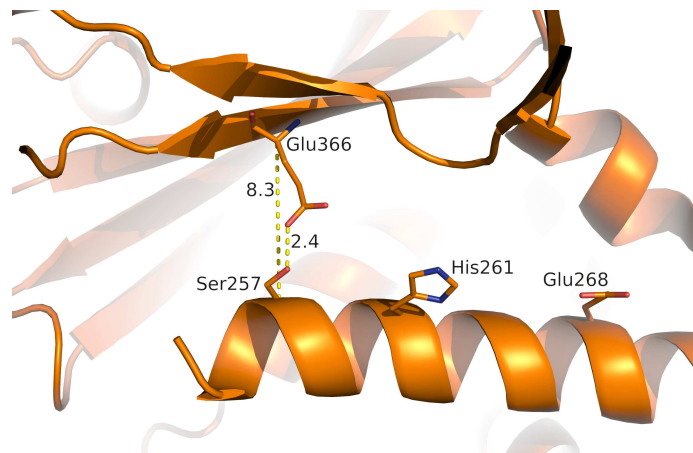
Komplex mit einem USP28-Konstrukt zu lösen, sodass die Frage nach der exakten Bindestelle einiger dieser Inhibitoren geklärt werden kann. Im Rahmen eines strukturbasierten Ligand-Designs werden die Kristallstrukturen von Vismodegib und AZ1 in Kapitel 4.2.1 diskutiert. Abbildung 1.8 zeigt dieselbe Bindestelle zum Vergleich für bereits publizierte Kristallstrukturen. Interessanterweise sind alle drei Verbindungen in der Furche zwischen der zentralen  $\alpha$ -Helix 5 und den  $\beta$ -Faltblättern, analog zum AIM in den tetrameren Strukturen von USP25 (siehe Abbildung 1.8b). Diese Helix ist ebenfalls verschoben im Vergleich zu bisher publizierten Kristallstrukturen (vergleiche Abbildung 1.8c mit Abbildung 1.8a), um die genannte Furche zu vergrößern, ebenfalls in Analogie zu den USP25-Tetrameren. Der gefundene Bindemodus von Vismodegib ist in Übereinstimmung mit den HDX-MS-Ergebnissen von Wang *et al.*,<sup>79</sup> allerdings nicht mit deren Docking-Ergebnissen, da diese nicht den Shift der zentralen  $\alpha$ -Helix berücksichtigten bzw. vorhersagen konnten. Ebenso ist der vorhergesagte Bindemodus von Liu *et al.*<sup>81</sup> inkorrekt.

Die gefundene Bindestelle zeigt großes Potenzial für das Design von USP28/USP25 - Inhibitoren, stellt aber auch ein gewisses Dilemma dar: Es gibt keinen Unterschied zwischen USP25 und USP28 innerhalb dieser Furche, sodass es sehr schwer ist, Selektivität zwischen den beiden USPs innerhalb dieser Bindetasche zu erreichen. Gleichzeitig könnte ein USP28-Inhibitor, der diese Bindestelle als Ziel hat, mit der Tetramerisierung von USP25 interferieren, und somit eine partielle Aktivierung von USP25 bewirken.

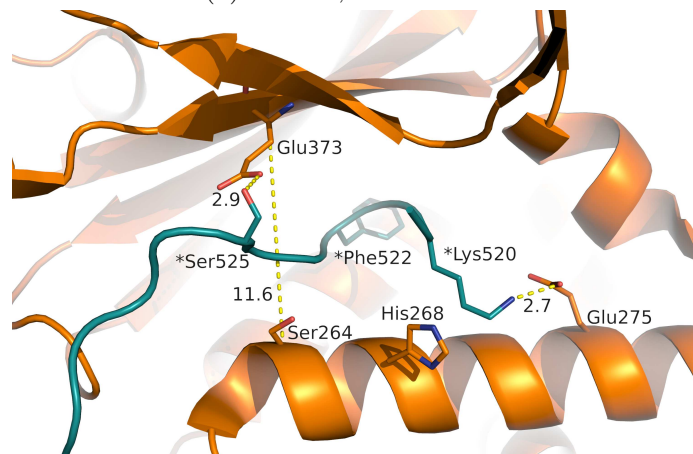
Möglicherweise kann die unterschiedliche Oligomerisierung genutzt werden, um diese Problematik zu lösen. Da im dimeren USP28 die Furche nicht wie im tetrameren USP25 blockiert wird, sollte USP28 kinetisch leichter zugänglich sein und schneller inhibiert werden als USP25. Um auf diese Art und Weise eine Selektivität zu erreichen, müsste der entsprechende Inhibitor einige Kriterien erfüllen:

- hohe Affinität, damit geringe Konzentrationen an Ligand bereits Wirkung zeigen und off-Target-Effekte an USP25 durch unnötig große Mengen an Inhibitor vermieden werden;
- eine irreversible Bindung, damit keine Abdissoziation von USP28 möglich ist und somit zu keinem späteren Zeitpunkt ungewollte USP25-Interaktionen entstehen;
- eine schnelle *On*-Kinetik zum Target, um die Selektivität durch die Oligomerisierung möglichst optimal auszunutzen.

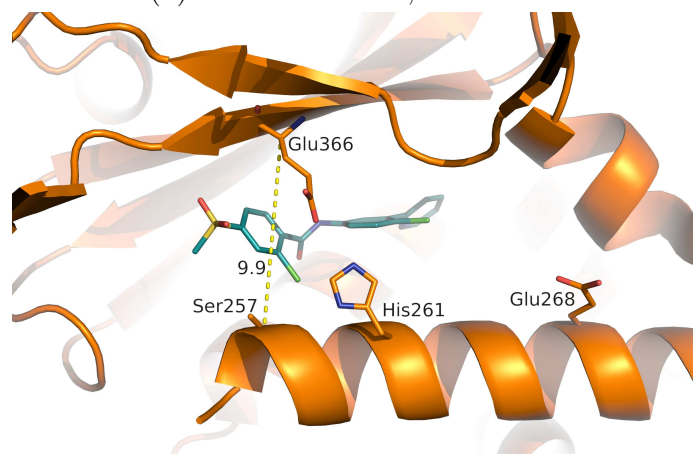
Diese Anforderungen können potenziell von kovalenten Inhibitoren erfüllt werden.



(a) USP28, PDB 6HEJ



(b) USP25 mit AIM, PDB 6H4J



(c) USP28 mit Vismodegib, PDB 8HJE

Abbildung 1.8.: Bindungstasche von USP28 und USP25. Am unteren Bildrand ist die  $\alpha 5$ -Helix erkennbar. Um den Shift der  $\alpha 5$ -Helix deutlich zu machen, wurde der Abstand zwischen den  $\alpha$ -Kohlenstoffatomen von Glu373/366 und Ser264/257 gemessen. Die Abstände sind in Å angegeben.

### 1.2.3.2. *Chlamydia*-Deubiquitinase 1 (ChlaDUB1)

*Chlamydia trachomatis* ist für eine der häufigsten Geschlechtskrankheiten der Welt verantwortlich, die Chlamydien-Infektion.<sup>83,84</sup> Dieses gramnegative Bakterium lebt intrazellulär und produziert Effektor-Proteine, um die Verteidigungsmechanismen der Zelle auszuschalten und so die eigene Reproduktion zu gewährleisten. Einer dieser Verteidigungsmechanismen, auf den infizierte Zellen zurückgreifen können, ist der programmierte Zelltod (Apoptose). Zellen, die mit *Chlamydia trachomatis* infiziert sind, sind deutlich resistenter gegen Apoptose, da in ihnen das Protein Mcl1 stabilisiert wird, welches den Zelltod verhindert.<sup>85</sup> Diese Stabilisierung wird über die vom Bakterium produzierte *Chlamydia*-Deubiquitinase 1 (ChlaDUB1) erreicht.<sup>86</sup> ChlaDUB1 unterbindet den proteasomalen Abbau von Mcl1, verhindert so die Apoptose und trägt damit zum Überleben des Bakteriums innerhalb der Zelle bei. Die Prävalenz von antibiotikaresistenten Erregern ist auch im Kontext der Chlamydien bemerkbar.<sup>87</sup> ChlaDUB1 stellt daher ein potenzielles Target dar, das bisher genutzte Therapie-Strategien nicht adressieren. Außerdem ist ChlaDUB1, wie viele Deubiquitinasen, eine Cysteinprotease, weswegen die kovalente Inhibition des aktiven Zentrums ein naheliegendes Vorgehen ist.<sup>88</sup>

## 2. Zielsetzung

In den vergangenen Jahrzehnten haben kovalente Inhibitoren aufgrund ihrer vielseitigen Vorteile stetig an Bedeutung und Verbreitung gewonnen. Die Modellierung dieser Inhibitoren ist allerdings problematisch: Quantenmechanische Methoden zur vollständigen Beschreibung der Reaktion eines Liganden mit dem Protein sind sehr rechenintensiv, während kovalente Docking-Ansätze nicht die korrekte Orientierung des Warheads innerhalb der Bindetasche untersuchen. Ziel dieser Doktorarbeit ist es, einen Beitrag zur verbesserten Beschreibung von kovalenten Inhibitoren durch computergestützte Methoden zu leisten. Im Rahmen von vier Teilprojekten soll die Reaktivität und der Zustand des Liganden vor der Reaktion in den Mittelpunkt gerückt werden, um ein besseres Design von zukünftigen Verbindungen zu ermöglichen.

Da für den Zustand vor der Reaktion des kovalenten Liganden die experimentelle Datenlage gering ist, soll eine Fallstudie an CatK durchgeführt werden. Für die zahlreichen in der PDB hinterlegten Kristallstrukturen von CatK im Komplex mit einem kovalenten Inhibitor sollen nichtkovalente Modelle generiert und durch MD-Simulationen analysiert werden. Besonderer Fokus soll hierbei auf den Einfluss des Protonierungszustandes der katalytischen Einheit gelegt werden, und wie dieser die Positionierung des Warheads beeinflusst.

Um USP28 bevorzugt gegenüber USP25 zu adressieren, wären hochaffine, schnell agierende Inhibitoren vorteilhaft. Daher sollen in dieser Arbeit ausgehend von unpublizierten Kristallstrukturen verschiedene Design-Strategien verfolgt werden, um dieses Ziel zu erreichen. Besonderes Augenmerk wird hierbei auf die Möglichkeiten der kovalenten Inhibition gelegt, da dieser Ansatz potenziell die Anforderungen am besten erfüllen kann. Durch die Kombinationen von verschiedenen Docking-Protokollen und Rechnungen basierend auf Dichtefunktionaltheorie (DFT) zur Abschätzung der Reaktivität sollen Strukturvorschläge für die Inhibition erarbeitet werden. Zusätzlich sollen molekulardynamische Studien durchgeführt werden, um mögliche Unterschiede in der Beweglichkeit von USP25 und USP28 zu erkennen, die eventuell für eine selektive Adressierung interessant sein könnten.

Für die Inhibitoren der ChlaDUB1 ist die unterschiedliche Reaktivität der Nitril-Warheads ausschlaggebend für die Inhibitionsgeschwindigkeit. Hier soll ein Protokoll basierend auf quantenmechanischen Berechnungen etabliert werden, das erlaubt, die Reaktivität der Liganden relativ zueinander einzustufen zu können. Insbesondere soll ein Vergleich der theoretisch



---

berechneten Reaktivität und den experimentell bestimmten Geschwindigkeitskonstanten erfolgen.

Ein kovalent reagierender Photoswitch, der als Inhibitor der BChE fungiert, zeigt als *cis*-Isomer eine deutlich verbesserte Aktivität als seine *trans*-Form. Hier ist der nichtkovalente Zustand des Liganden von besonderem Interesse, da der Photoswitch noch nicht abgespalten ist. Daher soll hier über nichtkovalentes Docking eine qualitative Erklärung für den Unterschied in der Aktivität der beiden Photoisomere gefunden werden. Für dieses Docking ist es notwendig, eine sinnvolle Beschreibung des Azobenzols und der Stereochemie der Erkennungseinheit zu finden.

## 3. Methoden

Im Folgenden werden die theoretischen Grundlagen der drei wichtigsten Methoden dargestellt, die in dieser Arbeit verwendet wurden. *Molekulares Docking* (Kapitel 3.1) fand in Kapitel 4.4, Kapitel 4.1.7.1 und in Kapitel 4.1.7.2 Anwendung, um Modelle von Ligand-Protein-Komplexen zu generieren. *Molekulardynamik-Simulationen*, kurz auch MD-Simulationen (Kapitel 3.2), werden genutzt, um das dynamische Verhalten von Proteinen und kleinen Molekülen in Lösung zu beschreiben, wie in Kapitel 4.1. *Quantenmechanische Berechnungen* (Kapitel 3.3) werden benötigt, um energetische Unterschiede bei Reaktionen vorherzusagen, was bei den Projekten in Kapitel in Kapitel 4.2.4.5 und 4.3 Anwendung fand.

### 3.1. Molekulares Docking

Molekulares Docking stellt ein wichtiges Werkzeug im Rahmen des Wirkstoffdesigns dar, um die Struktur von Protein-Ligand-Komplexen vorauszusagen. Erste Versuche, mit Hilfe computergestützter Methoden die Pose (d. h. die Konformation und Position) von kleinen Molekülen innerhalb einer Bindetasche vorherzusagen, sind bereits in den 1970er bzw. 1980er Jahren beschrieben. Seitdem sind eine Vielzahl an Docking-Programmen publiziert worden.<sup>22,89-94</sup> Die seit den Entstehungsjahren des Dockings deutlich gestiegenen Rechenkapazitäten erlauben einen immer höheren Durchsatz mit Molekülen in diesen Programmen, was das *virtuelle Screening* von ganzen Molekül-Datenbanken erlaubt.<sup>95</sup>

Prinzipiell setzt sich das Docking dabei aus zwei Schritten zusammen. Zuerst wird eine Pose des Moleküls in der Bindetasche generiert. Hierbei haben sich unterschiedliche Vorgehensweisen bewährt. Systematische Suchalgorithmen generieren und optimieren möglichst viele Konformationen des Liganden in der Bindetasche, um den Konformationsraum möglichst vollständig abzubilden und die energetisch günstigste Pose zu identifizieren. Stochastische Suchen bestehen aus zufälligen Änderungen der Konformation und Position, um Verbesserungen der Pose zu erzielen. Durch eine große Anzahl an Iterationen dieses Vorgehens soll die bestmögliche Pose gefunden werden.

Mit zunehmender Anzahl an Freiheitsgraden innerhalb des Liganden, d. h. mit der Zahl an rotierbaren Bindungen, steigt der zu untersuchende Suchraum exponentiell an. Es wird in diesem Zusammenhang auch von einer *kombinatorischen Explosion* gesprochen.<sup>96</sup> Je größer

der Suchraum ist, desto schwieriger ist es, den Bindemodus korrekt vorherzusagen. Daher muss typischerweise ein Kompromiss zwischen der Genauigkeit der Suche und der damit assoziierten Rechenzeit gefunden werden.

Der zweite Schritt im Rahmen des Dockings ist die Bewertung der erhaltenen Pose mit Hilfe einer Scoring-Funktion. Die thermodynamische Grundlage für diese Bewertung bildet das Gleichgewicht aus Ligand L, Protein P und Protein-Ligand-Komplex [PL] (siehe Gleichung 3.1), das über die Dissoziationskonstante  $K_D$  beschrieben wird (siehe Gleichung 3.2). Laut der Gibbs-Helmholtz-Gleichung (siehe Gleichung 3.3) entspricht diese Dissoziationskonstante einer freien Bindungsenthalpie unter Standardbedingungen,  $\Delta G_D^0$ .<sup>97</sup>



$$K_D = \frac{[P][L]}{[PL]} \quad (3.2)$$

$$\ln K_D = -\frac{\Delta G_D^0}{RT} \quad (3.3)$$

Die Anforderungen an Scoring-Funktionen sind:<sup>98,99</sup>

- *Docking Power*: Posen, die ähnlich zur experimentellen Kristallstruktur sind, sollen besser bewertet werden als solche, die von ihr abweichen.
- *Ranking Power*: Eine Reihe an Liganden soll anhand des Scores gemäß ihrer Affinität sortiert werden können.
- *Scoring Power*: Der Score sollte mit experimentell bestimmten Affinitäten möglichst gut korrelieren.

Prinzipiell existieren unterschiedliche Ansätze zur Parametrisierung von Scoring-Funktionen. *Kraftfeldbasierte* Scoring-Funktionen nutzen klassische Molekularmechanik, um die erhaltenen Docking-Lösungen zu bewerten. *Empirische* Scoring-Funktionen hingegen nutzen einzelne unabhängige Deskriptoren, die so zueinander gewichtet wurden, dass ein Datensatz an experimentellen Bindungsaffinitäten möglichst gut reproduziert wird. *Wissensbasierte* Scoring-Funktionen wählen einen statistischen Ansatz: Die grundlegende Hypothese hier ist, dass, gemäß der Boltzmann-Verteilung, energetisch günstige Abstände für ein Paar aus Atomtypen häufiger in Kristallstrukturen beobachtet werden können als energetisch ungünstige.<sup>99</sup> Mithilfe einer ausreichend großen Datenbank wie z. B. der PDB<sup>100</sup> können so statistische Funktionen zur Bewertung von Liganden parametrisiert werden.

### 3.1.1. Lamarckscher Genetischer Algorithmus und AutoDock

Ein weit verbreiteter Algorithmus, der im Rahmen von Optimierungen genutzt wird, ist der genetische Algorithmus (GA). Die grundlegende Idee hierbei ist, mögliche Lösungen zu einem konkreten Optimierungsproblem als Individuen einer Population aufzufassen. Diese Population soll mehrere der Evolution nachempfundene Iterationen (*Generationen*) mit statistischen Ereignissen durchlaufen, durch welche neue Individuen generiert werden. Am Ende einer Iteration erfolgt eine Bewertung aller Individuen. Schlecht bewertete Individuen sterben daraufhin aus, sodass die Anzahl an Individuen in der Population konstant gehalten wird, während besser bewertete Individuen die nächste Iteration durchlaufen. Dieser Prozess wird entweder für eine festgesetzte Anzahl an Iterationen durchlaufen oder bis eine Konvergenz innerhalb der Population zu beobachten ist. Das bestbewertete Individuum am Ende des GA stellt die gefundene Lösung für das Optimierungsproblem dar. Im evolutionären Bild hat sich dieses Individuum als das am besten angepasste herauskristallisiert.

Im Rahmen des Dockings von Liganden in die Bindetasche eines Proteins stellen mögliche Bindungsposen die Individuen dar, während die zu lösende Frage ist, welche Bindungspose die energetisch günstigste ist. Um eine Bindungspose vollständig zu beschreiben, müssen alle Freiheitsgrade definiert sein, d. h. Position und Orientierung des Liganden und Torsionen aller drehbaren Bindungen, sowohl auf Seite des Liganden als auch auf der Seite des Proteins, falls dessen Beweglichkeit in gewissem Maße berücksichtigt werden soll; sie stellen die „Gene“ des entsprechenden Individuums dar. Die Bewertung von Individuen am Ende jeder Iteration erfolgt unter Zuhilfenahme von Scoring-Funktionen, die in diesem Zusammenhang oft als Fitness-Funktionen bezeichnet werden.

Abhängig vom konkreten Algorithmus variieren die verwendeten Operationen, die die Evolution nachahmen sollen. Die zwei häufigsten sind jedoch bei nahezu jedem GA zu finden (siehe Abbildung 3.1): *Mutation* ist die zufällige Änderung eines Gens eines Individuums, z. B. die zufällige Änderung eines Diederwinkels innerhalb des Moleküls (siehe Abbildung 3.1a). Hierbei wird ein neues Individuum generiert. *Crossover* hingegen stellt den teilweisen Austausch an Genen zwischen zwei Individuen dar, vergleichbar zur geschlechtlichen Vermehrung; z. B. tauschen zwei mögliche Konformationen des Liganden untereinander den Wert eines bestimmten Diederwinkels aus (siehe Abbildung 3.1b). Dabei werden zwei neue Individuen generiert.

Faszinierenderweise ist die Idee eines evolutionär arbeitenden Optimierungsalgorithmus fast so alt wie Computer selbst. Bereits 1950 beschrieb Alan Turing in seiner berühmten Publikation *Computing Machinery and Intelligence*<sup>101</sup> (dieselbe Publikation, in der er den Turing-

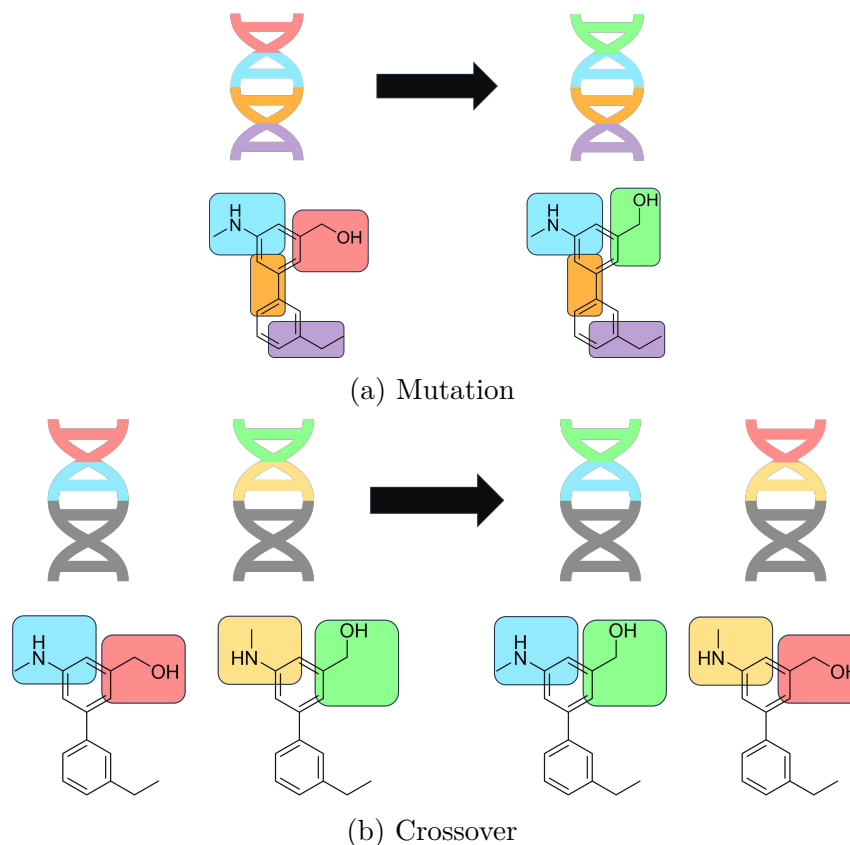


Abbildung 3.1.: Piktographische Darstellung der beiden häufigsten Operationen innerhalb eines genetischen Algorithmus.

Test vorschlag), einen lernenden maschinellen Prozess und zieht Parallelen zur Evolution: „One may hope, however, that this process will be more expeditious than evolution. The survival of the fittest is a slow method for measuring advantages. The experimenter, by the exercise of intelligence, may be able to speed it up. Equally important is the fact that he is not restricted to random mutations.“

Die Entwickler von AutoDock3<sup>93</sup> scheinen ähnliche Gedanken verfolgt zu haben. Sie implementierten eine besondere Variante des GA, die neben der Mutation und dem Crossover ein weiteres, zufällig stattfindendes Ereignis erlaubt: Mit einer gewissen Wahrscheinlichkeit werden Posen nach einer Iteration einer lokalen Energieminimierung unterzogen. Aus evolutionärer Sicht würde das bedeuten, dass die (Protein-)Umgebung in der Lage ist, direkt den genetischen Code zu optimieren - was der widerlegten Evolutionstheorie von Lamarck entspricht. Auch wenn eine so direkte Manipulation des Genoms nicht der Realität entspricht, ist es doch einleuchtend, dass ein derartiger Verlauf der Evolution eine wesentlich schnellere Anpassung an die Umgebung bedeuten würde, da nicht durch zufällige Mutationen und

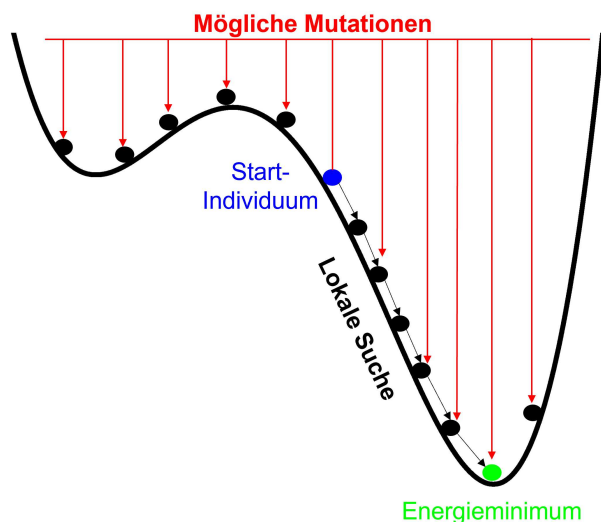


Abbildung 3.2.: Modellhafte Darstellung der Vorteile des LGA. Durch die lokale Suche kann schneller das energetische Minimum gefunden werden als durch Mutation.

„Trial-and-Error“ das passende Gen gefunden werden muss. Diese schnellere Konvergenz überträgt sich in das molekulare Bild des Dockings.

Abbildung 3.2 zeigt ein modellhaftes Potenzial entlang einer beliebigen Koordinate des zu optimierenden Systems. Obwohl die eingezeichnete Startkoordinate bereits nahe dem Energieminimum ist, würde diese nur dann das Minimum im Rahmen eines klassischen GA erreichen, wenn zufällig durch eine Mutation die richtige Koordinate getroffen wird oder durch ein Crossover mit einem geeigneten Individuum die nächste Generation einen besseren Wert erhält. Letzteres setzt voraus, dass der Crossover-Partner selbst irgendwann durch eine Mutation zufällig die richtige Koordinate eingenommen hat. Es leuchtet ein, dass der GA schneller konvergiert, wenn an dieser Stelle eine lokale Suche durchgeführt wird, d. h. den fallenden Werten der Koordinate gefolgt wird. Vor dem Hintergrund dieses Beispiels werden die Vorteile des Lamarckschen Genetischen Algorithmus (LGA) deutlich. Die Mutation bleibt aber dennoch ein wichtiges Ereignis innerhalb des LGA, da mit ihr energetische Barrieren übersprungen werden können und so ein breiteres Sampling möglich ist. Es wurde gezeigt, dass eine Optimierung nach jeder Iteration die Rechenzeit deutlich erhöht, aber kein besseres Ergebnis erzielt, wie wenn lediglich zufällig mit einer gewissen Wahrscheinlichkeit die Optimierung erfolgt, da bei einer Optimierung nach jeder Iteration auch viele nicht zum Ziel führende Lösungen minimiert werden müssen.<sup>93</sup>

Um zusätzlich die Geschwindigkeit des Dockings zu erhöhen, wird in AutoDock das Protein über verschiedene Gitter beschrieben, die bereits vor dem Docking generiert werden.<sup>93</sup> An jedem Gitterpunkt wird eine Sonde einer Atomsorte platziert und die Wechselwirkung mit dem Protein berechnet, sodass im Rahmen des Dockings dieser Wert nicht explizit berechnet werden muss, sondern zwischen zwei Gitterpunkten entlang einer Dimension interpoliert werden kann. Dies wird für jeden Atomtyp innerhalb des Liganden wiederholt. Nichtpolare Wasserstoffatome, die für gewöhnlich keine gezielten Wechselwirkungen ausbilden, werden implizit über die Kohlenstoffatome dargestellt. Der Abstand zwischen den einzelnen Gitterpunkten beträgt hierbei typischerweise 0.375 Å, ungefähr ein Viertel einer C-C-Bindungslänge, sodass eine ausreichende Auflösung des Proteins erreicht wird.<sup>93</sup>

### 3.1.2. WIDOCK-Protokoll

WIDOCK<sup>102</sup> ist eine Weiterentwicklung bereits zuvor publizierter Docking-Protokolle<sup>103,104</sup> für AutoDock4,<sup>94</sup> dem Nachfolger von AutoDock3. Dieses Protokoll soll ein virtuelles Screening von kovalenten Verbindungen ermöglichen. Die grundlegende Idee hierbei ist, zwei neue Atomtypen für die beiden miteinander reagierenden Atome zu definieren. Diese neuen Atomtypen sind identisch zu den ursprünglichen, weisen jedoch zueinander ein modifiziertes Lennard-Jones-Potenzial auf. Der Gleichgewichtsabstand dieses Atom-Paars wird so gewählt, dass er einer kovalenten Bindung entspricht. Das Minimum des Lennard-Jones-Potenzials für z. B. ein Kohlenstoff- und ein Schwefelatom, die miteinander reagieren sollen, wird von 4.0 Å auf 1.8 Å verschoben (vergleiche Abbildung 3.3).

Über die Tiefe des Minimums innerhalb des Lennard-Jones-Potenzials kann der energetische Gewinn der Reaktion dargestellt werden. Auf diese Art und Weise ist es möglich, die unterschiedlichen Reaktivitäten von verschiedenen Warheads innerhalb des Dockings zu berücksichtigen. In der ursprünglichen Veröffentlichung des WIDOCK-Protokolls erfolgt diese Skalierung zwischen 0.175 und 1. Ersteres entspricht etwa dem Minimum eines unmodifizierten Potenzials. Entsprechende Reaktivitäten können aus Experimenten, z. B. Glutathion-Reaktivitätsstudien, oder Berechnungen, z. B. DFT-Rechnungen der Addukt-Bildung, gewonnen werden. Zusätzlich wird die Seitenkette der reaktiven Aminosäure als flexibel definiert, um die korrekte Positionierung des modifizierten Kohlenstoffatoms zu erleichtern.

Die Wechselwirkungen zwischen unreaktiven und reaktiven Atomen bleiben unverändert. Dadurch untersucht dieses Protokoll immer noch den vollen Konformationsraum nichtkovalenter Bindungsposen, d. h., falls es schlicht nicht möglich ist, den Warhead nahe genug am reaktiven Protein-Atom zu platzieren (oder zumindest keine Pose im Rahmen des LGA

gefunden werden kann), wird keine kovalente Bindung erzwungen.

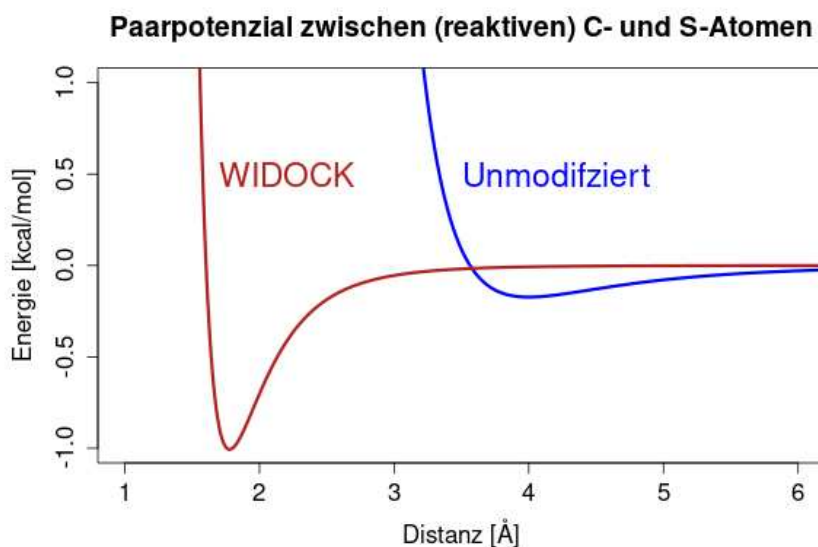


Abbildung 3.3.: Modellhafte Darstellung der Paarpotenziale zwischen reaktivem Kohlenstoff- und Schwefelatom mit den unmodifizierten AutoDock-Parametern und den WIDOCK-Parametern.

Wie bei jedem Modell werden hier Näherungen gemacht, die zu Problematiken führen können. So weist z. B. das genutzte Lennard-Jones-Potenzial kein lokales Maximum auf, das als Reaktionsbarriere genutzt werden kann. Das bedeutet, dass, auch wenn geometrisch eine kovalente Bindung im Rahmen des WIDOCK-Protokolls möglich erscheint, die zugrundeliegende Reaktion möglicherweise immer noch zu langsam abläuft, um im Experiment eine nennenswerte Modifikation der entsprechenden Aminosäure zu zeigen. Darüber hinaus wird im Rahmen dieser Methode nur ein einziges Potential modifiziert, d. h., dass keine Geometrieänderungen durch die Reaktion berücksichtigt werden. Dies ist problematisch für lineare Warheads, bei denen das reaktive Atom aufgrund der benachbarten Atome nicht zugänglich ist (siehe Kapitel 4.2.4).

## 3.2. MD-Simulationen

Um nicht nur die Flexibilität von kleinen Molekülen, sondern auch von Proteinen darzustellen, werden häufig MD-Simulationen genutzt. Im Folgenden werden hierfür die wesentlichen theoretische Grundlagen dargestellt. Darüber hinaus existieren auch spezielle MD-Methoden, um ein besseres Sampling des kompletten Konformationsraums des Proteins zu erreichen (wie



z. B. Simulationen mit beschleunigter Dynamik durch modifizierte Potenziale<sup>105</sup>) und Methoden, die eine bessere Beschreibung der Wechselwirkungen erlauben (wie z. B. Simulationen mit polarisierbaren Kraftfeldern<sup>106</sup> und MD-Simulationen in Kombination mit quantenmechanischen Berechnungen<sup>107</sup>). Für diese weiterführenden Methoden wird auf die Literatur verwiesen.

### 3.2.1. Kraftfelder

Für die Beschreibung von großen, flexiblen Systemen, wie Proteinen in Lösung, wird eine einfache und schnelle Methode benötigt, um Wechselwirkungen und Kräfte zwischen den einzelnen Atomen berechnen zu können. Rein quantenmechanische Ansätze sind hierfür nicht geeignet, da die Rechenzeit mit der Anzahl an benötigten Wellenfunktionen, d. h. effektiv mit der Anzahl an Elektronen des Systems skaliert und sie somit zu rechenintensiv sind.

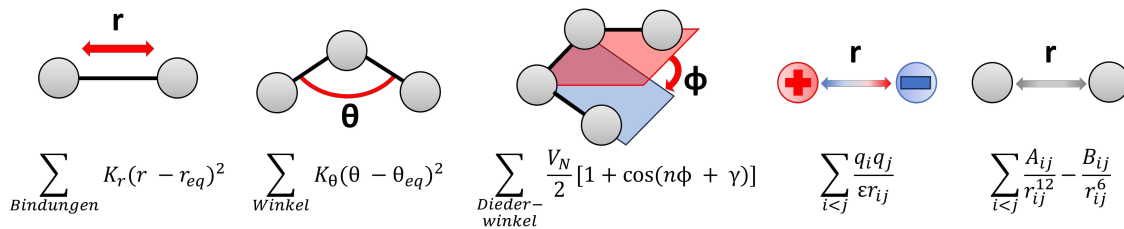


Abbildung 3.4.: Schematische Darstellung einzelner Terme, die in ein Kraftfeld einfließen. Die hier gezeigten Terme orientieren sich an einem *Assisted Model Building and Energy Refinement*-, kurz AMBER-Kraftfeld.<sup>108</sup> Bindungslängen  $r$  und -winkel  $\theta$  werden wie mechanische Federn über ein harmonisches Potenzial mit der Kraftkonstanten  $K_r$  bzw.  $K_\theta$  und den Gleichgewichtsabständen und -winkeln  $r_{eq}$  bzw.  $\theta_{eq}$  beschrieben. Zur Darstellung von Torsionen  $\phi$  werden periodische Kosinusfunktionen verwendet, die über die energetische Barriere  $V_N$ , die Periodizität  $n$  und die Phase  $\gamma$  definiert werden. Elektrostatische Wechselwirkungen werden über das Coulomb-Potenzial abgebildet, wofür die Partialladungen  $q$  und der Abstand  $r$  der jeweiligen Atome  $i$  und  $j$  und die Permittivität  $\epsilon$  benötigt werden. Van-der-Waals-Wechselwirkungen werden über das Lennard-Jones-Potenzial berechnet, wofür für das Atompaar  $i$  und  $j$  spezifische Parameter  $A$  und  $B$  benötigt werden.

Kraftfelder hingegen nutzen keine Wellenfunktionen zur Beschreibung von Atomen, sondern aus der klassischen Physik abgeleitete Terme. So werden z. B. Bindungslängen und -winkel über harmonische Potenziale beschrieben, für elektrostatische Interaktionen wird das Coulomb-Potenzial und für van-der-Waals-Wechselwirkungen das Lennard-Jones-Potenzial genutzt (siehe Abbildung 3.4).<sup>108</sup> Das reduziert den Rechenaufwand dramatisch, auch wenn diese Beschreibung einige Effekte nicht oder nur schlecht berücksichtigt, wie z. B.  $\sigma$ -Löcher (siehe Kapitel 3.2.5).

### 3.2.2. Bewegungsgleichungen

Die Grundlage für MD-Simulationen bilden die Newtonschen Gesetze. Alle Kräfte, die auf ein Atom  $i$  an der Position  $x$  laut der Berechnung des Kraftfeldes wirken, werden aufaddiert und ergeben eine Gesamtkraft  $F(x)$ . Diese Kraft  $F$  lässt sich über die Formulierung der Differentialgleichung des zweiten Gesetzes von Newton beschreiben (Gleichung 3.4):

$$F_i(t) = -\frac{\partial U(x_i(t))}{\partial x_i(t)} = m_i \frac{\partial^2 x_i(t)}{\partial t^2} = m_i \frac{\partial v_i(t)}{\partial t} = m_i a_i(t) \quad (3.4)$$

Eine Änderung der Position des Atoms  $i$  bewirkt eine Änderung der resultierenden Kraft  $F$ , und somit ebenfalls eine neue Geschwindigkeit  $v$ . Für eine MD-Simulation wird nun angenommen, dass diese Kraft  $F$  innerhalb eines kleinen Zeitschritts  $\Delta t$  konstant bleibt. So können über verschiedene Integrations-Algorithmen iterativ immer neue Positionen für das Atom, und dafür resultierende Kräfte und Geschwindigkeiten, berechnet werden. Exemplarisch ist hier der Verlet-Algorithmus gezeigt (Gleichung 3.5 und 3.6),<sup>109</sup> der zusätzlich den vorhergehenden Zeitschritt  $t - \Delta t$  nutzt, um die neue Position und die momentane Geschwindigkeit zu berechnen:

$$x_i(t + \Delta t) = 2x_i(t) - x_i(t - \Delta t) + \Delta t^2 a_i(t) \quad (3.5)$$

$$v_i(t) = \frac{x_i(t + \Delta t) - x_i(t - \Delta t)}{2\Delta t} \quad (3.6)$$

Die Wahl des entsprechenden Zeitschritts ist hierbei kritisch. Ein zu kleiner Zeitschritt ist mit einem langsamen Fortschreiten der Simulation verbunden, da nun, um denselben Zeitabschnitt zu simulieren, mehr Iterationen durchgeführt werden müssen. Ein zu großer Zeitschritt kann zu Instabilitäten innerhalb der Simulation führen, da z. B. zwei sich aufeinander zu bewegendes Teilchen nicht rechtzeitig ein Update der resultierenden Abstoßung erhalten und kollidieren. In der Praxis wird typischerweise ein Zeitschritt von 2 fs genutzt.

### 3.2.3. Wassermodele

In einer atomistischen MD-Simulation von Proteinen in Lösung sind explizite Wassermoleküle die am häufigsten vorkommenden Teilchen. Es wird daher eine gute Beschreibung benötigt, die gleichzeitig sehr effizient berechenbar ist. Dabei werden die Wassermodele anhand der Anzahl der Punkte im Raum unterschieden, die benutzt werden, um das Molekül zu beschreiben. Intramolekulare Bewegungen eines Wassermoleküls werden hierbei meist nicht

abgebildet, da diese zu hochfrequent sind und die Effizienz der Simulation verringern würden (siehe Kapitel 3.2.4). Eines der bekanntesten Wassermodelle ist TIP3P.<sup>110,111</sup> Hierbei wird das Wassermolekül über drei Punkte dargestellt, die das Wasserstoff- und Sauerstoffatom mit ihren entsprechenden Massen, Partialladungen und Lennard-Jones-Parameter darstellen. Auch wenn TIP3P aufgrund seiner Einfachheit immer noch breite Anwendung findet, wird es in vielen Punkten von anderen Wassermodellen inzwischen übertroffen.<sup>112</sup> Daneben gibt es noch das OPC3-Modell, bei dessen Parametrisierung darauf geachtet wurde, die Multipol-Momente eines Wassermoleküls möglichst gut zu reproduzieren.<sup>113</sup> Vierpunkt-Modelle verschieben den Ladungsschwerpunkt des Sauerstoffatoms weg von seinem Massenschwerpunkt, um die Elektronendichte besser zu repräsentieren. Fünfpunkt-Modelle gehen hierbei noch einen Schritt weiter und benutzen zwei Ladungsschwerpunkte, um die beiden freien Elektronenpaare des Wassermoleküls nachzuahmen. Es sollte hierbei aber betont werden, dass durch die erhöhte Anzahl der Punkte der Rechenaufwand einer Simulation dramatisch steigen kann, da typischerweise mehrere zehntausend Wassermoleküle in einer Simulation enthalten sein können.

#### 3.2.4. SHAKE-Algorithmus

Die Effizienz einer MD-Simulation wird bestimmt von der schnellsten Bewegung, die simuliert werden muss. Dies sind die Streckschwingungen in Wasserstoffbindungen, aufgrund der geringen Masse des Wasserstoffatoms. Da diese Bewegungen keine relevanten Informationen für Konformationsänderungen des Proteins oder eines Liganden beinhalten, wurden Algorithmen entwickelt, um diese Bindungen zu fixieren. Die beiden bekanntesten sind hierbei SHAKE<sup>114</sup> und RATTLE.<sup>115</sup> Diese tragen maßgeblich dazu bei, die Schnelligkeit der Simulation zu steigern, da so ein größerer Zeitschritt verwendet werden kann.

#### 3.2.5. Beschreibung der Partialladungen

Für die Beschreibung der elektrostatischen Wechselwirkungen werden die Partialladungen jedes einzelnen Atoms benötigt. Dies ist problematisch, da diese keine Eigenwerte der Wellenfunktion darstellen, daher keine physikalischen Observablen sind und nicht aus Experimenten bestimmt werden können. Im Gegensatz dazu stellt das elektrostatische Potenzial eines Moleküls eine physikalische Observable dar und kann direkt über QM-Rechnungen bestimmt werden. Basierend auf unterschiedlichen Fitting-Methoden können dann die Partialladungen jedes Atoms so gewählt werden, dass das elektrostatische Potenzial des Moleküls möglichst

gut abgebildet wird. Im Folgenden sind zwei Grenzfälle dargestellt, die so unzureichend beschrieben werden:

- Ist ein Molekül oder Atom leicht polarisierbar, ändert sich das elektrostatische Potenzial abhängig von der Umgebung, was sich dann entsprechend in den Partialladungen des Atoms widerspiegeln sollte. Die in dieser Arbeit verwendeten Kraftfelder sind allerdings nicht darauf ausgelegt, derartige Änderungen abzubilden.
- Bei der Bestimmung von Partialladungen wird bei den hier verwendeten Kraftfeldern implizit angenommen, dass ein Punkt im Raum ausreichend ist, um das elektrostatische Potenzial um ein Atom herum zu beschreiben. Bei stark polarisierten Halogenatomen ist zum Beispiel allerdings bekannt, dass die Ladungsverteilung nicht homogen um das Atom liegt, sondern dass gegenüberliegend zur  $\sigma$ -Bindung eine Lücke mit geringerer Elektronendichte vorliegt, das sogenannte  $\sigma$ -Loch. Derartige Effekte werden in klassischen MD-Simulationen nicht beschrieben.

Zur Parametrisierung von den in dieser Arbeit verwendeten AMBER-Kraftfelder wurde ursprünglich das elektrostatische Potenzial mit der HF/6-31G\*-Methode bestimmt,<sup>116,117</sup> da durch die Verwendung der Hartree-Fock-Methode eine Fehlerkompensation der Solvatationsenthalpie erreicht wurde. Folglich ist es ratsam, für Simulationen mit diesen Kraftfeldern neue Moleküle auf dieselbe Art und Weise zu parametrisieren. Für die Berechnung der Partialladungen wird das *restrained electrostatic potential*-Fitting (RESP-Fitting)<sup>116</sup> verwendet.

### 3.2.6. Simulationsboxen und periodische Randbedingungen (PBC)

Die Ränder eines Systems sind kritisch für die Simulation, da hier die „Unendlichkeit“ des Solvens dargestellt werden soll, ohne tatsächlich die Position von unendlich vielen Wassermolekülen berechnen zu müssen. Um dies zu ermöglichen, wird sobald ein Wassermolekül den Rand des Systems überschreitet, es auf der gegenüberliegenden Seite wieder eingefügt. Mathematisch betrachtet bedeutet dieses Spiegeln des verlassenden Wassermoleküls, dass an die betrachtete Simulationsbox identische, sich periodisch wiederholende Abbilder grenzen. Es wird daher auch von periodischen Randbedingungen (*periodic boundary conditions*, PBC) gesprochen.

Dadurch ist es möglich, dass ein Protein mit einem seiner Abbilder, zumindest indirekt, interagiert, wenn die Ränder der Simulation sehr nahe am Protein liegen. Derartige Interaktionen wären ein Artefakt der Simulationsbedingungen, da dies eine Interaktion mit sich selbst darstellt und für Proteine in Lösung ungewollt ist. Hier muss folglich eine Balance

zwischen einer möglichst kleinen Simulationsbox, um den Rechenaufwand gering zu halten, und einer ausreichend großen Box getroffen werden. Die Wahl einer passend geformten Box ist hierbei hilfreich. Rein theoretisch ist jede symmetrische geometrische Form möglich, die eine lückenlose Füllung des Raumes erlaubt. Die Box sollte jedoch an die Form des Proteins angepasst werden, um möglichst wenig Wassermoleküle zu benötigen. Neben würfelförmigen Boxen werden daher auch z. B. stumpfe Oktaeder verwendet.

### 3.2.7. Ewald-Summation

Mathematisch betrachtet erstrecken sich die nichtkovalenten Terme der Wechselwirkung zweier Atome bis ins Unendliche. Im Rahmen der Simulation würde dies eine große Anzahl an Berechnungen bedeuten, die für jedes Atom notwendig sind. Während dieser Umstand bei dem mit  $r^{-6}$  schnell abfallenden Lennard-Jones-Potenzial durch die Nutzung eines einfachen Grenzwerts gelöst werden kann, ist das Coulomb-Potenzial hier problematischer, da dieses lediglich mit  $r^{-1}$  abfällt und somit deutlich weitreichender ist. Eine direkte Summation der Wechselwirkungen konvergiert daher nur sehr langsam. Eine häufig genutzte mathematische Lösung für periodische Systeme ist die sogenannte Ewald-Summation.<sup>118</sup> Hierbei wird das zu berechnende Potenzial modifiziert und in zwei Reihen zerlegt. Der erste Term summiert hierbei im realen Raum, der zweite im reziproken. Dadurch, dass der reziproke Term über die *Fast Fourier Transformation* ausgewertet werden kann, kann die Rechenzeit deutlich reduziert werden, sodass die Summation nur noch mit  $N\log N$  statt mit  $N^2$  skaliert.<sup>118</sup>

### 3.2.8. Strukturaufbereitung und Equilibrierung

Startpunkt für eine MD-Simulation bilden für gewöhnlich Kristallstrukturen von Proteinen aus der *Protein Data Bank* (PDB),<sup>100</sup> die über Röntgenstrukturanalyse aufgenommen wurden. Diese Strukturen müssen vor der Simulation aufbereitet werden, z. B. durch Ergänzen von nicht aufgelösten Aminosäuren, dem Hinzufügen der über Röntgenbeugung nicht erfassbaren Protonen und dem Entfernen von Puffermolekülen, die für die Kristallisation notwendig waren. Anschließend erfolgt eine Energieminimierung des Systems mit Hilfe des verwendeten Kraftfeldes. So werden die experimentell aufgenommenen Bindungslängen und -winkeln an die idealen Werte des Kraftfeldes angepasst und evtl. vorhandene ungünstige sterische Wechselwirkungen reduziert. Es sollte ebenfalls beachtet werden, dass die verwendeten Kristallstrukturdaten bei 100 K aufgenommen wurden, die MD-Simulation aber bei höheren Temperaturen, z. B. bei 300 K durchgeführt werden soll. Darüber hinaus gilt, dass die aufgenommenen Kristallstrukturen keinerlei Information zu den Geschwindigkeiten der Atome

besitzen (d. h. vom Standpunkt der Thermodynamik 0 K darstellen). Initiale Geschwindigkeiten werden daher meist über die Maxwell-Boltzmann-Verteilung zufällig zugewiesen, sodass die experimentellen 100 K reproduziert werden.

Vor der eigentlichen Simulation erfolgt dann eine *Equilibrierung* (zum Teil auch *Thermalisierung* genannt), bei der das System von 100 K auf die Simulationstemperatur erwärmt wird. Diese Equilibrierung erfolgt im kanonischen Ensemble, auch *NVT*-Ensemble genannt. Das bedeutet, dass das Volumen  $V$  und die Teilchenanzahl  $N$  konstant gehalten werden, während die Temperatur  $T$  über ein externes Wärmebad konstant gehalten oder schrittweise modifiziert wird. Nach Anpassung der Temperatur muss die Größe der Simulationsbox angepasst werden, um die Dichte von Wasser zu reproduzieren. Diese Equilibrierung erfordert das *NPT*-Ensemble (isothermales-isobares Ensemble). Die Algorithmen, die bei diesen Simulationsschritten die Temperatur und den Druck des Systems kontrollieren, werden hierbei Thermostat und Barostat genannt. Es wurde eine Reihe an unterschiedlichen Algorithmen für diese Aufgaben entwickelt, wie z. B. das Berendsen-Thermostat,<sup>119</sup> das Nosé-Hoover-Thermostat<sup>120</sup> oder der Andersen-Barostat.<sup>121</sup>

Die eigentliche Simulation erfolgt typischerweise ebenfalls im *NPT*-Ensemble. Nach einer festgelegten Anzahl an Simulationsschritten werden die Koordinaten des Systems abgespeichert (z. B. nach 1 oder 10 ps). Diese Momentaufnahmen bilden letztendlich die Trajektorie des Systems, die analysiert werden kann.

### 3.2.9. Analyse von MD-Simulationen

Nach erfolgter Simulation müssen die erhaltenen Trajektorien analysiert und evaluiert werden. Je nach Fragestellung werden hierbei spezifische Verfahren verwendet. Im Folgenden werden zwei der gebräuchlichsten Methoden dargestellt.

#### 3.2.9.1. Root Mean Square Deviation (RMSD)

Um die Ähnlichkeit zwischen zwei Posen eines Moleküls quantitativ auszudrücken, wird häufig die Wurzel der mittleren quadratischen Abweichung (*Root Mean Square Deviation*, kurz RMSD) herangezogen, siehe Gleichung 3.7. So kann zwischen zwei Strukturen  $a$  und  $b$ , die beide dieselbe Anzahl an Atomen  $N$  aufweisen, eine Distanz berechnet werden.

$$RMSD(a, b) = \sqrt{\frac{1}{N} \sum_{i=1}^N [(x_{i,a} - x_{i,b})^2 + (y_{i,a} - y_{i,b})^2 + (z_{i,a} - z_{i,b})^2]} \quad (3.7)$$

Ein mögliches Vorgehen beim Auswerten einer MD-Simulation ist es, eine Struktur als Re-

ferenz festzulegen (z. B. der Startpunkt der Simulation, oder eine Kristallstruktur), alle Momentaufnahmen der Trajektorie bestmöglich zu überlagern, um Rotationen und Translationen des Systems zu entfernen, und anschließend den RMSD-Wert über die Simulationszeit zu betrachten. So können die strukturellen Änderungen im Verlauf der Trajektorie erkannt werden.

Wichtig ist hierbei, dass zwei Strukturen, die denselben RMSD-Wert zu einer Referenz aufweisen, nicht zwangsläufig identisch sein müssen. Für eine tiefgehendere Analyse ist es daher oft hilfreich, mehrere Referenzen zu wählen. Für MD-Simulationen können z. B. zweidimensionale RMSD-Plots erstellt werden, bei denen jeder Zeitpunkt innerhalb der Trajektorie einmal als Referenz genutzt wird, um zu jedem anderen Zeitpunkt in der Trajektorie den RMSD zu berechnen. Auf diese Art und Weise können graphisch Zeitabschnitte bestimmt werden, in denen eine Konformation beibehalten wurde.

### 3.2.9.2. Principal Component Analysis (PCA)

MD-Simulationen von Proteinen können bei bereits wenigen Aminosäuren eine Vielzahl an unterschiedlichen Bewegungen und Konformationen im Verlauf der Trajektorie aufweisen. Dabei zu eruieren, welche Dynamik innerhalb der Trajektorie die relevanteste Bewegung darstellt, ist alles andere als trivial.

Eine Methode zur Reduktion der Komplexität eines Datensatzes ist die Hauptkomponentenanalyse, oder *Principal Component Analysis* (PCA). Ziel der PCA ist die Reduktion der Dimensionalität eines Datensatzes, um möglichst wenig Variablen zur Darstellung der Daten zu benötigen.

Um die Position von  $N$  Atomen innerhalb eines dreidimensionalen Raumes eindeutig zu beschreiben, werden insgesamt  $3N$  Koordinaten benötigt. Eine Trajektorie dieser Atome kann daher als eine Matrix dargestellt werden, die  $3N$  Spalten aufweist und so viele Zeilen wie Zeitpunkte  $a, b, c, \dots$  innerhalb der Trajektorie, siehe Gleichung 3.8. Eine Spalte dieser Matrix  $X_1, Y_1, Z_1, \dots$  beinhaltet alle Werte einer Koordinate, die im Laufe einer Simulation eingenommen werden.

$$\begin{aligned}
 & \begin{pmatrix} x_{a,1} & y_{a,1} & z_{a,1} & x_{a,2} & y_{a,2} & z_{a,2} & \cdots & x_{a,N} & y_{a,N} & z_{a,N} \\ x_{b,1} & y_{b,1} & z_{b,1} & x_{b,2} & y_{b,2} & z_{b,2} & \cdots & x_{b,N} & y_{b,N} & z_{b,N} \\ x_{c,1} & y_{c,1} & z_{c,1} & x_{c,2} & y_{c,2} & z_{c,2} & \cdots & x_{c,N} & y_{c,N} & z_{c,N} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \end{pmatrix} \\
 & = \begin{pmatrix} X_1 & Y_1 & Z_1 & X_2 & Y_2 & Z_2 & \cdots & X_N & Y_N & Z_N \end{pmatrix}
 \end{aligned} \tag{3.8}$$

Viele dieser Koordinaten werden stark miteinander korreliert sein, da die Atome sich nicht komplett frei und zufällig bewegen können, sondern z. B. über Bindungen zusammengehalten werden. Mathematisch bedeutet das, dass die einzelnen Koordinaten nicht statistisch voneinander unabhängig sind, sondern eine Kovarianz  $cov(X, Y)$  miteinander besitzen. Die Kovarianz kann dabei über die Erwartungswerte (bzw. Mittelwerte) der Variablen  $E(X)$  und  $E(Y)$  sowie dem Erwartungswert des Produkts der Variablen  $E(XY)$  definiert werden (siehe Gleichung 3.9).

$$cov(X, Y) = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y) \quad (3.9)$$

Der Startpunkt für eine PCA bildet die Kovarianz-Matrix  $C$  (siehe Gleichung 3.10), die alle Kovarianzen zwischen zwei Variablen für das  $N$ -atomige System beinhaltet.

$$C = \begin{pmatrix} cov(X_1, X_1) & cov(X_1, Y_1) & cov(X_1, Z_1) & \cdots & cov(X_1, X_N) & cov(X_1, Y_N) & cov(X_1, Z_N) \\ cov(Y_1, X_1) & cov(Y_1, Y_1) & cov(Y_1, Z_1) & \cdots & cov(Y_1, X_N) & cov(Y_1, Y_N) & cov(Y_1, Z_N) \\ cov(Z_1, X_1) & cov(Z_1, Y_1) & cov(Z_1, Z_1) & \cdots & cov(Z_1, X_N) & cov(Z_1, Y_N) & cov(Z_1, Z_N) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ cov(X_N, X_1) & cov(X_N, Y_1) & cov(X_N, Z_1) & \cdots & cov(X_N, X_N) & cov(X_N, Y_N) & cov(X_N, Z_N) \\ cov(Y_N, X_1) & cov(Y_N, Y_1) & cov(Y_N, Z_1) & \cdots & cov(Y_N, X_N) & cov(Y_N, Y_N) & cov(Y_N, Z_N) \\ cov(Z_N, X_1) & cov(Z_N, Y_1) & cov(Z_N, Z_1) & \cdots & cov(Z_N, X_N) & cov(X_N, Y_N) & cov(Z_N, Z_N) \end{pmatrix} \quad (3.10)$$

Im Rahmen der PCA werden nun die Eigenvektoren  $V$  ermittelt, die die Kovarianz-Matrix  $C$  zu einer diagonalen Matrix  $D$  transformieren (siehe Gleichung 3.11). Diese Eigenvektoren stellen eine Art alternatives Koordinatensystem dar, über das jeder Zustand des Systems definiert werden kann, analog zu den ursprünglichen x-, y-, und z-Koordinaten der Atome. Mit jedem Eigenvektor ist ein Eigenwert assoziiert, der sich auf der Diagonalen von  $D$  befindet und die Varianz entlang des entsprechenden Eigenvektors widerspiegelt. Große Eigenwerte bedeuten, dass entlang dieser Koordinate große Schwankungen innerhalb des Datensatzes vorliegen.

$$D = V^{-1}CV \quad (3.11)$$

Um nun eine Reduktion der Dimensionalität des Datensatzes zu erhalten, werden die Ei-



genvektoren anhand ihrer Eigenwerte sortiert und nur diejenigen mit den größten Eigenwerten betrachtet, die sogenannten Hauptkomponenten (*Principal Components*, PCs). Dadurch können die größten Schwankungen innerhalb des Datensatzes, d. h. im Kontext von MD-Simulationen die markantesten Bewegungen, mit relativ wenigen Dimensionen abgebildet werden, wobei nur ein Minimum an Information verloren geht. Dies hilft dabei, Trajektorien zu interpretieren und verschiedene Zustände des Systems zu identifizieren. Beispiele für PCA-Analysen können in Kapitel 4.2.6 gefunden werden, siehe Abbildung 4.77.

### 3.3. Quantenmechanische Methoden

Im Weiteren folgt eine kurze Einführung in die Grundlagen der Quantenmechanik. Aufgrund der Größe und Komplexität des Feldes beschränkt sich diese Einführung auf die für diese Arbeit wesentlichen Aspekte; nämlich, auf welchen Grundlagen die angewendeten Methoden basieren und was ein Anwender bei der Verwendung dieser Methoden beachten sollte. Für weiterführende Informationen, insbesondere für die praktische Umsetzung der Rechnungen, wird auf das Buch *Exploring Chemistry with Electronic Structure Methods* von Foresman *et al.* verwiesen.<sup>25</sup> Alle quantenmechanischen Berechnungen wurden mit Gaussian09 durchgeführt.<sup>25, 122</sup>

#### 3.3.1. Hartree-Fock-Theorie und der Self-Consistent-Field-Ansatz

Um eine Wellenfunktion aufzustellen, die ein mehratomiges Teilchen beschreibt, werden häufig Molekülorbitale  $\phi_i$  (MOs) genutzt. Diese werden durch eine Linearkombination der Atomorbitale  $\chi_\mu$  mit den Koeffizienten  $c_{i\mu}$  konstruiert. Die Gesamtheit der Atomorbitale wird als Basis bezeichnet und ähnelt in ihrer Form den Eigenfunktionen des Wasserstoffatom-Hamiltonians (siehe Gleichung 3.12).

$$\phi_i = \sum_{\mu} c_{i\mu} \chi_{\mu} \quad (3.12)$$

Um das Pauli-Prinzip zu erfüllen, muss die Wellenfunktion antisymmetrisch sein. Dies wird erreicht, indem die Molekülorbitale mit der Spinfunktion  $\Omega$  (bzw. dessen Eigenfunktionen  $\alpha$  und  $\beta$ ) zu Spinorbitalen  $\Phi$  zusammengefasst (siehe Gleichung 3.13) werden:

$$\Phi(x_i) = \phi(r_i) \Omega(\omega_i) \quad (3.13)$$

und in Form einer Slater-Determinante kombiniert werden (Gleichung 3.14).<sup>123</sup>

$$\Psi(x) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \Phi_1(x_1) & \Phi_2(x_1) & \Phi_3(x_1) & \cdots & \Phi_n(x_1) \\ \Phi_1(x_2) & \Phi_2(x_2) & \Phi_3(x_2) & \cdots & \Phi_n(x_2) \\ \Phi_1(x_3) & \Phi_2(x_3) & \Phi_3(x_3) & \cdots & \Phi_n(x_3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Phi_1(x_n) & \Phi_2(x_n) & \Phi_3(x_n) & \cdots & \Phi_n(x_n) \end{vmatrix} \quad (3.14)$$

Die Elektronen 1, 2, ...  $n$  verfügen dabei über die Raumkoordinaten  $r_1, r_2, \dots, r_n$  und über die Spinkoordinaten  $\omega_1, \omega_2, \dots, \omega_n$  bzw. die Spinraumkoordinaten  $x_1, x_2, \dots, x_n$ .

Bis hierhin wurde nur beschrieben, wie eine formal korrekte Wellenfunktion eines Systems aufgebaut werden kann. Es ist allerdings noch nicht klar, wie diese Wellenfunktion konkret aussieht, da die Koeffizienten  $c_{i\mu}$  zur Konstruktion der MOs noch unbekannt sind. Einen Ansatz hierfür bietet das Variationsprinzip (siehe Gleichung 3.15). Dieses besagt, dass die minimale Energie eines Systems nur mit der korrekten Wellenfunktion erhalten wird; mit anderen Wellenfunktionen werden immer höhere Werte erhalten.<sup>25,123</sup>

$$\frac{\int \Psi^* \hat{H} \Psi d\tau}{\int \Psi^* \Psi d\tau} = E_{\text{variational}} \geq E_{\text{exact}} \quad (3.15)$$

Somit ist der Pfad zur richtigen Wellenfunktion klar gezeichnet: Eine initiale, geratene Wellenfunktion gibt einen Startpunkt vor, für den  $E_{\text{variational}}$  berechnet werden kann und der anschließend variiert wird. Solange dabei Wellenfunktionen gefunden werden, die einen niedrigeren Wert für die Energie ergeben, ist noch nicht die korrekte Wellenfunktion gefunden und es muss weiter variiert werden.

Fock und Hartree sind diesen Pfad vorausgegangen und entwickelten (jeweils) ein System aus Eigenwertgleichungen, bei dem die Eigenfunktionen die bestmöglichen Spinorbitale  $\Phi_i$  sind. Die Eigenwerte dieser Gleichungen werden Orbitalenergien  $\epsilon_i$  genannt (Gleichung 3.16).<sup>25,123</sup>

$$\hat{F}\Phi_i = \hat{T}\Phi_i + \hat{V}_{NE} + \hat{J}\Phi_i - \hat{K}\Phi_i = \epsilon_i\Phi_i \quad (3.16)$$

Hierbei wird der Fock-Operator  $\hat{F}$  aus dem Operator der kinetischen Energie  $\hat{T}$  (der über die zweifache Ableitung nach dem Ort,  $\nabla^2$ , definiert wird),

$$\hat{T} = -\frac{1}{2}\nabla^2$$

dem Operator für die Anziehung zwischen Elektronen und Kernen mit der Ladung  $Z$ ,

$$\hat{V}_{NE} = \sum_A^{\text{Kerne}} \frac{-Z_A}{r_{iA}}$$

dem Coulomb-Operator

$$\hat{J}\Phi_i(x_1) = \left( \sum_j^{\text{occupied}} \int \Phi_j^*(x_2) \frac{1}{r_{12}} \Phi_j(x_2) dx_2 \right) \Phi_i(x_1)$$

sowie dem Austausch-Operator

$$\hat{K}\Phi_i(x_1) = \left( \sum_j^{\text{occupied}} \int \Phi_j^*(x_2) \frac{1}{r_{12}} \Phi_i(x_2) dx_2 \right) \Phi_j(x_1)$$

zusammengesetzt, jeweils in atomaren Einheiten. Beachtenswert ist hierbei, dass sowohl  $\hat{J}$  als auch  $\hat{K}$  die Spinorbitale benötigen, deren Form erst bestimmt werden soll. Die Fock-Gleichung bietet daher die Grundlage für ein iteratives Optimierungsverfahren, das als *Self-Consistent-Field-Ansatz* (SCF-Ansatz) bezeichnet wird. Nach dem Lösen der Fock-Gleichung wird ein neues Set an Spinorbitalen erhalten, die erneut in die Fock-Gleichung eingesetzt werden, bis sich konvergierte Lösungen ergeben. Aus diesen optimierten Orbitalen resultieren anschließend nach dem Variationsprinzip die optimierten Energien. Es sollte betont werden, dass  $\hat{J}$  und  $\hat{K}$  effektive Ein-Elektronen-Operatoren sind, da über die Spin-Raum-Koordinaten des zweiten Elektrons  $x_2$  integriert wird. Dies bedeutet, dass ein Elektron, das ein Spinorbital  $\Phi_i$  besetzt, lediglich das gemittelte Feld aller anderen Elektronen erfährt, was die größte Näherung in der HF-Theorie darstellt.<sup>25,123</sup>

### 3.3.2. Dichtefunktionaltheorie (DFT)

Die Abweichung, die durch die Annahme eines gemittelten Feldes in der HF-Theorie entsteht, ist groß genug, dass der sinnvolle Anwendungsbereich der HF-Methode sehr begrenzt ist. Methoden, die versuchen, die in der HF-Theorie nicht berücksichtigten Wechselwirkungen zu erfassen (sogenannte Post-HF- oder Post-SCF-Methoden), sind allerdings sehr rechenintensiv, vor allem für Systeme mit einer großen Anzahl an Elektronen. Dies ist einer der Gründe, warum die Dichtefunktionaltheorie (DFT) breite Anwendung findet. Grundlage für die DFT sind die Hohenberg-Kohn-Theoreme.<sup>124,125</sup> Diese besagen, dass ein System im Grundzustand über dessen Elektronendichteverteilung  $\rho$  anstelle der Wellenfunktion vollständig beschrieben

werden kann. Diese kann wieder über eine Basis konstruiert werden:

$$\rho(\vec{r}_g) = \sum_i^{\text{occupied}} \phi_i^2(\vec{r}_g) \quad (3.17)$$

Es gilt analog wie zuvor das Variationsprinzip (vergleichbar zu Gleichung 3.15). Hierbei sollte betont werden, dass der Index  $g$  des Positionsvektors anzeigen soll, dass der Raum auf einem Gitter  $g$  betrachtet wird. Somit ist  $\rho(\vec{r}_g)$  die Wahrscheinlichkeit, ein Elektron im Gitterelement zu finden, das durch  $\vec{r}_g$  definiert wird. Schematisch wird die Energie innerhalb der DFT durch Gleichung 3.18 berechnet:

$$E_{DFT}[\rho(\vec{r})] = T[\rho(\vec{r})] + V_{NE} + J[\rho(\vec{r})] + c_{HF}K + c_{DFT}E_{XC}[\rho(\vec{r})] \quad (3.18)$$

Auffällig ist hierbei die Ähnlichkeit zum Ansatz der HF-Theorie, da die ersten vier Terme analog zu Gleichung 3.16 lauten, mit der Ausnahme, dass sie in der DFT Funktionale der Elektronendichte sind und nicht Operatoren, die auf ein Orbital wirken. Diese Ähnlichkeit erklärt den vergleichbaren Rechenaufwand der beiden Ansätze. In der ursprünglichen DFT wird der Hartree-Fock-Austausch-Term  $K$  nicht berücksichtigt ( $c_{HF} = 0$ ,  $c_{DFT} = 1$ ); bei sogenannten Hybrid-Funktionalen wird der Term teilweise berücksichtigt, da dieser bereits in der HF-Theorie exakt bestimmt wird, in DFT allerdings nicht. Neu ist der Austausch-Korrelation-Term  $E_{XC}$ , der jene Elektron-Elektron-Wechselwirkungen und Fehlerkorrekturen der anderen vier Terme beschreiben soll, die in der HF-Theorie nicht erfasst werden. Die exakte Form, wie  $E_{XC}$  aussieht, variiert zwischen den einzelnen Funktionalen, da verschiedene Terme für die lokale und die nichtlokale Austausch- und Korrelationsenergie verwendet werden. Klassische Funktionale leiten diesen Term physikalisch her, während neuere DFT- und vor allem Hybridfunktionale, empirisch parametrisiert werden. Die Parameter, mit denen die einzelnen Terme gewichtet werden, stammen dann z. B. aus verschiedenen experimentellen Datensätzen. Derartiges Fitting macht es schwer nachvollziehbar, welche physikalische Grundlage die Unterschiede zwischen den einzelnen Funktionalen haben. Zusätzlich sollte bei jedem Funktional hinterfragt werden, ob der für das Fitting verwendete Datensatz für die eigene Problematik repräsentativ ist. Das wohl bekannteste Beispiel für ein (Hybrid)-Funktional ist B3LYP.<sup>126–128</sup>

Wesentlicher Vorteil von DFT-Ansätzen ist ihre Schnelligkeit im Vergleich zu post-HF-*ab initio*-Methoden bei gleichzeitig deutlich verbesserter Genauigkeit im Vergleich zur HF-Theorie. Allerdings weist auch diese Methode einige wesentliche Schwachpunkte auf, die gerade für die Thematik von nichtkovalenten Komplexen zweier miteinander reagierender

Spezies von Relevanz sind. Nicht alle DFT-Funktionale beschreiben das Abklingen des Austauschpotentials bei großen Abständen korrekt (gemeint sind solche Abstände, die für nicht-kovalente Komplexe relevant wären), da  $K$  nicht oder nicht vollständig berücksichtigt wird.<sup>25</sup> Dispersionswechselwirkungen werden ebenfalls nicht abgebildet.<sup>129</sup> Für beide Problematiken wurden entsprechende Korrekturen entworfen, z. B. B3LYP mit Coulomb-Dämpfung, CAM-B3LYP genannt,<sup>130</sup> oder das Funktional  $\omega$ B97xD.<sup>131</sup>

### 3.3.3. Geometrieoptimierung und Frequenzberechnung

Das Variationsprinzip bietet die Grundlage, die Orbitalfunktionen eines Moleküls zu optimieren und die mit der Struktur assoziierte Energie zu bestimmen. Die auf diese Art und Weise bestimmte Energie des Moleküls ist allerdings abhängig von der zu Beginn festgesetzten Geometrie des Moleküls. Bei einer Geometrieoptimierung ist das Ziel, die Koordinaten der Atome innerhalb des Moleküls so zu variieren, dass die im SCF-Ansatz bestimmte Energie minimiert wird.<sup>25</sup> Hierbei wird der Gradient und die zweite Ableitung nach der Energie, die Hesse-Matrix, der Wellenfunktion genutzt, um den nächsten Schritt der Geometrieänderung zu bestimmen. Der Gradient beinhaltet Informationen über die Kräfte, die auf jedes Atom wirken, während die Hesse-Matrix die Informationen zu den unterschiedlichen Kraftkonstanten innerhalb des Moleküls beinhaltet.<sup>25</sup> Hierbei ist erwähnenswert, dass die explizite Berechnung der Hesse-Matrix aufgrund der vielen Freiheitsgrade eines mehratomigen Teilchens für jeden Punkt einer Geometrieoptimierung äußerst rechenintensiv ist. Typischerweise werden daher zu Beginn Kraftkonstanten mit Hilfe eines Kraftfeldes abgeschätzt und diese im Laufe der Optimierung aktualisiert.<sup>25</sup> Anspruchsvolle Optimierungen, die z. B. zum Ziel haben, einen Übergangszustand zu bestimmen, können allerdings davon profitieren, zumindest am Start explizit die Hesse-Matrix berechnen zu lassen.<sup>25</sup> Eine Geometrieoptimierung gilt dann als konvergiert und vollendet, wenn die verbleibenden Kräfte innerhalb des Moleküls und die Änderungen der Koordinaten im Vergleich zum vorherigen Schritt unterhalb eines vorgegebenen Grenzwertes liegen.<sup>122</sup> Aus mathematischer Sicht treffen diese Kriterien aber potenziell auf jeden stationären Punkt einer Potenzialhyperfläche zu. Daher ist es nötig, die mathematische Natur des gefundenen Punktes zu überprüfen (d. h. zu testen, ob ein Minimum oder Maximum vorliegt). Dies erfolgt über die Berechnung der Vibrationsfrequenzen und der dazugehörigen Kraftkonstanten, also über die Hesse-Matrix.<sup>25</sup> Werden hierbei keine negativen Eigenwerte gefunden, liegt ein Minimum vor, falls (genau) eine negative Frequenz angezeigt wird, handelt es sich bei der gefundenen Geometrie um einen Übergangszustand, der zwei Minima miteinander verbindet.

### 3.3.4. Intrinsic-Reaction-Coordinate-Berechnungen

Wird ein Übergangszustand in Isolation betrachtet, ist nicht zwangsläufig offensichtlich, ob er für die Reaktion, die betrachtet werden soll, relevant ist. *Intrinsic reaction coordinate*-Berechnungen,<sup>132</sup> kurz IRC-Berechnungen, erlauben es, ausgehend von einem Übergangszustand eines Moleküls die beiden Minima zu identifizieren, die durch den Übergangszustand miteinander verbunden werden. Dafür wird die Reaktionskoordinate schrittweise in beide Richtungen der Reaktion verfolgt, bis das Potential abflacht. Da der Algorithmus, der hierfür genutzt wird, nicht dafür ausgelegt ist, ein Minimum zu identifizieren, ist es notwendig, die Geometrie der beiden so erhaltenen Endpunkte erneut zu optimieren.<sup>25</sup> Mithilfe der so bestimmten Endpunkte der Reaktion und des Übergangszustandes können nun Aktivierungsenergien für die Hin- und Rückreaktion sowie die frei werdende Energie bestimmt werden.

---

## 4. Ergebnisse

### 4.1. CatK: Simulation von nichtkovalenten Modellen basierend auf kovalenten Protein-Ligand-Komplexen

Für die Beschreibung des Zustandes vor der Reaktion zwischen einem kovalenten Inhibitor und einem Protein existieren fast keine experimentellen strukturellen Daten. Kristallstrukturen von kovalenten Inhibitoren, wie sie z. B. in der *Protein Data Bank* (PDB)<sup>100</sup> hinterlegt sind, zeigen häufig nur den kovalent gebundenen Zustand nach der Reaktion.

Diese Studie hatte zum Ziel, ein Modell für den nichtkovalenten Zustand vor der Reaktion aus jenen Kristallstrukturen zu generieren, und dessen dynamisches Verhalten im Rahmen einer MD-Simulation zu untersuchen. Insbesondere war hierbei der Abstand zwischen den beiden Reaktionspartnern von Interesse, d. h. dem elektrophilen Kohlenstoffatom des Warheads und dem nukleophilen Schwefelatom des katalytischen Cysteins. Der Einfachheit halber wird dieser Abstand im Folgenden oft nur als „reaktive Distanz“ bezeichnet. Für papainähnliche Cysteinproteasen, und insbesondere CatK, existieren viele publizierte Kristallstrukturen mit einem kovalenten Inhibitor, was eine gute Datengrundlage für ein derartiges Unterfangen darstellt.

Einige der hier untersuchten Simulationen erfolgten bereits in der vorangegangenen Masterarbeit.<sup>133</sup> Im Rahmen der jetzigen Arbeit wurden die zugrunde liegenden Kristallstrukturen aber nochmal kritisch bewertet und eine Selektion basierend auf einigen Qualitätskriterien durchgeführt. Falls im Rahmen dieser Inspektion Fehler in der Präparation der Simulation auffällig wurden, wurden entsprechend verbesserte Simulationen durchgeführt. Das genutzte MD-Protokoll ist in Kapitel 4.1.2 festgehalten. Um den Einfluss des Protonierungszustandes besser nachvollziehen zu können, wurde besonderes Augenmerk auf die Solvatation der katalytischen Einheit gelegt (siehe Kapitel 4.1.4). Für einige Komplexe wurde exemplarisch eine verlängerte Simulation durchgeführt, um eventuelle Abweichungen zu kürzeren Sampling-Zeiten zu erkennen (siehe Kapitel 4.1.5). Anders als in der vorangegangenen Masterarbeit war hier nicht nur CatK Teil der Studie, sondern auch andere Cysteinproteasen, wie z. B. das verwandte CatS (siehe Kapitel 4.1.6.1) oder die SARS-CoV-2-Main Protease (siehe Kapitel 4.1.6.2). Darüber hinaus wurde auch getestet, ob derartige nichtkovalente Modelle sich

ebenfalls über Docking generieren lassen (siehe Kapitel 4.1.7.1), und wie die Einbindung der hier gefundenen Ergebnisse in einem virtuellen Screening aussehen könnte (siehe Kapitel 4.1.7.2).

Die hier gezeigten Analysen wurden bereits zum Teil an anderer Stelle publiziert.<sup>134</sup>

#### 4.1.1. Auswahl der Kristallstrukturen

Die PDB wurde nach kovalenten Ligand-CatK-Komplexen durchsucht. Dabei wurden folgende Kriterien an die Strukturen gestellt:

- Die Auflösung der Struktur sollte kleiner als 2.5 Å sein.
- Die Bindungslänge zwischen Cystein-Schwefelatom und vormals elektrophilen Kohlenstoffatom sollte  $1.8 \pm 0.1$  Å betragen.
- Es sollte nicht gegen mehr als eine der Lipinski-Regeln verstoßen werden.<sup>135</sup>
- Es wurden nur Kristallstrukturen einbezogen, die bis zum 31.03.2019 veröffentlicht wurden.

Die ersten beiden Kriterien sollten sicherstellen, dass nur Kristallstrukturen mit ausreichender Qualität für Simulationen genutzt werden, während durch das Einhalten der Lipinski-Regeln gewährleistet werden soll, dass Verbindungen aussortiert werden, die wenig Ähnlichkeit zu typischen Wirkstoffen haben. Eine Auflistung aller PDB-Dateien, die diese Kriterien erfüllen, ist in Tabelle 4.1 zu sehen, die dazugehörigen Ligandstrukturen in Abbildung 4.1. In Tabelle 4.1 ist ebenfalls zu lesen, ob die dazugehörigen Simulationsergebnisse aus der vorhergehenden Masterarbeit übernommen oder verbessert wurden, oder ob es sich um neue Simulationen handelt.

Ein Großteil der hier diskutierten Simulationen erfolgten bereits im Rahmen der vorhergehenden Masterarbeit.<sup>133</sup> Für 1ATK, 4X6H und 4X6I wurden allerdings Mängel im Setup identifiziert, weswegen im Rahmen dieser Arbeit korrigierte Simulationen erfolgten. Simulationen zu 1U9X und 6QBS waren im Rahmen der Masterarbeit nicht erfolgt und wurden nun durchgeführt. Die entsprechenden Ergebnisse werden in Kapitel 4.1.3 diskutiert.



4.1. CatK: Simulation von nichtkovalenten Modellen basierend auf kovalenten Protein-Ligand-Komplexen

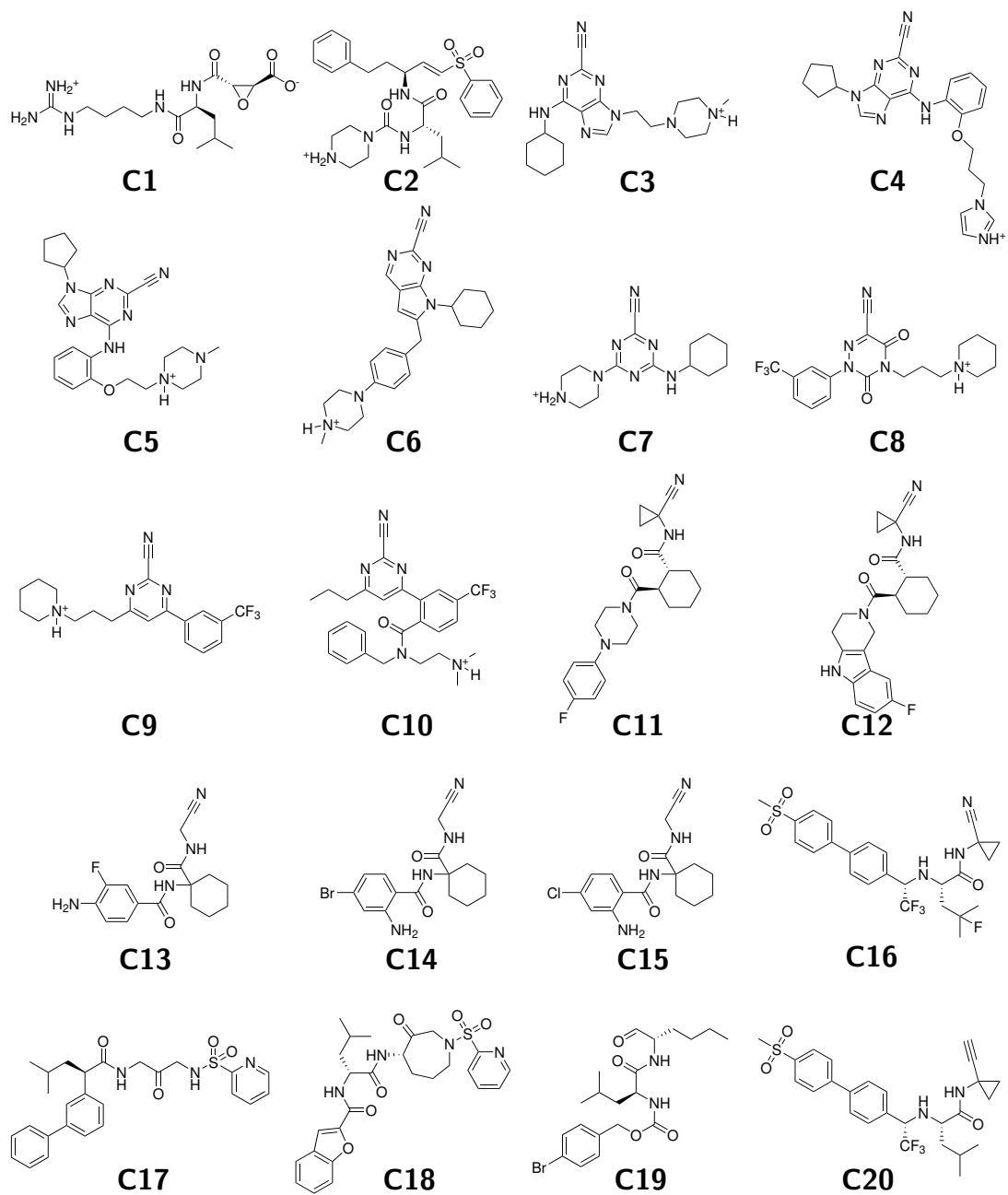


Abbildung 4.1.: Alle Liganden der CatK-Studie. Die Liganden werden bereits in der Prototypierungsstufe gezeigt, in der sie letztlich auch simuliert wurden.

PDB	Ligand	Warhead	Publi- kation	Auflö- sung [Å]	C-S-Bin- dungslänge [Å]	Lipinski- verstöße	Alt/ Neu?
1ATK <sup>136</sup>	<b>C1</b>	Epoxid	1998	2.20	1.8	1	Korrigiert <sup>133</sup>
1MEM <sup>36</sup>	<b>C2</b>	Vinylsulfon	1998	1.80	1.8	1	Alt <sup>133</sup>
1U9V <sup>137</sup>	<b>C3</b>	Nitril	2005	2.20	1.8	0	Alt <sup>133</sup>
1U9W <sup>137</sup>	<b>C4</b>	Nitril	2005	2.30	1.7	0	Alt <sup>133</sup>
1U9X <sup>137</sup>	<b>C5</b>	Nitril	2005	2.10	1.7	0	Neu <sup>133</sup>
2R6N <sup>138</sup>	<b>C6</b>	Nitril	2007	1.95	1.8	0	Alt <sup>133</sup>
3KW9 <sup>139</sup>	<b>C7</b>	Nitril	2010	1.80	1.8	0	Alt <sup>133</sup>
3KWB <sup>140</sup>	<b>C8</b>	Nitril	2010	2.02	1.8	0	Alt <sup>133</sup>
3KWZ <sup>139</sup>	<b>C9</b>	Nitril	2010	1.49	1.8	0	Alt <sup>133</sup>
3O1G <sup>141</sup>	<b>C10</b>	Nitril	2010	1.65	1.8	0	Alt <sup>133</sup>
4DMX <sup>35</sup>	<b>C11</b>	Nitril	2012	1.70	1.8	0	Alt <sup>133</sup>
4DMY <sup>35</sup>	<b>C12</b>	Nitril	2012	1.63	1.8	0	Alt <sup>133</sup>
4X6H <sup>142</sup>	<b>C13</b>	Nitril	2015	1.00	1.8	0	Korrigiert <sup>133</sup>
4X6I <sup>142</sup>	<b>C14</b>	Nitril	2015	1.87	-*	0	Korrigiert <sup>133</sup>
4X6J <sup>142</sup>	<b>C15</b>	Nitril	2015	1.59	1.7	0	Alt <sup>133</sup>
5TDI <sup>143</sup>	<b>C16</b>	Nitril	2017	1.40	1.8	1	Alt <sup>133</sup>
1BGO <sup>144</sup>	<b>C17</b>	Vinylsulfon	1999	2.30	1.9	0	Alt <sup>133</sup>
1NLJ <sup>145</sup>	<b>C18</b>	Vinylsulfon	2003	2.40	1.8	1	Alt <sup>133</sup>
1SNK <sup>146</sup>	<b>C19</b>	Aldehyd	2004	2.40	1.8	1	Alt <sup>133</sup>
6QBS <sup>147</sup>	<b>C20</b>	Alkin	2019	1.70	1.8	1	Neu <sup>133</sup>

\*4X6I wurde aus Vergleichsgründen aufgenommen, siehe 4.1.3.3

Tabelle 4.1.: Die für diese Arbeit in Betracht gezogenen PDB-Dateien. Die für die Auswahl betrachteten Parameter (Publikationsjahr, Auflösung, C-S-Bindungslänge und Lipinskiverstöße des Liganden) sind ebenfalls gegeben. In der letzten Spalte wird zusätzlich angegeben, ob die dazugehörige Simulation aus der Masterarbeit übernommen (alt), nochmal korrigiert hier (korrigiert) oder vor dieser Arbeit noch nicht erfolgt ist (neu). Die entsprechenden Ligand-Strukturen können in Abbildung 4.1 gesehen werden.

#### 4.1.2. MD-Protokoll

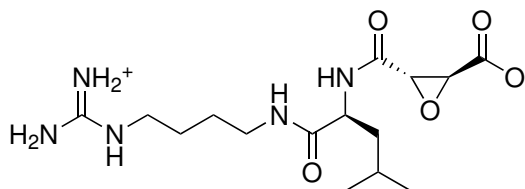
Die PDB-Strukturen wurden im *Molecular Operating Environment*, kurz MOE,<sup>148</sup> in nicht-kovalente Strukturen überführt. Protein, Ligand und Wassermoleküle der Struktur wurden übernommen, während andere Moleküle und Ionen entfernt wurden. Anschließend wurden Wasserstoffatome durch die Funktion *protonate3D* von MOE 2018 bei einem pH von 7.0 ergänzt,<sup>149</sup> was zu geladenen Aspartat-, Glutamat-, Lysin- und Argininseitenketten führte. Die für die Katalyse relevanten Aminosäuren Cys25 und His162 wurden hierbei in zwei Protonierungszuständen modelliert: zwitterionisch, d. h. mit negativ geladenem Schwefelatom und protoniertem Imidazolring (im Folgenden dem AMBER-Namensschema nach als

CYM/HIP bezeichnet), und als neutrale Aminosäuren (CYS/HIS). Disulfidbrücken wurden zwischen Cys22 und Cys63, Cys56 und Cys96 sowie Cys155 und Cys204 gesetzt. In PDB 1U9W fehlten einige Atome einer flexiblen Kette des Liganden. Diese wurden in Analogie zur Kristallstruktur 1U9X modelliert. Die kovalente Bindung zwischen Inhibitor und Protein wurde aufgelöst und der Zustand vor der Reaktion hergestellt. Anschließend erfolgte eine lokale Energieminimierung durch das Kraftfeld MMFF94x<sup>150</sup> bis zu einem RMS-Gradienten von 0.001 kcal/(mol·Å). Die Partialladungen der Ligandatome wurden anschließend durch RESP-Fitting<sup>116</sup> mit *antechamber* basierend auf dem elektrostatischen Potenzial bestimmt, welches wiederum durch Gaussian09 auf dem HF/6-31G\* Niveau berechnet wurde. Kraftfeldparameter für den Liganden wurden aus dem *general AMBER force field*, GAFF,<sup>117</sup> entnommen, während für die Beschreibung des Proteins das AMBER-Kraftfeld ff14sb<sup>108</sup> genutzt wurde. Durch das AMBER-Tool *tleap*<sup>151</sup> wurden anschließend Protein, nichtkovalenter Ligand und Wassermoleküle vereint zu einem Komplex, der in *sander*<sup>151</sup> durch Nutzung eines verallgemeinerten, impliziten Born-Solvensmodell<sup>152</sup> für 2000 Schritte minimiert wurde. Diese Minimierung dient nur zur Anpassung des Komplexes an das Kraftfeld und wurde daher bewusst kurz gehalten. Im Anschluss wurde der Komplex in eine Box aus TIP3P-Wassermolekülen<sup>110,111</sup> gefügt, mit einem minimalen Abstand von 10 Å zwischen Protein und Box. Das System wurde dann durch Chloridionen neutralisiert. Die so erhaltene Simulationsbox wurde dann schrittweise unter periodischen Randbedingungen im NVT-Ensemble thermal equilibriert. Zuerst wurde das Solvens von 100 K auf 300 K innerhalb von 250 000 Schritten à 2 fs (entsprechend 0.5 ns) aufgeheizt, während Proteinatome durch ein harmonisches Potenzial mit einer Kraftkonstante von 100 kcal/(mol·Å<sup>2</sup>) festgehalten wurden. Anschließend wurde dieses Potenzial auf 10 kcal/(mol·Å<sup>2</sup>) gelockert, wobei das Solvens innerhalb von 0.25 ns zurück auf 100 K gekühlt wurde. Die Idee hinter diesem zweiten Schritt ist, dass das bis jetzt noch „kalte“ Protein, dessen Struktur bei 100 K aufgenommen wurde und das bisher noch durch ein starres Potenzial festgehalten wurde, möglichst schonend die Temperatur des Solvens annimmt, ohne dass weiter aufgeheizt wird. Auf diese Art und Weise sollen mögliche Artefakte durch zu schnelles Aufheizen vermieden werden. Zuletzt wurden die Constraints des Proteins auf 1 kcal/(mol·Å<sup>2</sup>) gesenkt, während das System innerhalb von 0.5 ns nun erneut zurück auf 300 K erhöht wurde. Das System wurde anschließend für 0.5 ns im NPT-Ensemble unter Nutzung des Berendsen-Barostats und einer Druck-Relaxationszeit von 2 ps simuliert, um einen Druck von 1 bar und eine entsprechend angepasste Boxgröße zu erhalten. Von diesen equilibrierten Systemen ausgehend wurden jeweils drei Trajektorien mit 10 ns Länge simuliert. Alle Simulationen erfolgten mit *pmemd*<sup>153</sup> von AMBER18 unter Nutzung einer Langevin-Dynamik mit einer Kollisionsfrequenz von 3 ps<sup>-1</sup>. Kovalente Bin-

dungen von Wasserstoffatomen wurden mithilfe des SHAKE-Algorithmus<sup>114</sup> fixiert, sodass ein Zeitschritt von 2 fs genutzt werden konnte. Nichtkovalente Wechselwirkungen wurden über die Particle Mesh Ewald Methode<sup>118</sup> beschrieben, wobei der Cutoff hierbei auf 10 Å gesetzt wurde. Zur Auswertung der Simulationen wurde für jedes Triplikat der Abstand zwischen dem reaktiven Schwefelatom von Cys25 und dem elektrophilen Kohlenstoffatom des Liganden gemessen. Falls nicht anders angegeben wird dieser Abstand als „*moving average*“ über 25 Schritte dargestellt.

### 4.1.3. Neu erfolgte Simulationen

#### 4.1.3.1. Verbesserung der 1ATK-Simulation



**C1**

Abbildung 4.2.: Ligand von 1ATK.

In der Kristallstruktur 1ATK, in der CatK mit Verbindung **C1** kristallisiert ist, weist eines der Stickstoffatome innerhalb der Guanidinsubstruktur eine Orientierung auf, die die dazugehörigen Protonen orthogonal zur Ebene der funktionellen Gruppe stehen lassen würden. Da diese unphysikalische Ausrichtung des Atoms in der ersten Simulation innerhalb der Masterarbeit für das nichtkovalente Modell des Liganden übernommen wurde, erfolgten nun korrigierte Simulationen, die durch Energieminimierung mit MMFF94x zu einem Gradienten von  $0.001 \text{ kcal}/(\text{mol}\cdot\text{\AA})$  die korrekte Planarität des Guanidins bereits von Anfang an aufwiesen. Abbildung 4.3 zeigt die reaktive Distanz im Laufe der Simulation.

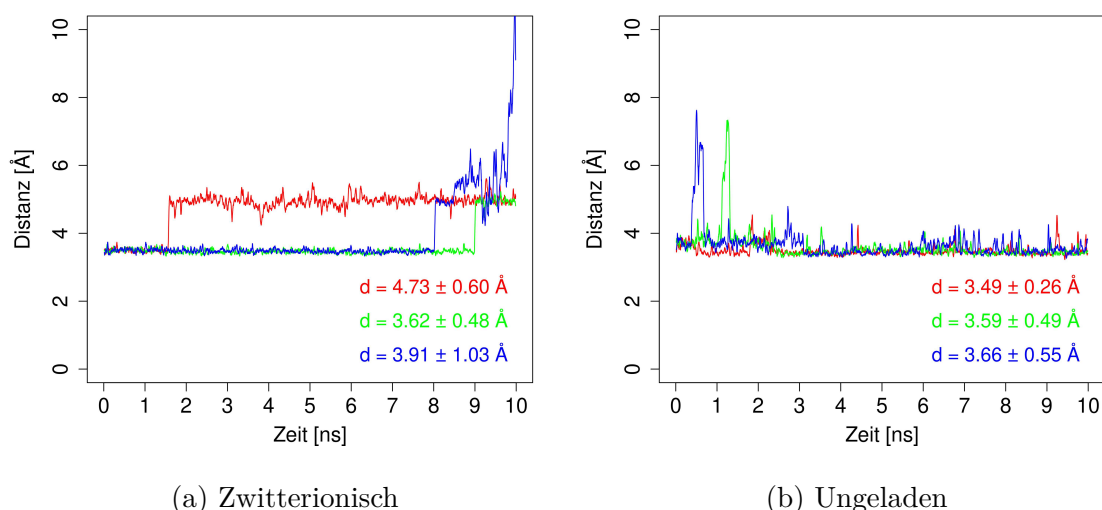


Abbildung 4.3.: Reaktive Distanz von 1ATK im Verlauf der drei Simulationen pro Protonierungszustand.

Sowohl im zwitterionischen Zustand als auch im ungeladenen sind über mehrere Nanosekunden hinweg kurze Abstände zwischen dem katalytischen Cystein und dem elektrophilen Kohlenstoffatom beobachtbar. Allerdings kann für zwei Replikate der zwitterionischen Simulation beobachtet werden, wie der Warhead die Bindetasche verlässt.

#### 4.1.3.2. 1U9X, ein Arylnitril-Inhibitor

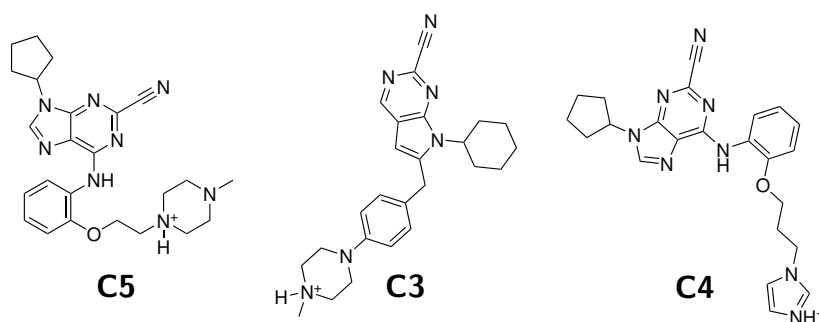


Abbildung 4.4.: Liganden von 1U9X, 1U9V und 1U9W.

1U9X<sup>137</sup> ist ein Komplex von Cathepsin K und einem purinbasierten Nitril-Inhibitor (**C5**), der strukturell sehr ähnlich ist zu den Liganden in 1U9V und 1U9W (**C3** und **C4**, siehe Abbildung 4.4).<sup>137</sup> Auffällig an der Kristallstruktur ist, dass zwei Konformationen der flexiblen Ringsysteme des Liganden modelliert wurden, um der in der Röntgenstrukturanalyse beobachteten Elektronendichte gerecht zu werden. Dies ist wiederum in starker Analogie zu 1U9W, wo der Imidazolring von **C4**, der an einem flexiblen Linker gebunden ist, nicht beobachtbar war. Zur Vorbereitung der Struktur wurde Konformation A von 1U9X orientiert (wobei Konformation B eine genauso valide Wahl gewesen wäre, da beide Konformationen denselben Besetzungsgrad aufweisen). Nach Simulation mit unterschiedlich protonierten katalytischen Einheiten konnten folgende Abstände beobachtet werden (siehe Abbildung 4.5a und Abbildung 4.5b). Zum Vergleich werden ebenfalls die Abstände von 1U9V gezeigt (Abbildung 4.5c und Abbildung 4.5d). Die Ähnlichkeit zwischen den Liganden spiegelt sich ebenfalls in den beobachteten Abständen wider. Erneut zeigt der neutrale Zustand der Bindungstasche im Vergleich zur zwitterionischen Simulation stabilere Abstände zwischen dem reaktiven Cystein und dem Warhead.

4.1. *CatK*: Simulation von nichtkovalenten Modellen basierend auf kovalenten Protein-Ligand-Komplexen

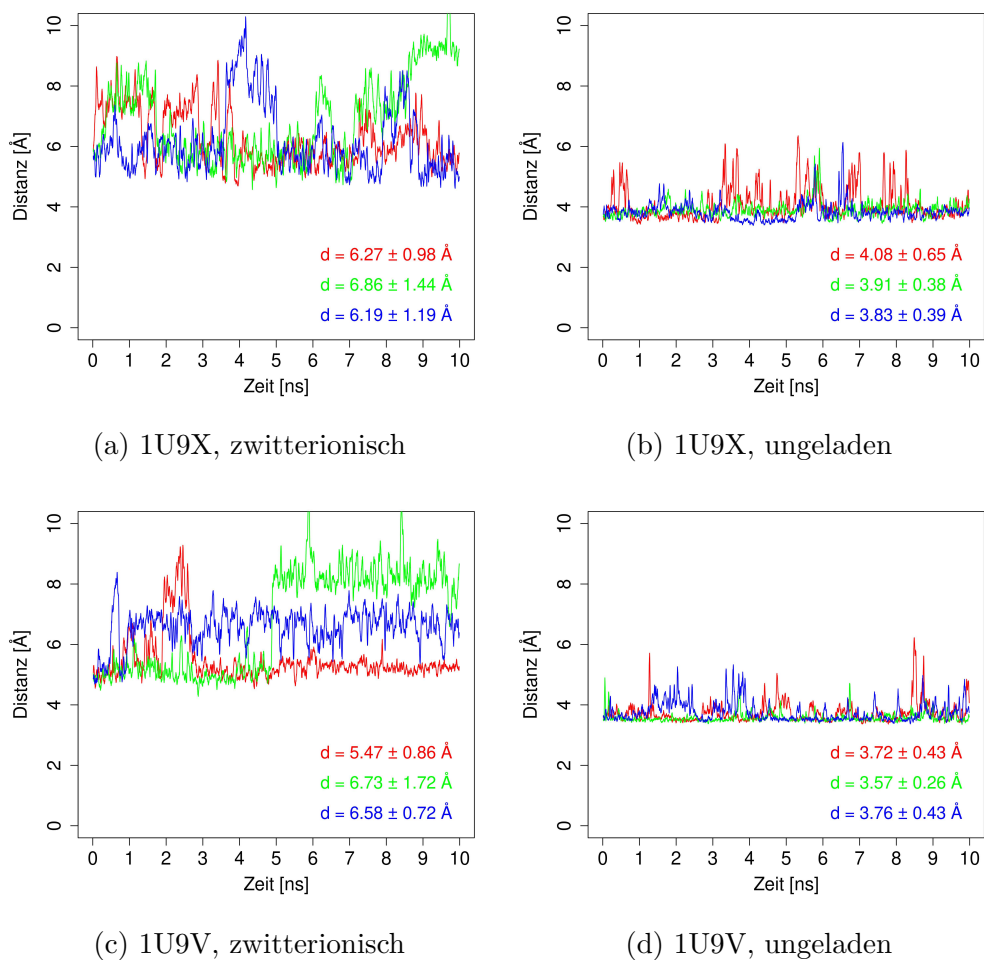


Abbildung 4.5.: Reaktive Distanz von 1U9X und 1U9V im Verlauf der drei Simulationen.

4.1.3.3. Verbesserung der Simulation von 4X6H und 4X6I

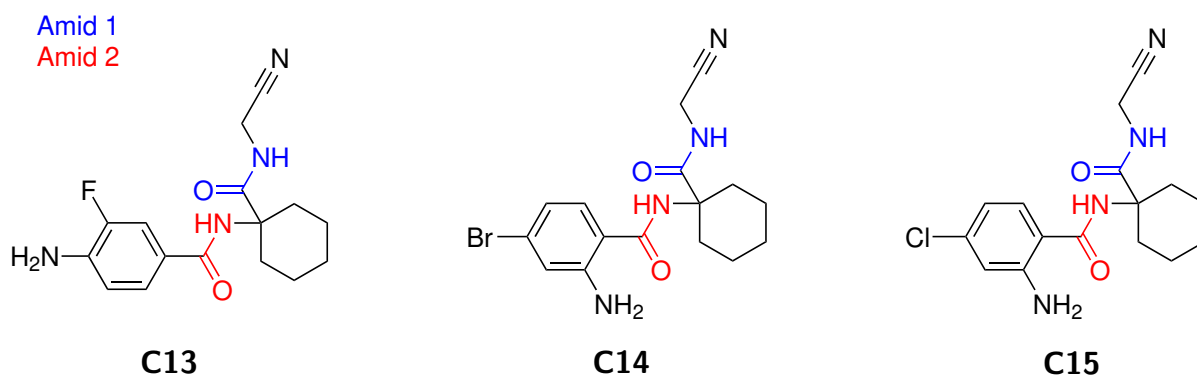


Abbildung 4.6.: Liganden von 4X6H, 4X6I und 4X6J.

Die drei CatK-Inhibitoren, die in 4X6H, 4X6I und 4X6J im Komplex mit CatK kristallisiert sind (**C13**, **C14** und **C15**, siehe Abbildung 4.6), stellen zusammen ein scheinbares Paradoxon dar (siehe Tabelle 4.2).<sup>142</sup> Alle drei Strukturen unterscheiden sich lediglich in der Substitution des terminalen Phenylrings. Daher liegt es nahe, von einer ähnlichen Reaktivität für den Nitril-Warhead am anderen Ende des Moleküls auszugehen. Dies spiegelt sich auch in den vergleichbaren  $K_i$ -Werten der Verbindungen wider. Die Kinetiken konnten experimentell nicht exakt bestimmt werden, sind allerdings qualitativ vergleichbar.<sup>142</sup> Faszinierenderweise sind jedoch deutliche Unterschiede in den Kristallstrukturen bemerkbar: Der 4X6J-Ligand **C15** (Auflösung 1.59 Å) ist kovalent gebunden, während in 4X6I **C14** (Auflösung 1.87 Å) keine kovalente Bindung zwischen den beiden reaktiven Gruppen aufzeigt. Die aufgenommene Elektronendichte von **C13** (PDB 4X6H, Auflösung 1.00 Å) ließ die Modellierung beider Zustände zu, mit 79 % im kovalenten Zustand und 21 % im nichtkovalenten. Wie bereits in der Masterarbeit diskutiert<sup>133</sup> konnten stark abweichende Partialladungen bei

	4X6H	4X6I	4X6J
Auflösung [Å]	1.00	1.87	1.59
Raumgruppe	$P2_12_12_1$	$P2_12_12_1$	$P12_11$
$K_i$ [nM]	$22 \pm 1$	$10 \pm 3$	$16 \pm 2$
$k_{on}$ [ $M^{-1}s^{-1}$ ]	ungefähr $3 \cdot 10^4$		
$k_{off}$ [ $s^{-1}$ ]	ungefähr $1 \cdot 10^{-4}$		
Kovalent gebunden?	79 %	0 %	100 %

Tabelle 4.2.: Fakten zu den 4X6-Komplexen.

den Liganden **C13** und **C14** im Vergleich zu **C15** festgestellt werden, was zu Differenzen in den ausgebildeten Wasserstoffbrückenbindungen führte. Eine mögliche Erklärung hierfür war die unterschiedliche Orientierung der Warheads. Da in den Kristallstrukturen von 4X6H und 4X6I bereits nichtkovalente Strukturen modelliert waren, wurden diese als Startpunkt genutzt. Der Warhead der Liganden deutet dadurch in eine andere Richtung als in 4X6J, zum terminalen aromatischen System hin bzw. in Richtung der Amidbindungen statt von diesen weg. Diese entgegengesetzten Konformationen können zu signifikant unterschiedlichen elektrostatischen Potenzialen in der HF-Rechnung führen, sodass letztendlich abweichende Partialladungen für die Atome der Amidbindungen ermittelt werden.

Bei den in dieser Arbeit erneut durchgeführten Simulationen wurde sichergestellt, dass für die Gaussian-Rechnungen und RESP-Ausgleichsrechnung von **C13** und **C14** die Orientierung des Warheads der von **C15** in 4X6J entspricht. Tabelle 4.3 enthält alle Partialladungen der entsprechenden Amidbindungen von **C13**, **C14** und **C15**. Die in dieser Arbeit berech-



neten Partiaalladungen sind vor allem für N2 deutlich kleiner (vom Betrag her) als zuvor in der Masterarbeit und nun eher vergleichbar mit den Partiaalladungen von **C15**. Es sollte betont werden, dass nur für die Berechnung des elektrostatischen Potenzials die geänderte Orientierung genutzt wurde, für die eigentliche Simulation wurde weiterhin mit der Warhead-Ausrichtung in Richtung des Cysteins gearbeitet.

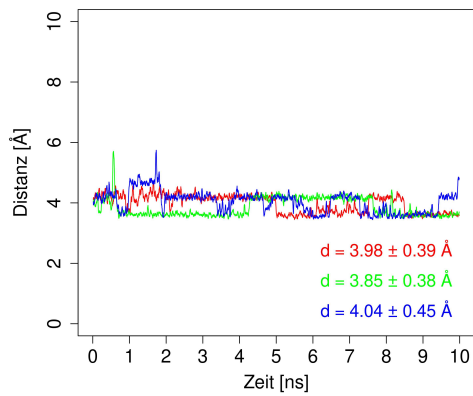
	<b>C13</b> (alt)	<b>C13</b> (neu)	<b>C14</b> (alt)	<b>C14</b> (neu)	<b>C15</b>
O1	-0.57	-0.57	-0.58	-0.58	-0.58
C1	0.61	0.58	0.76	0.63	0.66
N1	-0.50	-0.51	-0.69	-0.54	-0.55
O2	-0.55	-0.53	-0.55	-0.50	-0.51
C2	0.66	0.54	0.76	0.38	0.42
N2	-0.48	-0.22	-0.69	-0.06	-0.10

Tabelle 4.3.: Partiaalladungen in den Amidbindungen der 4X6-Liganden.

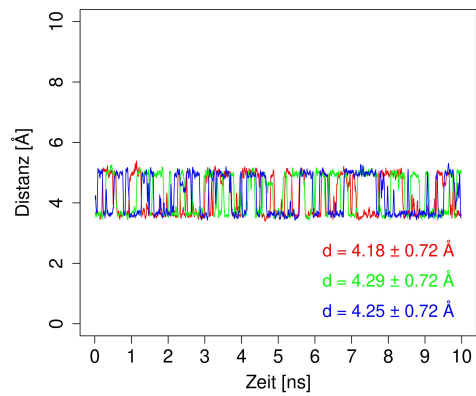
Abbildung 4.7 zeigt die Distanz zwischen elektrophilem Kohlenstoffatom und nukleophilem Schwefelatom für den Datensatz aller drei simulierten Komplexe. In Analogie zu den anderen Nitril-Simulationen sorgt die zwitterionische Protonierung für starke Schwankungen innerhalb des reaktiven Abstandes. Auffällig ist aber, dass (mit den nun korrigierten Partiaalladungen) bei der CYS/HIS-Protonierung qualitativ keine Unterschiede mehr zwischen den drei Komplexen beobachtet werden konnten. Zwei dominante Zustände sind erkennbar: ein kurzer Abstand bei ca. 3.7 Å, und ein weiter entfernter bei ca. 5 Å. Letzterer entspricht der nichtkovalenten Warhead-Orientierung, wie er z. B. in der Kristallstruktur von 4X6I beobachtet werden kann, während ersterer der Warhead-Ausrichtung von 4X6J entspricht, bei der das Nitril zum katalytischen Cystein deutet.

Die MD-Simulationen erlauben daher ebenfalls nicht, eine Aussage darüber zu treffen, warum **C15** kovalent binden sollte, **C14** allerdings nicht. Da für alle drei Simulationen der Warhead vergleichbar oft in Kontakt mit dem katalytischen Cystein ist, liegt es nahe, davon auszugehen, dass die Reaktion für alle Liganden gleich gut möglich ist. Dies ist qualitativ in Übereinstimmung mit den ähnlichen  $K_t$ -Werten und Geschwindigkeitskonstanten. Das einzige Experiment, das daher nahelegt, dass die drei Liganden sich unterschiedlich verhalten, ist somit die (Röntgenbeugung)-Kristallstrukturanalyse.

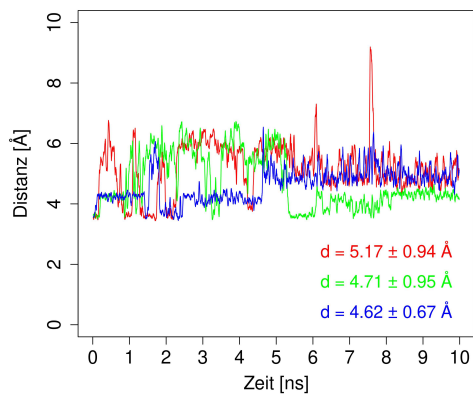
Daher wurden die Kristallstrukturen nochmal kritisch überprüft. 4X6H und 4X6I zeigen beide eine orthorhombische Raumgruppe auf. Hier entsteht in der Nähe der Bindetasche ein Kontakt mit einer benachbarten Einheitszelle, wodurch eine Aminosäure des Symmetriepartners, Asn117\*, in Richtung des Warheads ragt (Abstand 3.4 Å in 4X6I, 3.7 Å in 4X6H).



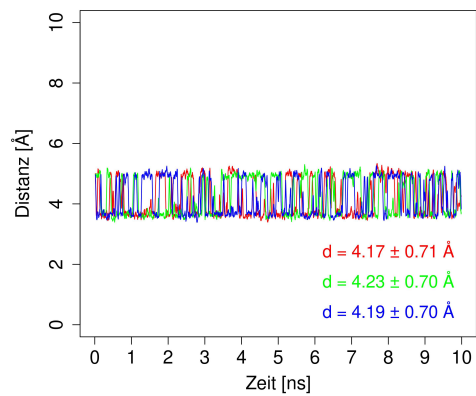
(a) 4X6H, zwitterionisch



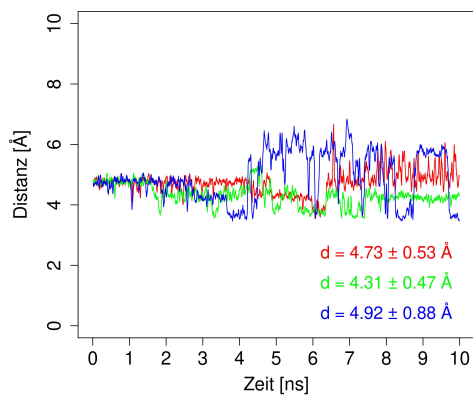
(b) 4X6H, ungeladen



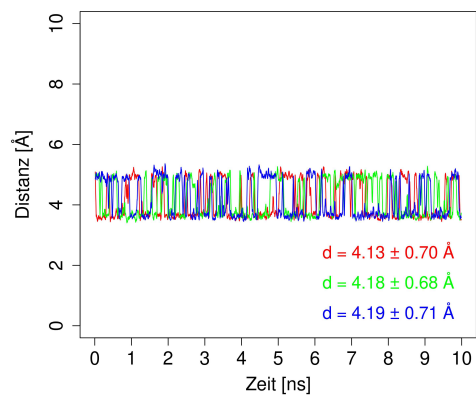
(c) 4X6I, zwitterionisch



(d) 4X6I, ungeladen



(e) 4X6J, zwitterionisch



(f) 4X6J, ungeladen

Abbildung 4.7.: Simulationen der 4X6-Reihe.

Verbrückend zwischen Asn117\* und dem Nitril befindet sich ein Wassermolekül (Abstand zum Nitril 3.0 Å in 4X6I, 2.6 Å in 4X6H). 4X6J kristallisierte hingegen monoklin und zeigte diese Kontakte aufgrund der anderen Kristallpackung nicht auf. Abbildung 4.8 zeigt diese Interaktionen für 4X6J und 4X6I zur Illustration.

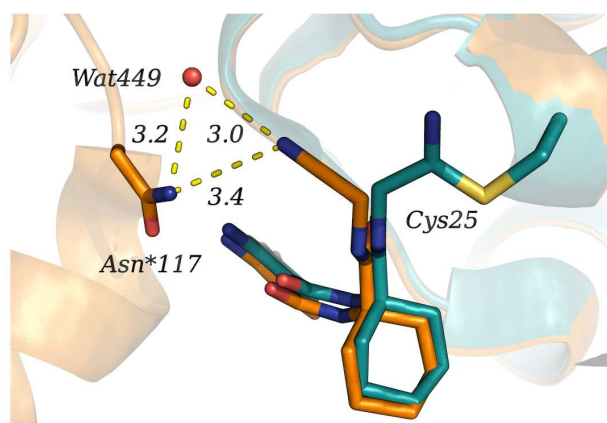


Abbildung 4.8.: Vergleich der Kristallstrukturen von 4X6J (petrol) und 4X6I (orange).

Basierend auf diesen Abständen sollte von schwachen Wasserstoffbrückenbindungen ausgegangen werden, die das Nitril in seinem nichtkovalenten Zustand stabilisieren könnten. Möglicherweise könnten idealere Winkel und Abstände für diese Brücken in die Kristallstruktur modelliert werden, wenn diese Interaktion explizit im Refinement berücksichtigt würde (da Asn117 weit entfernt von der Bindetasche liegt wurde dessen Orientierung wahrscheinlich nicht näher betrachtet). Bei 100 K, der Temperatur des Beugungsexperiments, könnte diese schwache Stabilisierung möglicherweise ausreichen, um die nichtkovalente Orientierung zeitgleich mit der kovalenten Form sichtbar oder gar prävalent werden zu lassen. Bei 4X6I sind kürzere Abstände beobachtbar als bei 4X6H, und in Analogie ist bei 4X6I hauptsächlich die nichtkovalente Form beobachtbar, während diese bei 4X6H nur ca. 20 % ausmacht.

Diese stabilisierenden Wasserstoffbrückenbindungen aufgrund des Kristallkontakts werden nicht in Lösung ausgebildet, was erklären würde, warum in keinem nasschemischen Experiment ein Unterschied zwischen den drei Verbindungen beobachtet werden kann. Diese Beobachtungen stellen zwar keinen harten Beweis dar, sind aber nach momentanen Stand die beste Erklärung, warum 4X6J in der Kristallstruktur lediglich kovalent sichtbar ist, 4X6H und 4X6I hingegen auch nichtkovalent.

#### 4.1.3.4. 6QBS, ein Alkin-Inhibitor

6QBS<sup>147</sup> ist die Kristallstruktur von Cathepsin K, das durch den Alkin-Inhibitor **C20** irreversibel komplexiert wurde. Die Autoren dieser Struktur stellten eine These sehr ähnlich zu der vorliegenden Arbeit auf: Reaktivität wird nicht nur durch elektronische Eigenschaften wie Elektrophilie eines Warheads getrieben, sondern auch durch die erreichbare Nähe der miteinander reagierenden Gruppen. Die Liganden, die sie zur Untermauerung dieser These synthetisierten, basieren auf Odanacatib **C16** als Grundstruktur (der Ligand aus der Kristallstruktur 5TDI, siehe Abbildung 4.9), substituieren aber das Nitril als Warhead mit einem Alkin. Zur Verbesserung der Löslichkeit wurde das Fluoratom der *iso*-Butylgruppe durch ein Wasserstoffatom ersetzt.<sup>147</sup>

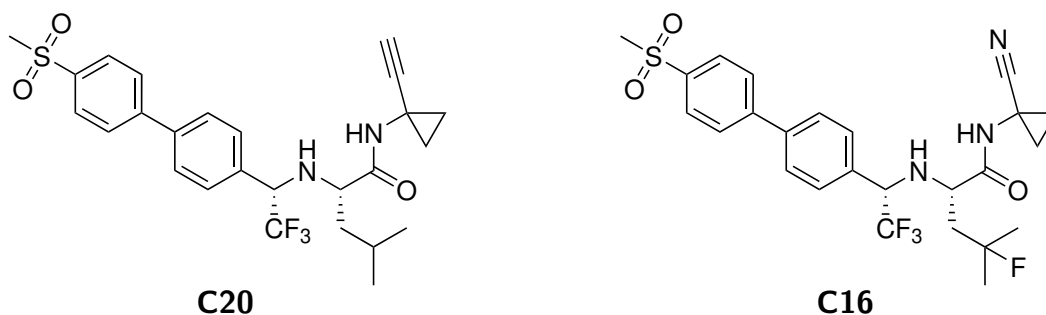


Abbildung 4.9.: Liganden von 6QBS und 5TDI.

Abbildung 4.10 zeigt die Abstände der reaktiven Atome während der zwitterionischen und neutralen Simulation, im Vergleich zu 5TDI. Zusätzlich werden in Abbildung 4.11 zur besseren Illustration Histogramme der Abstände der 6QBS- und 5TDI-Simulationen gezeigt. In der zwitterionischen Simulation (Abbildung 4.10a und Abbildung 4.11a) ist vor allem ein Abstand bei ca. 4.5 Å dominant. Derselbe Abstand ist auch in der ungeladenen Simulation beobachtbar, allerdings ist zusätzlich ein weiterer Abstand knapp unter 4 Å vertreten (Abbildung 4.10b und Abbildung 4.11b). Dieser Zustand ist ebenfalls in der ungeladenen Simulation von 5TDI beobachtbar (Abbildung 4.10d und Abbildung 4.11d). Das bedeutet, dass der Warhead von **C20** zwischen zwei unterschiedlichen Zuständen alterniert, sehr vergleichbar zu den Simulationen der 4X6-Reihe (siehe Kapitel 4.1.3.3). Hier scheint jedoch einer der beiden Zustände deutlich gegenüber dem anderen präferiert zu sein. Es ist nur anhand der Abstandsbetrachtung nicht nachvollziehbar, warum der Warhead in **C16**, der als Design-Grundlage für **C20** genutzt wurde, dieses alternierende Verhalten nicht aufzeigt. Um diese Unterschiede besser verstehen zu können, wurde die Torsionsbarriere der entsprechenden Bindung für kleine Modellsysteme der Warheads von 4X6H/I/J, 4DMX bzw. 5TDI (identi-

#### 4.1. CatK: Simulation von nichtkovalenten Modellen basierend auf kovalenten Protein-Ligand-Komplexen

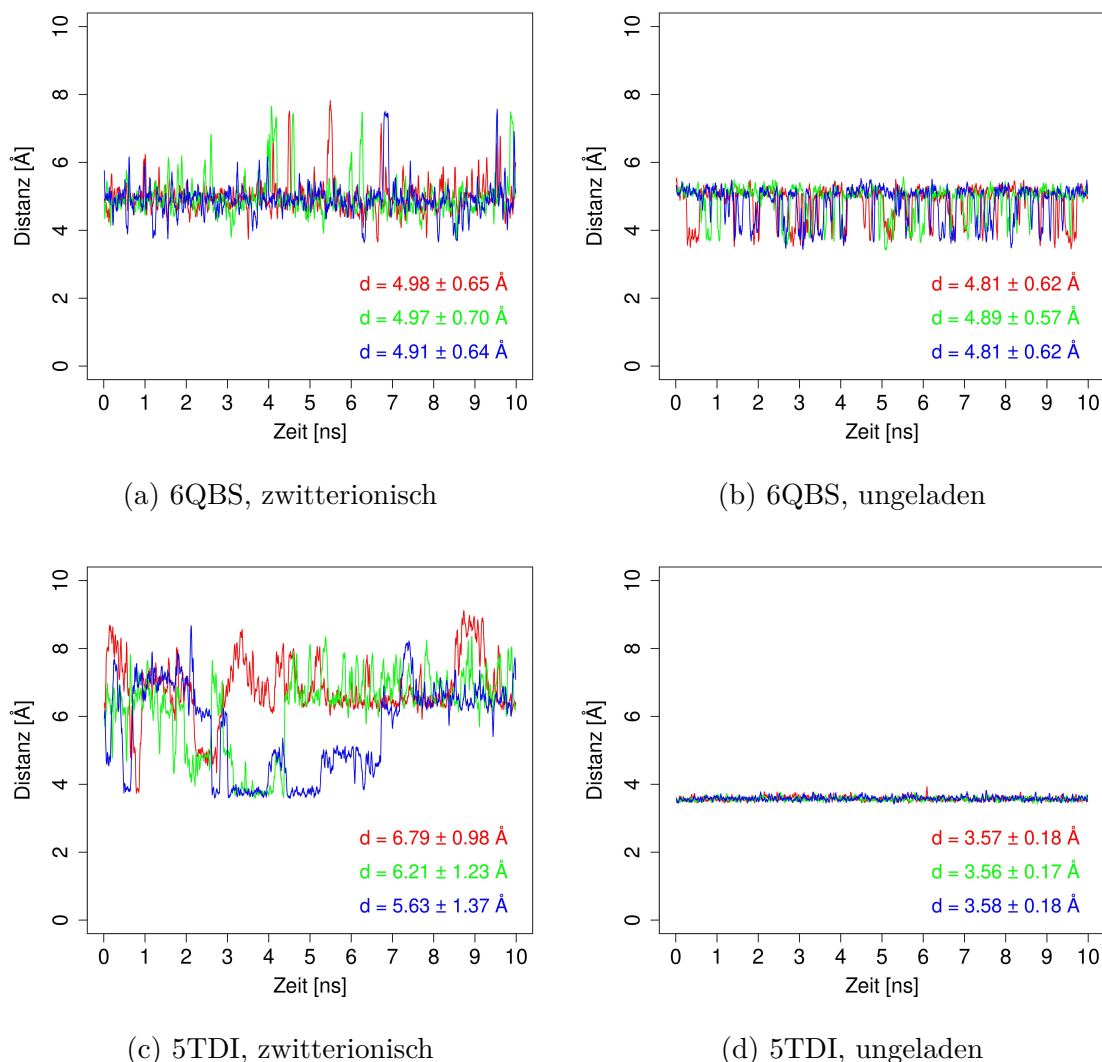


Abbildung 4.10.: Reaktiver Abstand in den 6QBS-Simulationen im Verlauf der Zeit und im Vergleich zu 5TDI.

sche Warheads) und 6QBS berechnet (siehe Abbildung 4.12b). Hierfür wurde in Gaussian09 auf dem B3LYP/6-311+G\*\* Level ein Scan des Diederwinkels in 10°-Schritten durchgeführt. Die erhaltenen Energie-Profile zeigen, dass der Warhead mit einem Cyclopropyl-Ring ein zusätzliches Maximum aufweist. Diese zusätzliche Barriere erklärt qualitativ, warum für 4DMX/5TDI keine Rotation beobachtbar ist, macht allerdings nicht klar, wieso 6QBS den weiter entfernten Zustand präferiert. Die in Kapitel 4.1.4 durchgeführte Analyse zeigt, dass der Warhead in der 5TDI-Struktur besser durch Wasserstoffbrückenbindungen stabilisiert werden kann. Das Alkin in 6QBS ist hierzu nicht in der Lage.

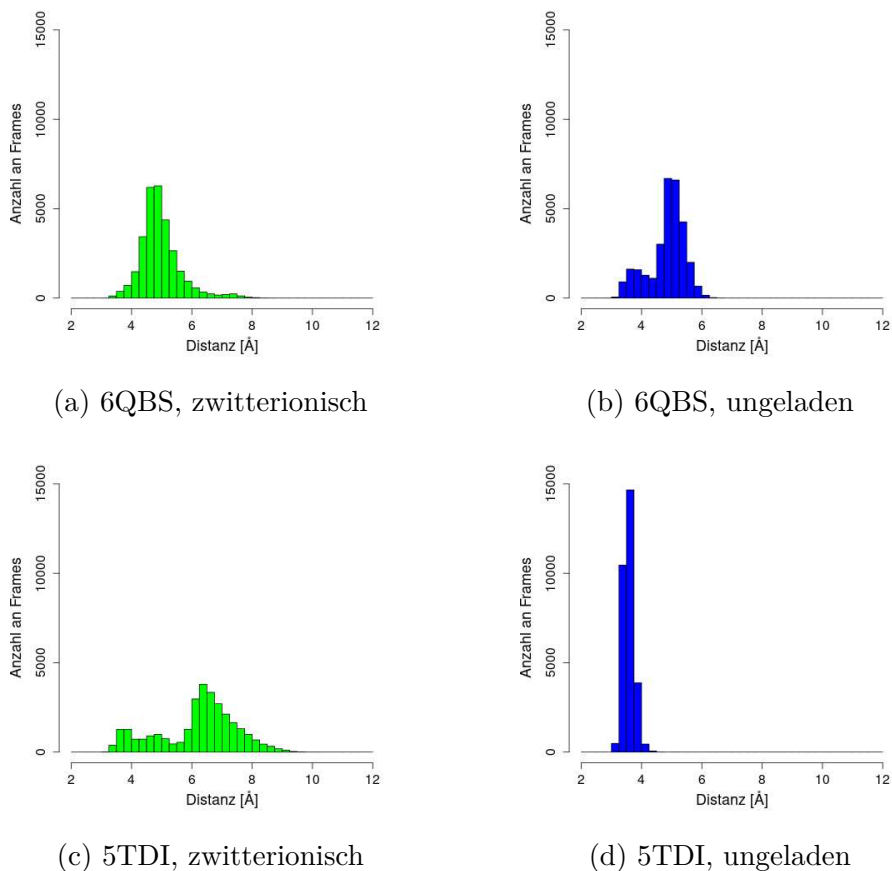


Abbildung 4.11.: Reaktiver Abstand in den 6QBS-Simulationen als Histogramm im Vergleich zu 5TDI.

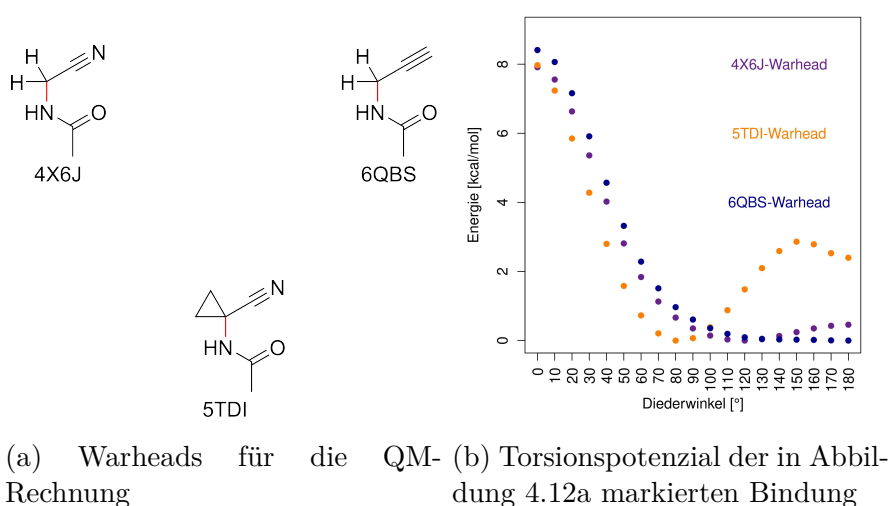


Abbildung 4.12.: Vergleich des Torsionspotenzials verschiedener Warheads.

#### 4.1.3.5. Zusammenfassung der gesamten Simulationsergebnisse

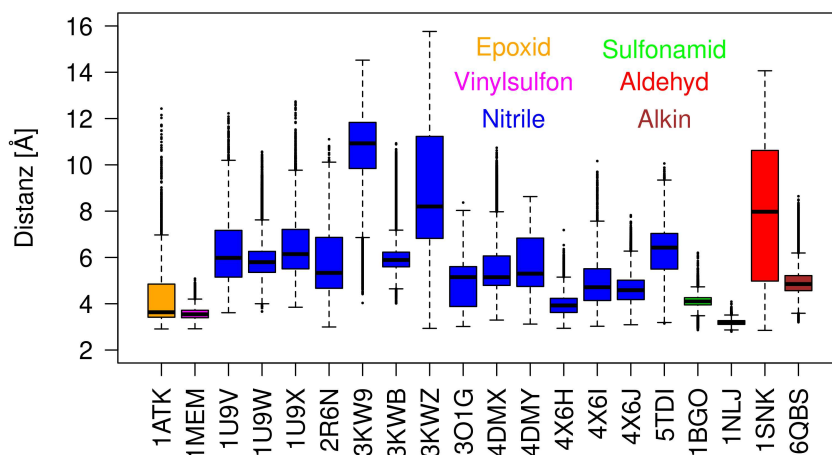
Zusammen mit den Simulationen, die bereits in der Masterarbeit erfolgt sind, ergibt sich Abbildung 4.13, die die Boxplots des reaktiven Abstandes in den Simulationen darstellt. Bei der zwitterionischen Protonierung (siehe Abbildung 4.13a) ergeben sich, insbesondere für die Nitrile, hohe Median-Werte und Interquartilabstände. Dies ist ein Zeichen dafür, dass der Warhead nicht stabil in der Bindetasche positioniert ist. Bei der ungeladenen Protonierung hingegen werden hauptsächlich Mediane bei etwa 3.7-4 Å gefunden und geringere Interquartilabstände beobachtet (siehe Abbildung 4.13b). Wenn davon ausgegangen wird, dass es für eine Reaktion notwendig ist, dass der Warhead und das Nukleophil nahe beieinander positioniert sind, würde somit die CYS/HIS-Protonierung der katalytischen Einheit ein plausibleres Bild für die Nitrile ergeben.

Für die anderen Warhead-Typen lassen sich nur bedingt Tendenzen ableiten, da meist nur ein Vertreter im Datensatz auftritt. Das Aldehyd, 1SNK, scheint allerdings, ebenfalls wie die Nitrile, deutlich geringere Abstände in der ungeladenen Bindetasche einzunehmen als in der zwitterionischen. Das Epoxid **C1** in 1ATK, das Vinylsulfon **C2** in 1MEM und die beiden Sulfonamide **C17** und **C18** in 1BGO und 1NLJ zeigen unabhängig vom Protonierungszustand niedrige Mediane für den hier diskutierten Abstand. Das Alkin **C20** in 6QBS hat bei beiden Zuständen Medianwerte über 4 Å; wie in Kapitel 4.1.3.4 allerdings diskutiert, setzen sich die Abstände im neutralen Zustand aus zwei unterschiedlichen Positionierungen des Warheads zusammen, von denen eine vergleichbar ist mit den Abständen der Nitrile.

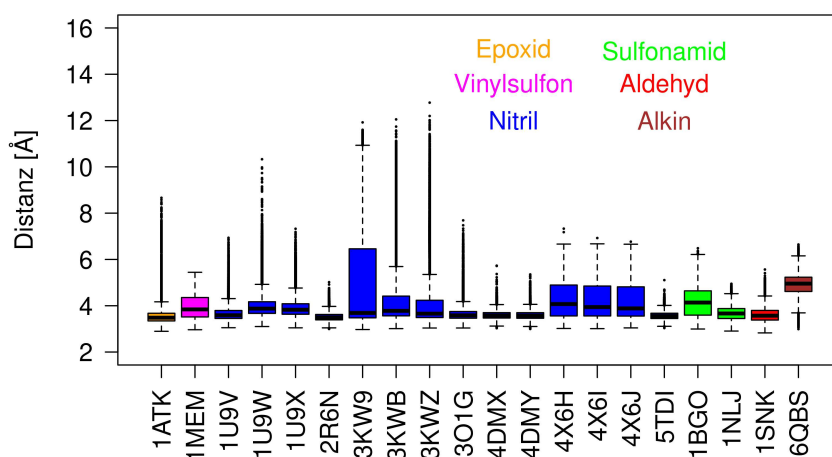
Für die hier betrachteten, durch verschiedenste Methoden optimierte Inhibitoren wird somit zusammenfassend bevorzugt ein Abstand von 3.7-4 Å oder niedriger beobachtet. Dies bedeutet aber nicht zwangsläufig, dass ein Ligand zwingend diesen Abstand in einer Simulation im Median erfüllen muss, um eine Reaktion zu durchlaufen. Es wäre theoretisch denkbar, dass mit weniger optimierten Liganden, die ihren Warhead weiter entfernt als 4 Å platzieren, im Experiment immer noch kovalente Inhibition beobachtbar ist. Die Simulationen der 4X6-Komplexe (siehe Kapitel 4.1.3.3) bzw. von 6QBS (siehe Kapitel 4.1.3.4) legen zusätzlich nahe, dass der geringe Abstand zur katalytischen Einheit nicht permanent aufrecht erhalten werden muss. Daher kann der hier wiederholt in Simulationen beobachtete Abstand von 3.7-4 Å eher als „optimaler“ Abstand verstanden werden, den ein kovalent reagierender Ligand erreichen können sollte, nicht als zwingend dauerhaft zu erfüllendes Kriterium.

Nichtsdestotrotz liegt es nahe, dass man diesen reaktiven Abstand im Rahmen eines virtuellen Screenings nutzen kann, um zu selektieren, welche Strukturen Potenzial haben, ihr Target kovalent zu inhibieren, und welche ihren Warhead ungünstig positionieren und eli-

miniert werden sollten. Dies könnte helfen, falsch-positive Vorhersagen zu vermeiden. Dieses Vorgehen wird in Kapitel 4.1.7.2 diskutiert.



(a) Zwitterionisch



(b) Ungeladen

Abbildung 4.13.: Boxplots des reaktiven Abstandes aller Simulationen.

#### 4.1.4. Analyse des Wasserstoffbrückennetzwerks der katalytischen Einheit

Die bisher gezeigten Ergebnisse legen nahe, dass zumindest für Nitrile eine stabilere Koordination des Warheads innerhalb der Bindetasche erreicht wird, wenn von einer CYS/HIS-Protonierung für die katalytische Einheit ausgegangen wird. Allerdings wird nicht klar, wor-



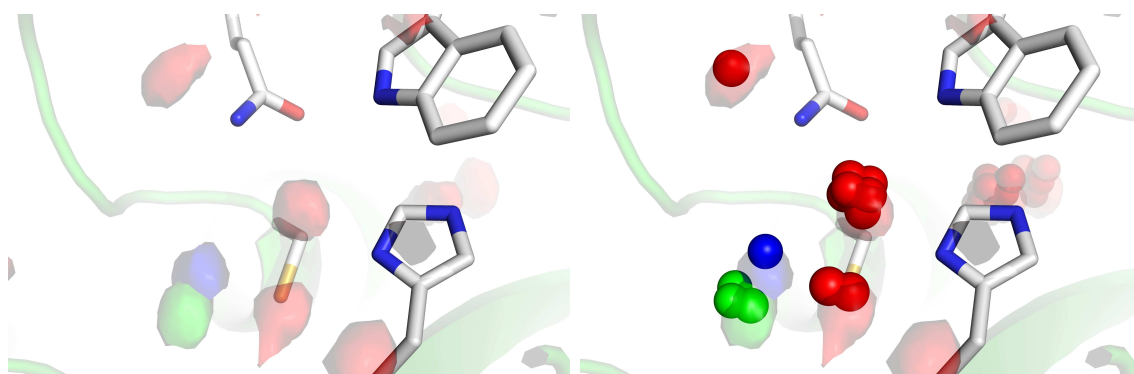
auf diese Stabilisierung beruht. Eine visuelle Inspektion der Simulationen legt nahe, dass das Solvatisierungsmuster aufgrund der unterschiedlichen Protonierung verändert ist. Um dies genauer zu untersuchen, wurden im wesentlichen zwei Ansätze genutzt: ein gitterbasierter Ansatz zur besseren Lokalisation der Wassermoleküle und ein zeitlich auflösender Ansatz, um die Fluktuation innerhalb des Netzwerks besser nachvollziehen zu können.

#### 4.1.4.1. Gitter-Ansatz

Für den Gitter-Ansatz wurden alle Replikate eines Simulationssatzes auf eine Referenz aligniert. Um den Vergleich zwischen den unterschiedlichen Simulationen zu erleichtern, wurde hierfür eine gemeinsame Referenz für alle Simulationen verwendet, nämlich die Momentaufnahme der 5TDI-Simulationen, bei denen das Protein-Rückgrat den geringsten RMSD-Wert zu den durchschnittlichen Koordinaten der Trajektorie aufwies (prinzipiell wären aber auch andere Simulationen oder andere Referenzen denkbar gewesen). Die Größe des Gitters wurde mit Hilfe des *cptraj*-Befehls *bounds* festgelegt.<sup>151,154</sup> Hierfür wurden die Extremwerte der Koordinaten des Liganden sowie der Aminosäuren Gln19, Cys25, His162 und Trp184 herangezogen. Anschließend wurden für jedes Sauerstoffatom eines Wassermoleküls sowie für das Nitril-Kohlenstoffatom und das Nitril-Stickstoffatom Dichten mit dem *grid*-Befehl berechnet.<sup>151</sup> Die Kantenlänge eines Gitterwürfels betrug dabei 0.5 Å. Zur besseren Lokalisation wurden Pseudo-PDB-Dateien erstellt, die alle Gitterkoordinaten enthielten, die mindestens 80% der maximal beobachteten Dichte aufwiesen. Für die Wasser-Dichten wurde dieser Prozentsatz auf 20% gesenkt, da strukturelles Wasser innerhalb des Proteins ein deutlich höheres Gridmaximum erreichte als das für diese Analyse interessante an der Oberfläche befindliche Wasser. Durch *fconv*<sup>155</sup> wurde für diese Pseudo-PDB-Dateien ein Clustering durchgeführt, um nahe beieinanderliegende Gitterpunkte zu gruppieren. Für jeden Cluster wurden repräsentative Koordinaten ermittelt, indem die Koordinaten basierend auf dem vorher bestimmten Prozentsatz gewichtet und anschließend gemittelt wurden. Die so erhaltenen Cluster werden in den Abbildungen 4.14 und 4.16 für das Wasser-Sauerstoffatom rot, für das Nitril-Kohlenstoffatom grün und für das Nitril-Stickstoffatom blau dargestellt.

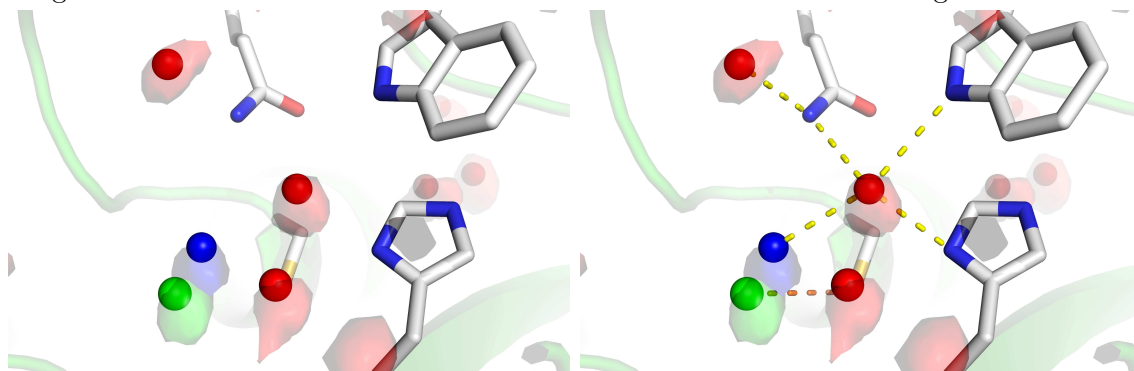
Abbildung 4.14 zeigt die eben beschriebenen Schritte im Rahmen dieser Auswertung. Das Ergebnis für jede einzelne Simulation kann im Anhang gefunden werden. Exemplarisch wird das Ergebnis von **C11**, dem Liganden in 4DMX (siehe Abbildung 4.15), in Abbildung 4.16 gezeigt.

Bei sowohl den zwitterionischen als auch den ungeladenen Simulationen ist ein zentraler Wasser-Cluster beobachtbar, der zwischen Gln19, His162 und Trp184 platziert ist. Die Distanzen zu diesen potenziellen Partnern für Wasserstoffbrückenbindungen liegen zwischen



(a) Die aus dem *grid*-Befehl erhaltene Dichte. Die farbigen Oberflächen markieren hier- bei Bereiche mit einer vierfachen Dichte im Vergleich zu reinem Wasser.

(b) Alle Gridpunkte, die 80% der maximal beobachteten Dichte innerhalb eines Grid- punkts erreicht haben. Für Wasser wurde dieser Prozentsatz auf 20% gesenkt.



(c) Nach dem Clustering mit *fconv* wird ein gewichteter Mittelwert aller Clusterkoordinaten gebildet.

(d) Basierend auf diesen Koordinaten können Distanzen bestimmt werden.

Abbildung 4.14.: Die einzelnen Schritte des für die Analyse genutzten Gitter-Ansatzes. Zusätzlich zu den generierten Dichten und Clustern sind die durchschnittlichen Koordinaten des Proteins gezeigt (d. h. keine tatsächliche Momentaufnahme der Simulation).

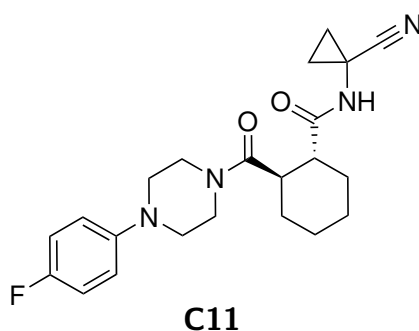


Abbildung 4.15.: Ligand von 4DMX.

#### 4.1. *CatK*: Simulation von nichtkovalenten Modellen basierend auf kovalenten Protein-Ligand-Komplexen

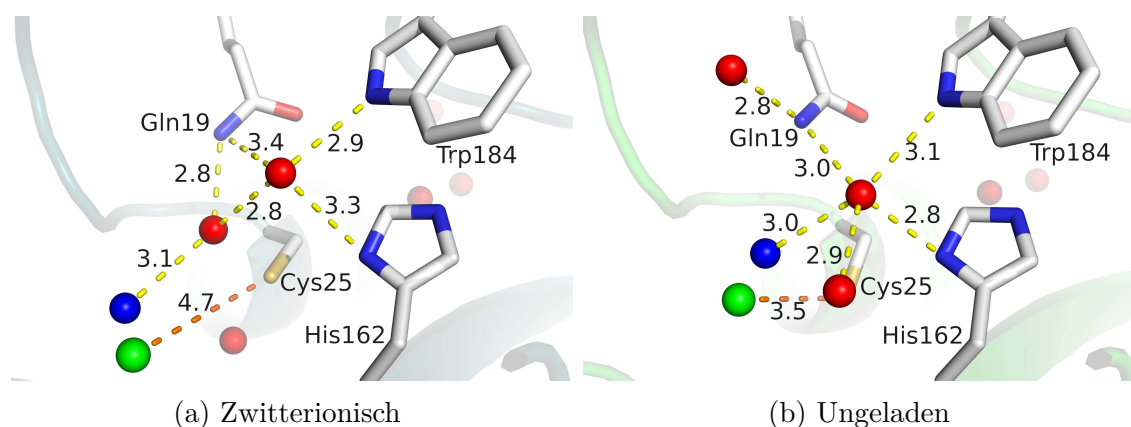


Abbildung 4.16.: Gitteranalyse der 4DMX-Simulationen. Rote Kugeln zeigen die Position der Wasser-Cluster, während blaue bzw. grüne Kugeln die Position des Clusters des Nitril-Stickstoff- und Kohlenstoffatoms markieren.

2.8 Å und 3.4 Å. In der zwitterionischen Simulation (Abbildung 4.16a) liegt zwischen diesem zentralen Wasser-Cluster und dem Nitril-Stickstoff-Cluster ein weiterer Wasser-Cluster vor, der somit nahe am negativ geladenen Cys25 platziert ist und den Weg zu diesem blockiert. Für das Nitril-Kohlenstoffatom kann in diesem Beispiel daher ein Abstand von 4.7 Å zum reaktiven Cystein beobachtet werden. Näher kann das Nitril nicht an Cys25 platziert werden, ohne mit dem Wasser-Cluster zu konkurrieren. Aufgrund dieser Distanz und den im Weg befindlichen Wassermolekülen scheint eine Reaktion unwahrscheinlich. Offensichtlich wird durch die zwitterionische Protonierung der katalytischen Diade die Koordination mit dem stärkeren Dipol der Wassermoleküle bevorzugt.

Dieses Bild ändert sich allerdings, wenn die Gitteranalyse des ungeladenen Zustands (Abbildung 4.16b) betrachtet wird. Es ist immer noch ein zentraler Wasser-Cluster zwischen Gln19, His162 und Trp184, allerdings ist der zweite Wasser-Cluster an Cys25 nicht beobachtbar, da nun das Schwefelatom nicht mehr negativ geladen ist. Stattdessen ist das Nitril in Richtung der reaktiven Aminosäuren verschoben. Der Stickstoff-Cluster ist in diesem Beispiel 3.0 Å von dem zentralen Wassermolekül entfernt, sodass Wasserstoffbrückenbindungen möglich sind. Durch diese Koordination ist das Kohlenstoffatom des Nitrils ebenfalls wesentlich näher an Cys25 platziert (3.5 Å), ohne dazwischen befindliche Wasser-Cluster. Ein Angriff des Schwefelatoms an das elektrophile Nitril-Kohlenstoffatom wäre somit denkbar.

Das oben beschriebene Bild ist qualitativ bei den Simulationen anderer Nitrile ebenfalls beobachtbar. Tabelle 4.4 beinhaltet die Abstände des Nitril-Kohlenstoffatom-Clusters zum reaktiven Cystein. (Es sollte hierbei betont werden, dass die Koordinaten des Cysteins, die für diese Abstandsmessung genutzt wurden, die durchschnittlichen Koordinaten während

aller drei Replikate sind und nicht aus dem zuvor beschriebenen Gitter-Ansatz stammen.) In nur wenigen Fällen wird der beobachtete Trend durchbrochen; die ungeladene Bindetasche führt zu kurzen Abständen zwischen den reaktiven Atomen, während die zwitterionische Protonierung zu deutlich erhöhten Abständen führt.

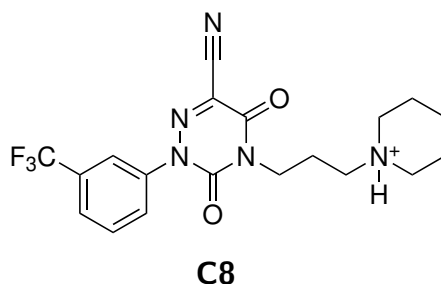


Abbildung 4.17.: Ligand in 3KWB.

Die CYS/HIS-Simulation von 3KWB unterscheidet sich von den anderen Trajektorien, da hier der Warhead stellenweise mit dem zentralen Wasser-Cluster konkurriert. Ein möglicher Grund hierfür könnte der sehr elektronenarme Heteroaromat des Liganden **C8** sein (siehe Abbildung 4.17), der das Dipolmoment des Nitrils erhöht. Die zwitterionische Simulation

Komplex	Ligand	CYM/HIP [Å]	CYS/HIS [Å]
1U9V	<b>C3</b>	5.0	3.4
1U9W	<b>C4</b>	5.7	3.7
1U9X	<b>C5</b>	5.6	3.7
2R6N	<b>C6</b>	3.6/4.8/7.2	3.5
3KW9	<b>C7</b>	9.8	3.3
3KWB	<b>C8</b>	5.9	3.8*
3KWZ	<b>C9</b>	7.0	3.3
3O1G	<b>C10</b>	5.1	3.5
4DMX	<b>C11</b>	4.7	3.5
4DMY	<b>C12</b>	4.5	3.7
4X6H	<b>C13</b>	3.8**	3.6
4X6I	<b>C14</b>	3.8**	3.6
4X6J	<b>C15</b>	4.5	3.6
5TDI	<b>C16</b>	6.6	3.5

Tabelle 4.4.: Distanz zwischen Schwefelatom von Cys25 und Nitril-Kohlenstoffatom-Cluster. \* 3KWB konkurriert in der neutralen Simulation mit dem zentralen Wassercluster. \*\* Die beiden Liganden besetzen nicht mehr korrekt die lipophile Tasche des Proteins.

von **C6** führte zu insgesamt drei unterschiedlichen Clustern an Nitril-Kohlenstoffatomen, da es aufgrund der starken Fluktuation des Warheads an mehreren Gitterpunkten etwa

gleich oft beobachtet werden kann. Bei den Simulationen von 4X6H und 4X6I (**C13** und **C14**) bleibt das Nitril-Kohlenstoffatom zwar in der Nähe des negativ geladenen Cysteins, allerdings ist der Ligand deutlich im Vergleich zu seiner Kristallstruktur-Pose verschoben; der Cyclohexyrling verlässt teilweise die tiefe lipophile Tasche.

Der zentrale Wasser-Cluster ist bei beiden Protonierungszuständen prominent und scheint indirekt Wechselwirkungen zwischen Gln19, His162 und Trp184 an andere Wasserstoffbrücken-Akzeptoren vermitteln zu können. Tabelle 4.5 unterstreicht, dass dieser Cluster in allen Simulationen an derselben Stelle beobachtbar ist. Im Vergleich zur ungeladenen Simulation ist der Wassercluster in der zwitterionischen Simulation allerdings ca. 0.6 Å weiter von His162 entfernt (von ca. 2.8 Å zu ca. 3.4 Å verschoben), möglicherweise, um den zweiten Wasser-Cluster zu ermöglichen.

Komplex	Zwitterionisch			Ungeladen		
	Gln19 [Å]	His162 [Å]	Trp184 [Å]	Gln19 [Å]	His162 [Å]	Trp184 [Å]
1U9V	3.4	3.4	2.9	3.1	2.8	3.1
1U9W	3.3	3.4	2.9	3.0	2.8	3.1
1U9X	3.4	3.4	2.9	3.0	2.8	3.2
2R6N	3.3	3.3	3.0	3.2	2.8	3.0
3KW9	3.3	3.6	2.9	3.1	2.8	3.0
3KWB	3.3	3.4	2.9	2.9	2.7	3.4
3KWZ	3.4	3.5	2.9	3.2	2.8	3.0
3O1G	3.4	3.5	2.9	3.6	2.8	2.9
4DMX	3.4	3.3	2.9	3.0	2.8	3.1
4DMY	3.4	3.5	2.9	3.1	2.8	3.1
4X6H	3.3	3.3	2.9	3.0	2.8	3.2
4X6I	3.3	3.4	2.9	3.0	2.8	3.1
4X6J	3.3	3.4	2.9	3.0	2.8	3.2
5TDI	3.3	3.3	3.0	3.1	2.8	3.2

Tabelle 4.5.: Distanz des zentralen Wasser-Clusters zu naheliegenden Aminosäuren.

Für andere Warhead-Typen wurden vergleichbare Auswertungen durchgeführt:

- Für 1SNK wurden das Kohlenstoff- und Sauerstoffatom des Aldehyds betrachtet.
- Für 1MEM wurden das  $\beta$ -Kohlenstoffatom und die beiden Sauerstoffatome des Sulfonylvinylsystems erfasst.
- Für die Analyse von 1BGO und 1NLJ wurden neben dem reagierenden Kohlenstoffatom die Sauerstoffatome des Sulfonamids betrachtet.

Abbildung 4.18 zeigt die Resultate der gitterbasierten Analyse für die Simulationen von 1SNK. Dasselbe Netzwerk an Interaktionen wie für die Nitrile kann für die ungeladenen Simulationen beobachtet werden, und für die zwitterionische Simulation ist ebenfalls ein zusätzliches Wassermolekül beobachtbar, das als Abstandshalter fungiert. Abbildung 4.19 zeigt die reaktiven Abstände während der Simulationen im Verlauf der Simulationszeit. Es wird deutlich, dass der Warhead innerhalb des Großteils der drei zwitterionischen Replika (Abbildung 4.19a) nicht stabil koordiniert ist, sondern frei beweglich. Einzige Ausnahme ist hierzu die zweite Hälfte der rot gezeichneten Trajektorie. Hier findet eine direkte Interaktion zwischen dem positiv geladenen His162 und dem Aldehyd-Sauerstoffatom statt. Diese Koordination wird vom gitterbasierten Auswertungsansatz als häufigste Position des Warheads erfasst und in Abbildung 4.18a dargestellt. Während den restlichen Simulationen ist eigentlich ein Wassermolekül an dieser Position. Dies erklärt die in rot dargestellte Kollision mit einem Wassercluster in Abbildung 4.18a.

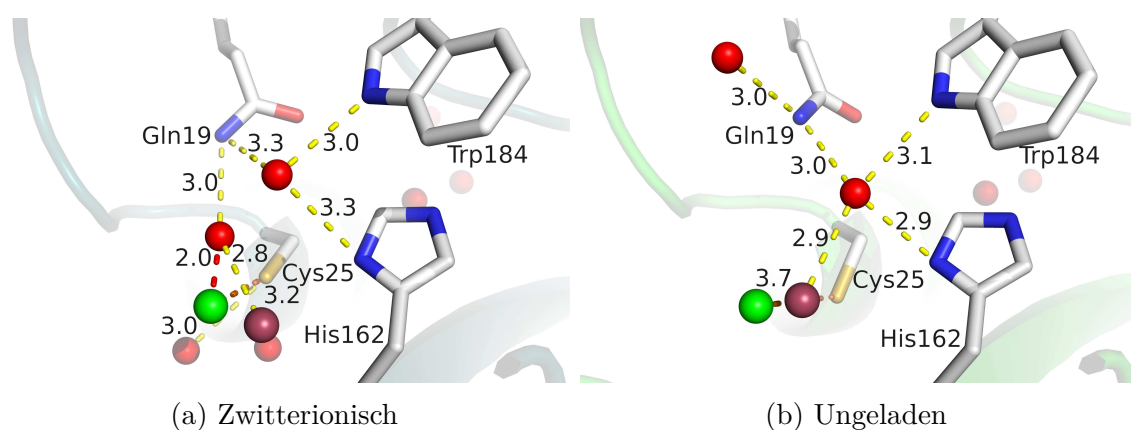


Abbildung 4.18.: Gitteranalyse der 1SNK-Simulationen. Die erfassten Wassercluster sind hellrot, der Cluster für das Aldehyd-Kohlenstoffatom ist grün, und der für das Aldehyd-Sauerstoffatom dunkelrot dargestellt.

Qualitativ werden somit für das Aldehyd die Ergebnisse der Nitrile reproduziert: Eine stabile Positionierung des Warheads im Einklang mit dem Solvatationsmuster wird während der ungeladenen Simulation erreicht (siehe Abbildung 4.19b), während die zwitterionische Protonierung zu einem fluktuierenden Aldehyd führt.

Die Liganden, die eine Sulfonylfunktion beinhalten, zeichnen hierbei jedoch ein komplett anderes Bild als die bisher diskutierten Warheads. Abbildung 4.20 zeigt dies anhand der Ergebnisse der Gitteranalyse für 1MEM.

Die Sulfonyl-Sauerstoffatome sind in der Lage, den bei den Nitrilen charakterisierten zentralen Wasser-Cluster zu verdrängen und simultane Wasserstoffbrückenbindungen zu Gln19

#### 4.1. CatK: Simulation von nichtkovalenten Modellen basierend auf kovalenten Protein-Ligand-Komplexen

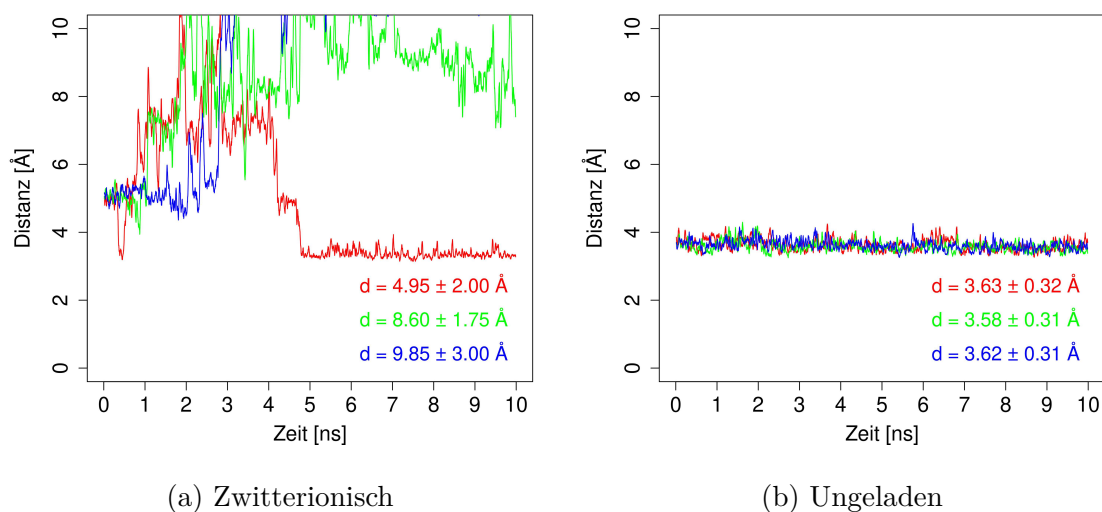


Abbildung 4.19.: Reaktive Distanz des 1SNK-Komplexes im Verlauf der Simulationszeit.

und Trp184 auszubilden. Dies kann unabhängig vom Protonierungszustand erfolgen (siehe Abbildung 4.20). Ähnliches trifft auf 1BGO (Abbildung 4.21) und 1NLJ (Abbildung 4.22) zu, allerdings scheint hier die Koordination schwächer zu sein, da zum Teil der zentrale Wasser-Cluster ebenfalls beobachtbar ist (was darauf hindeutet, dass zumindest kurzzeitig die Sulfonyleinheit durch das Solvens verdrängt wurde). Interessanterweise scheint dies nicht zwangsläufig zu starken Bewegungen des elektrophilen Kohlenstoffatoms zu führen, da dieses hauptsächlich immer noch in der Nähe von Cys25 beobachtbar ist.

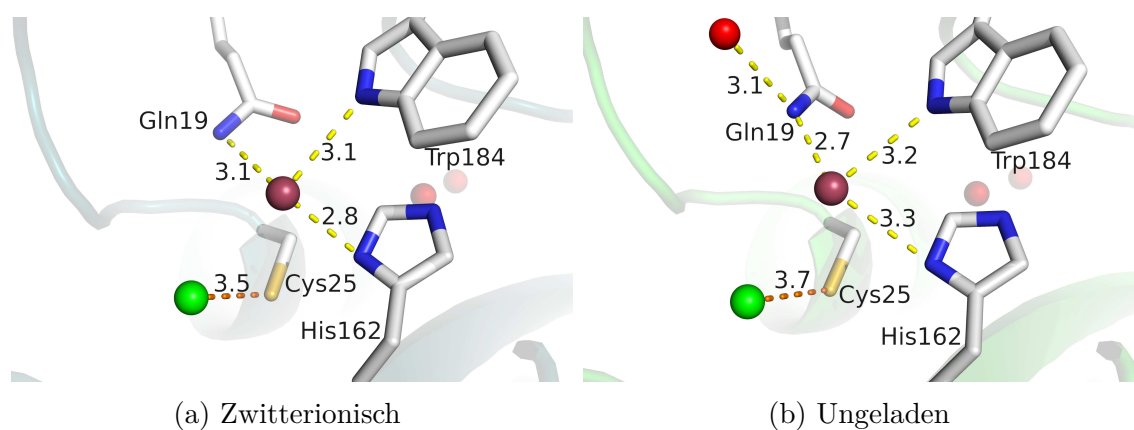


Abbildung 4.20.: Gitteranalyse der 1MEM-Simulationen. Der Cluster des elektrophilen Kohlenstoffatoms ist grün, der Cluster des Sulfonyl-Sauerstoffatoms dunkelrot dargestellt.

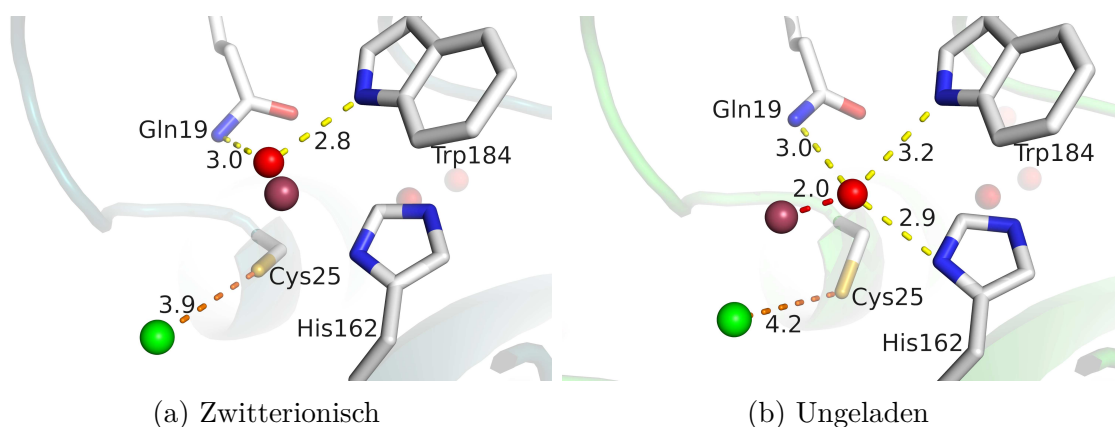


Abbildung 4.21.: Gitteranalyse der 1BGO-Simulationen. Der Cluster des elektrophilen Kohlenstoffatoms ist grün, der Cluster des Sulfonyl-Sauerstoffatoms dunkelrot dargestellt.

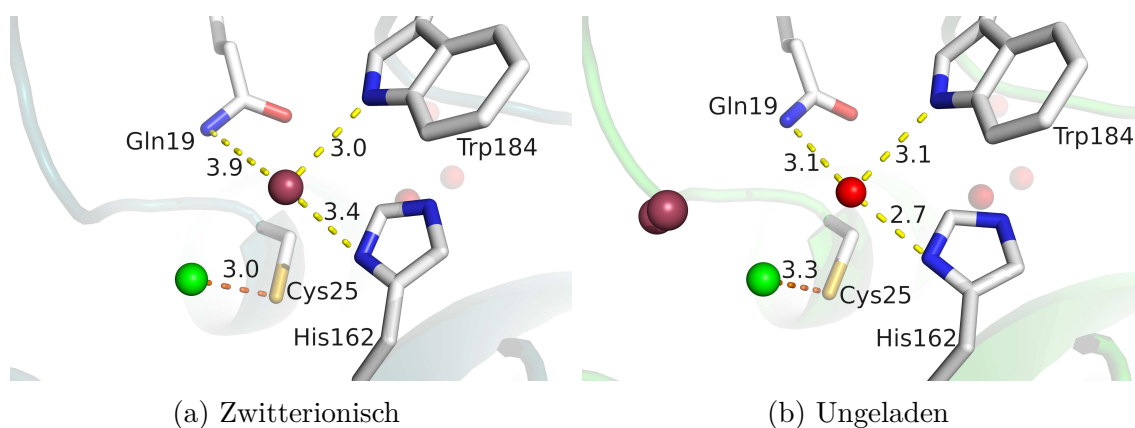


Abbildung 4.22.: Gitteranalyse der 1NLJ-Simulationen. Der Cluster des elektrophilen Kohlenstoffatoms ist grün, der Cluster des Sulfonyl-Sauerstoffatoms dunkelrot dargestellt.

#### 4.1.4.2. Zeitlich aufgelöster Ansatz

Auch wenn der Gitter-Ansatz sehr deutlich zeigt, wo sich ein entsprechendes Netzwerk aus Wasserstoffbrückenbindungen ausbildet, geht doch jegliche zeitliche Information verloren und es ist unklar, wie lang welche Interaktion zwischen Nitril, Wasser und Protein aufrecht erhalten bleibt. Darüber hinaus ist nicht mehr nachvollziehbar, wie viele Wassermoleküle tatsächlich zu dem beobachteten Cluster beigetragen haben. Die hohe Beweglichkeit des Solvens erschwert eine derartige Auswertung, da ein einzelnes Wassermolekül festgesetzte Kriterien für Wasserstoffbrückenbindungen, wie etwa Grenzwerte in Bezug auf Distanz und Winkel, selten erfüllt.

Tabelle 4.6 illustriert diesen Umstand. Lediglich während ca. 55 % (CYM/HIP) bzw. 34 % (CYS/HIS) aller Momentaufnahmen liegt ein Wassermolekül geometrisch so vor, dass es un-



4.1. *CatK: Simulation von nichtkovalenten Modellen basierend auf kovalenten Protein-Ligand-Komplexen*

		Zwitterionisch			Ungeladen		
		Rep. 1	Rep. 2	Rep. 3	Rep. 1	Rep. 2	Rep. 3
Trp184	Frames [%]	54	56	54	32	35	34
	$\emptyset$ Lebenszeit [ps]	2.4	2.5	2.4	1.6	1.7	1.7
His162	Frames [%]	5	4	3	69	69	67
	$\emptyset$ Lebenszeit [ps]	1.4	1.5	1.2	3.5	3.5	3.3
Verbrückend	Frames [%]	0	0	0	18	20	19
	$\emptyset$ Lebenszeit [ps]	0	0	0	1.3	1.3	1.3

Tabelle 4.6.: Prozentsatz der Simulationszeit, während dem Wasserstoffbrückenbindungen zu Trp184 sowie His162 als auch verbrückend vorliegen. Die „Lebenszeit“ gibt an, wie lange diese Wasserstoffbrückenbindungen im Schnitt bestehen. Diese Werte wurden unter Nutzung der Standard-Kriterien von *cpptraj* ermittelt.

ter den Standard-Kriterien von *cpptraj*<sup>154</sup> als Wasserstoffbrückenpartner von Trp184 identifiziert wird, d. h., dass der Abstand zwischen den Schweratomen 3.0 Å ist und ein Winkel von mindestens 135° zwischen den Schweratomen mit dem Wasserstoffatom als Scheitel vorliegt. Dies suggeriert, dass bei den verbleibenden Konfigurationen kein Wassermolekül vorhanden ist, was jedoch ein Trugschluss wäre: Es ist genauso gut möglich, dass das Wassermolekül nur knapp die entsprechenden Grenzwerte nicht erfüllt, z. B. 3.1 Å entfernt ist oder einen Winkel von 134° bildet. Besonders kritisch ist das bei verbrückenden Interaktionen, da hier zweimal zeitgleich die Kriterien erfüllt sein müssen. Ein denkbare Vorgehen wäre nun, die entsprechenden Standard-Kriterien zu überschreiben und durch lockerere Bedingungen zu ersetzen. Ein solches Vorgehen führt allerdings schnell zu Grenzwerten, die weder physikalisch noch statistisch begründet sind und nur noch vage an Wasserstoffbrückenbindungen erinnern. Daher wurde folgendes Vorgehen für die zeitaufgelöste Beschreibung der Solvation gewählt: Mit Hilfe des *hbond* - Befehls von *cpptraj*<sup>154</sup> wurden alle Wassermoleküle bestimmt, die zumindest kurzzeitig sowohl zu His162 als auch zu Trp184 einen Abstand von 3.2 Å oder kleiner aufwiesen. Es wurden keine Anforderungen an den konkreten Winkel gestellt. Für jedes so erfasste Wassermolekül wurde dann die tatsächliche Distanz zu Trp184 und His162 im Verlauf der Zeit betrachtet. Gemessen wurde zwischen dem Sauerstoffatom des Wassermoleküls, dem  $\epsilon$ -Stickstoffatom des Trp184 und dem  $\delta$ -Stickstoffatom des His162.

Abbildung 4.23 zeigt das Ergebnis dieser Analyse für die zwitterionischen Simulationen des 4DMX-Komplexes. Es wird klar, dass der in Abbildung 4.16a dargestellte Cluster durch viele unterschiedliche Wassermoleküle zustande kommt, die sich gegenseitig immer wieder verdrängen. Für Trp184 werden häufiger die Kriterien einer Wasserstoffbrückenbindung erfüllt (Tabelle 4.6) als für His162, was sich auch in Abbildung 4.23 widerspiegelt. Abbildung 4.16a

zeigt, dass die Orientierung von His162 nicht optimal ist, um durch sein Proton mit dem Wasser-Cluster zu interagieren. Nichtsdestoweniger gibt es keinen Zeitpunkt im Verlauf der drei Replika, an dem das an Trp184 koordinierende Wassermolekül nicht in der unmittelbaren Nähe von His162 anzufinden ist, was erneut die Prävalenz des in Abbildung 4.16a dargestellten zentralen Wasser-Clusters unterstreicht.

Für die ungeladenen Simulationen sind die Ergebnisse in Abbildung 4.24 zu sehen. Es sind viele Übereinstimmungen zu 4.23 erkennbar. Immer noch fluktuieren viele unterschiedliche Wassermoleküle durch den Raumbereich, der durch den Wasser-Cluster in Abbildung 4.16b beschrieben wird. Allerdings wird, im Gegensatz zu der zwitterionischen Simulation, häufiger eine Wasserstoffbrückenbindung zu dem nun deprotonierten His162 ausgebildet als zu Trp184. Der Wechsel des Protonierungszustands scheint auch einen Shift des koordinierenden Wassermoleküls zu verursachen, wie bereits in Tabelle 4.5 angedeutet. Abbildung 4.6 legt nahe, dass es im ungeladenen Zustand von His162 einfacher ist, eine gleichzeitige Wasserstoffbrückenbindung zu His162 und Trp184 aufzubauen.

#### 4.1. CatK: Simulation von nichtkovalenten Modellen basierend auf kovalenten Protein-Ligand-Komplexen

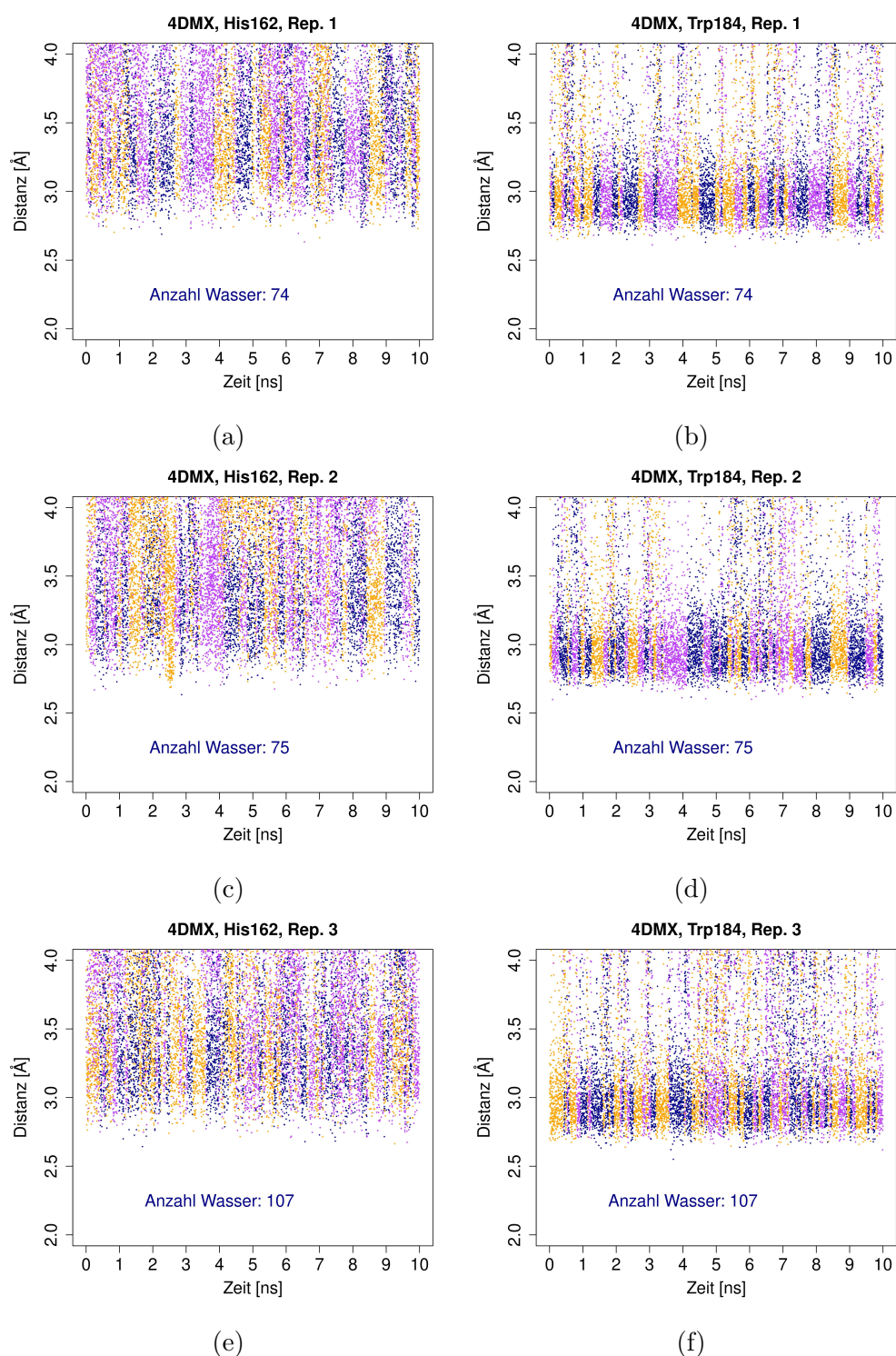


Abbildung 4.23.: Distanzen aller Wassermoleküle, die innerhalb der zwitterionischen 4DMX-Simulationen zu His162 und Trp184 in einem Abstand von 3.2 Å beobachtet werden konnten. Die Distanzen zwischen den Wassermolekülen wurden abwechselnd dunkelblau, orange und lila gezeichnet, um den Wechsel zwischen den Wassermolekülen zu verdeutlichen.

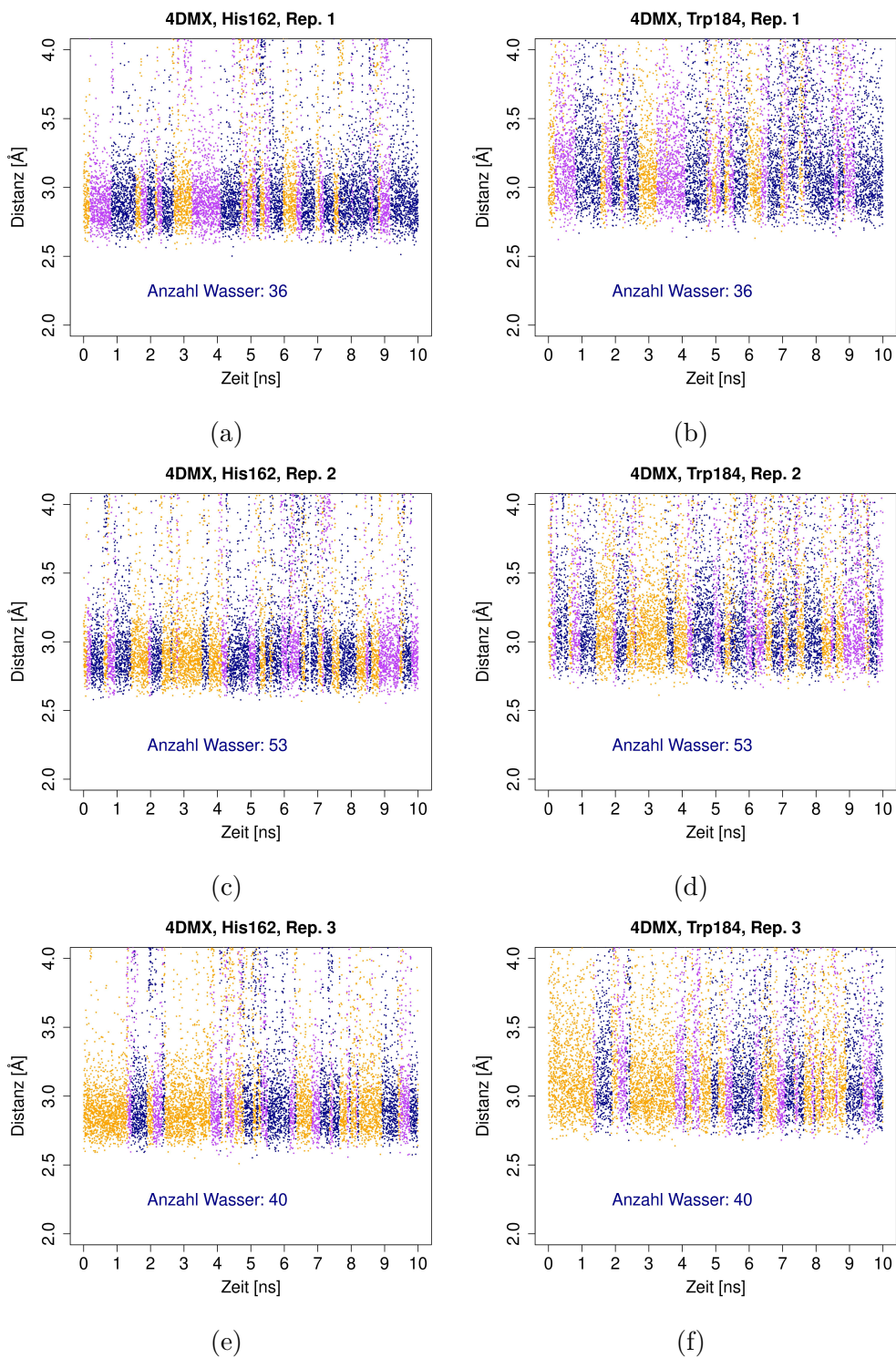


Abbildung 4.24.: Distanzen aller Wassermoleküle, die innerhalb der ungeladenen 4DMX-Simulationen zu His162 und Trp184 in einem Abstand von  $3.2 \text{ \AA}$  beobachtet werden konnten. Die Distanzen zwischen den Wassermolekülen wurden abwechselnd dunkelblau, orange und lila gezeichnet, um den Wechsel zwischen den Wassermolekülen zu verdeutlichen.

### 4.1.5. Verlängerte Simulationen

Da mit 3 x 10 ns Simulation nur ein begrenztes Sampling erreicht werden kann, wurden zusätzlich für einige Komplexe verlängerte Simulationen durchgeführt (3 x 500 ns für jeden Protonierungszustand), um zu testen, ob sich qualitativ signifikant andere Ergebnisse dadurch ergeben. Ausgewählt für einen ersten Test wurde hierfür 5TDI, da dieser Komplex den Liganden Odanacatib (**C16**) enthielt, der Phase III der klinischen Tests erreichte.

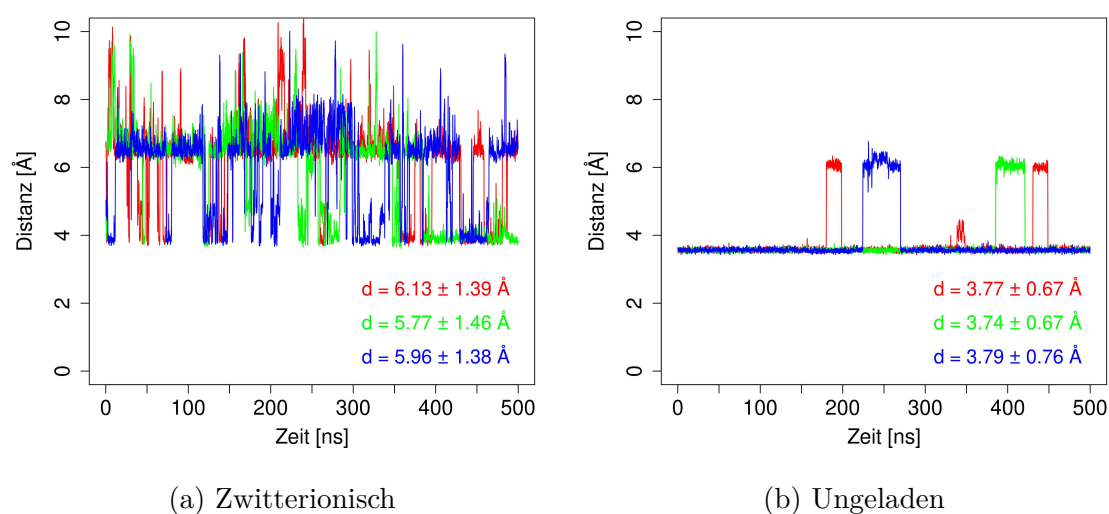


Abbildung 4.25.: Distanz zwischen den reaktiven Zentren für die 500 ns - Simulationen von 5TDI.

Abbildung 4.25 zeigt das Ergebnis dieses Testlaufs. Die längere Zeitskala zeigt temporäre Abweichungen von dem (scheinbar) bevorzugten Zustand. So kann z. B. in Abbildung 4.25b für bis zu ca. 50 ns der Warhead aus der Bindetasche hinausrotieren und eine weiter entfernte Position zum katalytischen Zentrum einnehmen. Anschließend wird aber die ursprüngliche Orientierung eingenommen. Qualitativ wird somit größtenteils das Ergebnis widerspiegelt, das bereits in den kurzen Simulationen beobachtbar ist: der zwitterionische Zustand zeigt erhöhte Abstände im Vergleich zum ungeladenen Zustand. Dieses Ergebnis wurde durch zusätzlich längere Simulationen für 1ATK, 1MEM, 4X6J und 6QBS bestätigt.

Abbildung 4.26 fasst die Ergebnisse von 1ATK mit dem Liganden (**C1**) zusammen. In den verlängerten zwitterionischen Simulationen (Abbildung 4.26a) ist deutlich die Dissoziation des Warheads erkennbar, da nun die Distanz zwischen dem Cys25 und dem elektrophilen Kohlenstoff mehr als 19 Å beträgt. Im Gegensatz dazu entfernt sich der Warhead in den neutralen Simulationen (Abbildung 4.26b) nur kurzzeitig aus der katalytischen Bindetasche. Ansätze eines möglichen Wegdriftens des Warheads in zwitterionischen 1ATK-Simulationen

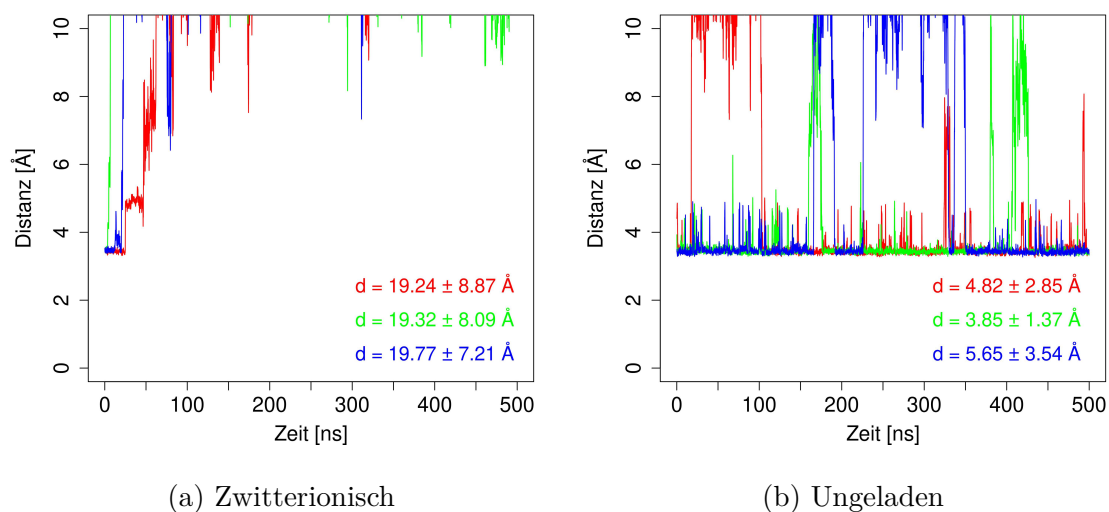


Abbildung 4.26.: Distanz zwischen den reaktiven Zentren für die 500 ns - Simulationen von 1ATK.

konnten bereits am Ende der kurzen Replikas erahnt werden (vergleiche Abbildung 4.3a) und konnten nun über die verlängerten Simulationen bestätigt werden.

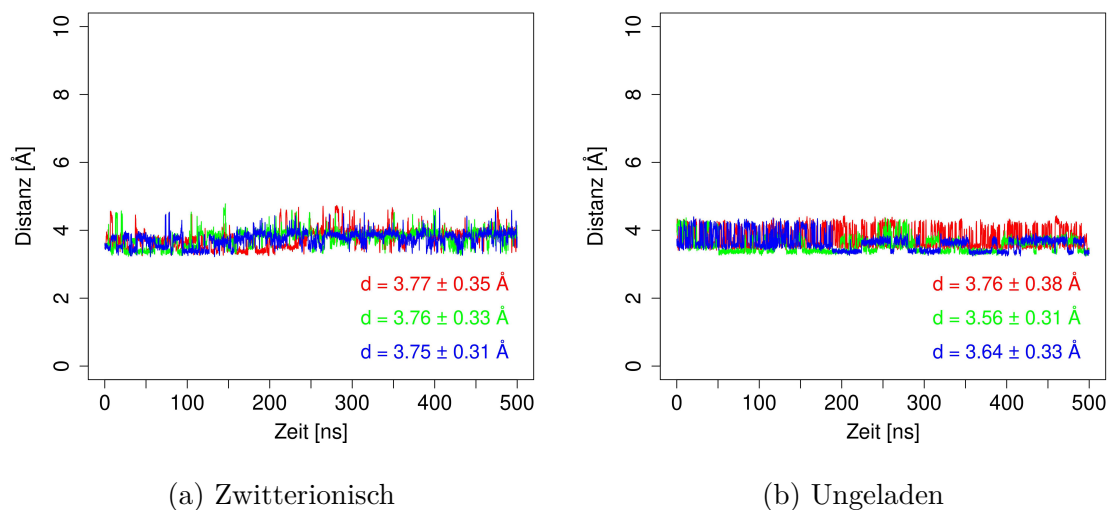


Abbildung 4.27.: Distanz zwischen den reaktiven Zentren für die 500 ns - Simulationen von 1MEM.

1MEM bzw. der Ligand **C2** zeigt, ebenfalls wie in den kurzen Simulationen, in beiden Protonierungszuständen kurze reaktive Abstände (Abbildung 4.27). Erneut ließe sich allein anhand der Simulationen nicht klar beurteilen, welcher Protonierungszustand für die Koordination

des Vinylsulfons günstiger wäre, da beide Zustände kurze Abstände zum reaktiven Zentrum erlauben.

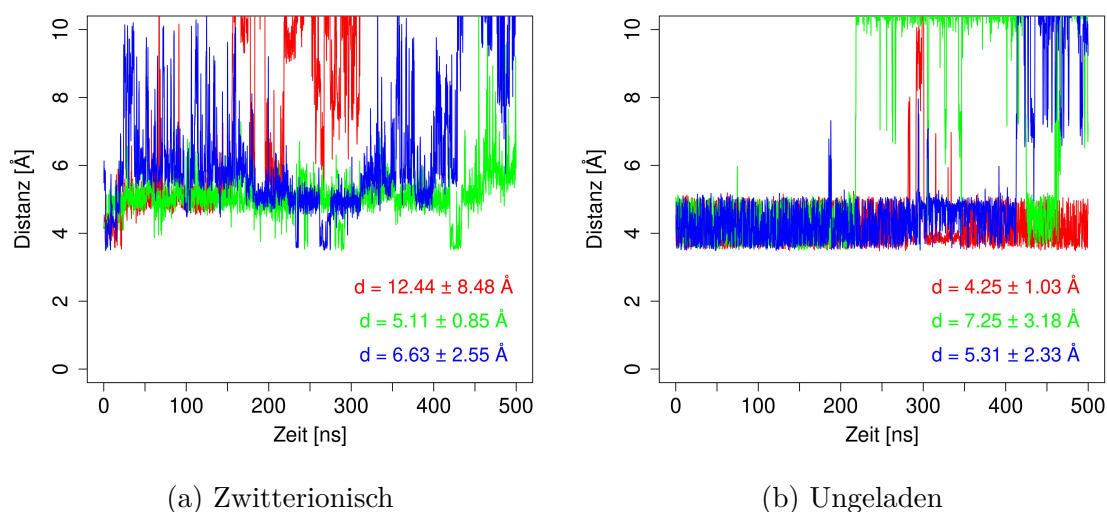


Abbildung 4.28.: Distanz zwischen den reaktiven Zentren für die 500 ns - Simulationen von 4X6J.

Abbildung 4.28 zeigt die verlängerte Simulation für 4X6J und dessen Ligand **C15**. In der zwitterionischen Simulation (Abbildung 4.28a) liegt über mehrere hundert Nanosekunden hinweg der Warhead ca. 5 Å vom katalytischen Cystein entfernt vor; in einem Replikon scheint der Ligand sogar noch weiter zu dissoziieren, sodass Abstände über 10 Å vorliegen. Dieses Verhalten konnte in den 3 x 10 ns langen Simulationen nicht beobachtet werden. In der ungeladenen Simulation (Abbildung 4.28b) kann über mehrere hundert Nanosekunden hinweg das oszillierende Verhalten des Warheads beobachtet werden, das auch schon in Abbildung 4.7f erkennbar war. Auch hier kann über den größeren Zeitraum der Beginn der Dissoziation des nichtkovalenten Liganden beobachtet werden, anders als in Abbildung 4.28a ist dieses Verhalten allerdings zeitlich begrenzt. Die verlängerte 4X6J-Simulation kann im ungeladenen Zuständen über mehrere hundert Nanosekunden hinweg das Verhalten innerhalb der kurzen Simulation reproduzieren, während in der zwitterionischen Simulation noch größere Abstände beobachtbar sind, weil der Ligand mit der Dissoziation beginnt.

6QBS (**C20**) hingegen zeigt im Gegensatz dazu in seinen verlängerten Simulationen (Abbildung 4.29) dasselbe Verhalten wie in den kürzeren Simulationen (Abbildung 4.10). Der Warhead alterniert in der ungeladenen Simulation in allen drei Replikaten für 500 ns lang zwischen zwei Zuständen (Abbildung 4.29b). Der kürzere Abstand, bei dem das Alkin auf Cys25 deutet, ist dabei erneut wesentlich seltener vertreten als der größere Abstand, bei

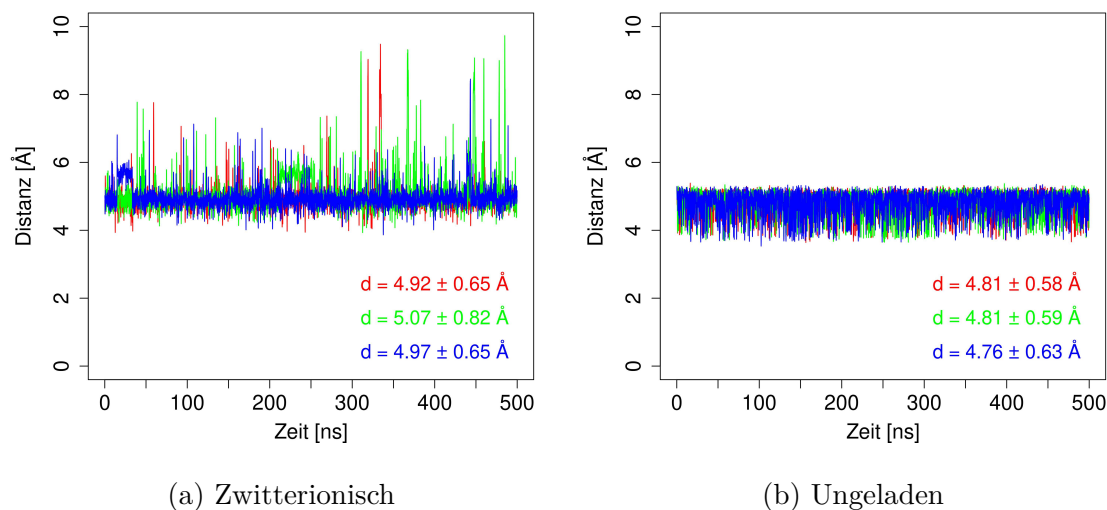


Abbildung 4.29.: Distanz zwischen den reaktiven Zentren für die 500 ns - Simulationen von 6QBS.

dem der Warhead vom katalytischen Zentrum wegdeutet. In der zwitterionischen Simulation (Abbildung 4.29a) hingegen ist der kurze Abstand, bei dem eine Reaktion möglich wäre, überhaupt nicht beobachtbar.

## 4.1.6. Vergleich zu anderen Cysteinproteasen

### 4.1.6.1. CatS

CatK gehört als papainähnliche Cysteinprotease zum Clan CA, dem größten Clan der Cysteinproteasen.<sup>156</sup> In diesem Clan sind vor allem die Aminosäuren innerhalb der katalytischen Bindetasche hoch konserviert. Abbildung 4.30 zeigt eine Überlagerung verschiedener Clanmitglieder, um diese Ähnlichkeit zu untermauern.

Es liegt daher nahe, davon auszugehen, dass die bisher gemachten Beobachtungen für CatK sich auf andere Cysteinproteasen übertragen lassen. Um diesen Punkt zu unterstreichen, wurde ein vergleichbarer Satz an Simulationen für CatS angefertigt, der in der Größe vergleichbar zum CatK-Datensatz ist (siehe Tabelle 4.7). Alle Liganden der CatS-Studie sind in Abbildung 4.31 gezeigt.

Abbildung 4.32 zeigt eine Zusammenfassung der reaktiven Abstände im Laufe der Simulationen. Qualitativ sind die Ergebnisse vergleichbar zu Abbildung 4.13: Es werden niedrigere Mediane und kleinere Interquartilabstände für die Nitrile und Aldehyde beobachtet, wenn von einer ungeladenen katalytischen Einheit ausgegangen wird. Das Vinylsulfon in 1NPZ



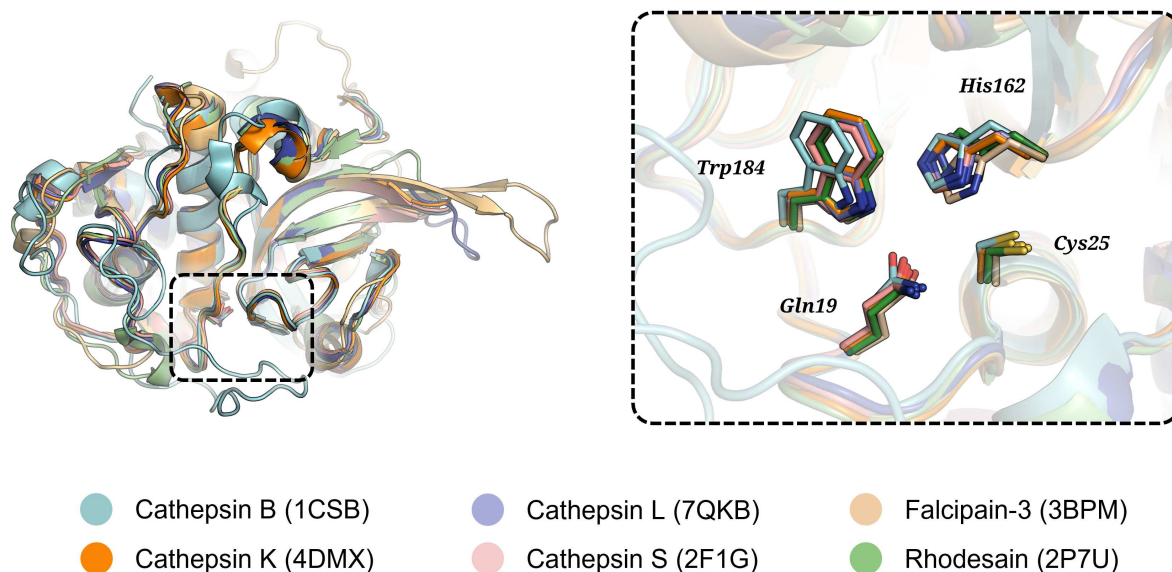


Abbildung 4.30.: Überlagerung verschiedener papainähnlicher Cysteinproteasen. Die Nummerierung von CatK wurde hierbei übernommen. Die Abbildung wurde freundlicherweise von Natalia Chen zur Verfügung gestellt, bereits an anderer Stelle publiziert und mit Genehmigung hier reproduziert.<sup>134</sup> Copyright 2023 American Chemical Society.

PDB	Ligand	Warhead	Publi- kation	Auflösung [Å]	C-S-Bindungs- länge [Å]	Lipinski- verstöße
1NPZ <sup>157</sup>	<b>C21</b>	Vinylsulfon	2003	2.00	1.8	1
1NQC <sup>157</sup>	<b>C22</b>	Aldehyd	2003	1.80	1.7	0
3OVX <sup>158</sup>	<b>C23</b>	Aldehyd	2010	1.49	1.8	0
1MS6 <sup>159</sup>	<b>C24</b>	Nitril	2003	1.90	1.9	0
2FQ9 <sup>160</sup>	<b>C25</b>	Nitril	2006	1.65	1.8	0
2FRA <sup>161</sup>	<b>C26</b>	Nitril	2006	1.90	1.8	0
2FRQ <sup>162</sup>	<b>C27</b>	Nitril	2006	1.60	1.8	0
2FT2 <sup>163</sup>	<b>C28</b>	Nitril	2006	1.70	1.8	0
2R9M <sup>164</sup>	<b>C29</b>	Nitril	2007	1.97	1.9	0
2R9N <sup>165</sup>	<b>C30</b>	Nitril	2007	2.00	1.9	0
2R9O <sup>166</sup>	<b>C31</b>	Nitril	2007	2.00	1.9	0
3N4C <sup>167</sup>	<b>C32</b>	Nitril	2011	1.90	1.8	0
2G7Y <sup>168</sup>	<b>C33</b>	Nitril	2006	2.00	-	0

Tabelle 4.7.: Simulierte CatS-Komplexe. 2G7Y wurde ebenfalls in die Studie aufgenommen, auch wenn hier keine kovalente Bindung modelliert war, da die 2Fo-fc-Map der Kristallstruktur durchaus eine Bindung zwischen Ligand und Protein andeutet.

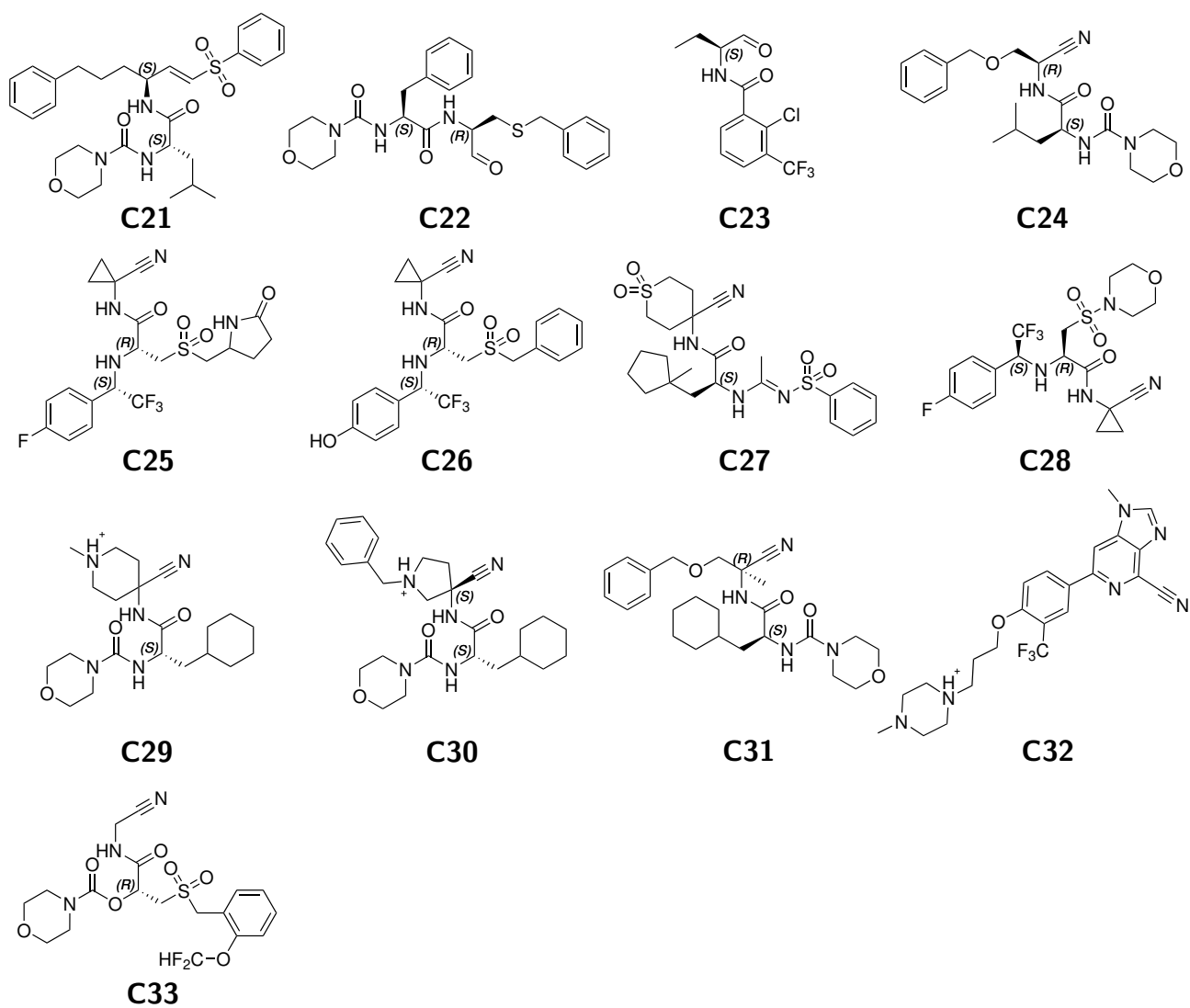
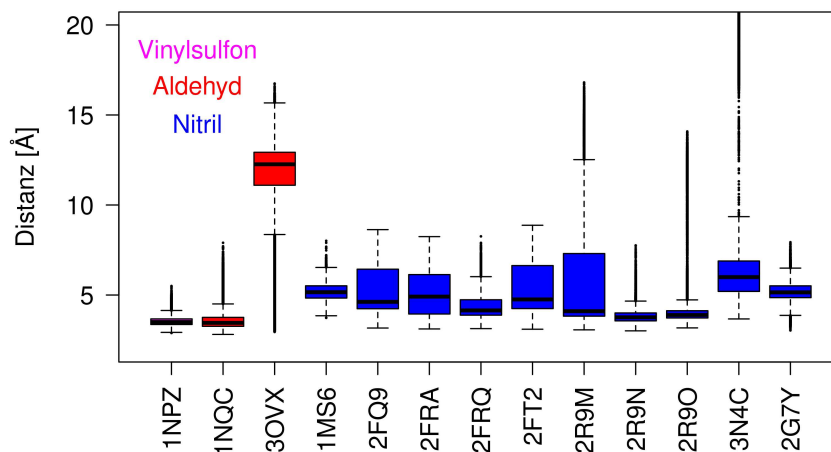
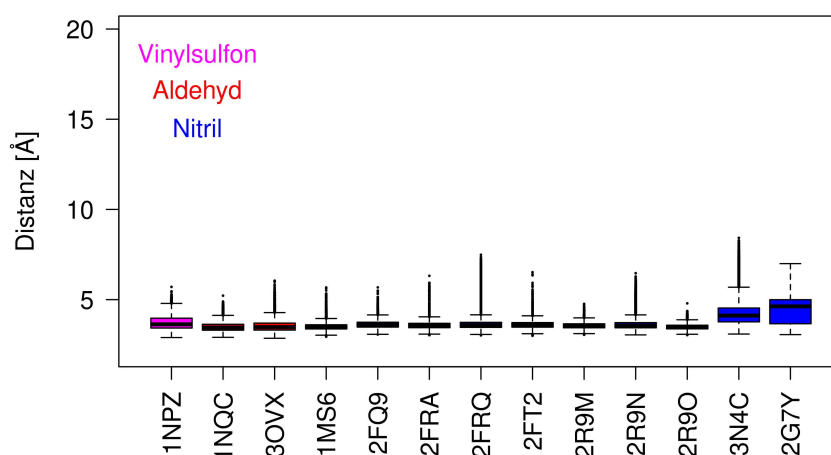


Abbildung 4.31.: Alle Liganden der CatS-Studie.

(C21) ist, unabhängig vom Protonierungszustand der katalytischen Einheit, mit seinem elektrophilen Kohlenstoffatom nahe am reaktiven Cystein.



(a) Zwitterionisch



(b) Ungeladen

Abbildung 4.32.: Boxplots der CatS-Simulationen.

Die Übereinstimmung der Ergebnisse zu CatK erstreckt sich ebenfalls auf die Resultate der gitterbasierten Wasseranalyse. Abbildung 4.33 zeigt exemplarisch das Netzwerk, das bei 1MS6 (CatS mit C24) für die zwitterionischen und für die ungeladenen Simulationen erhalten wurde. Die Ähnlichkeit zu Abbildung 4.16, der analogen Analyse für einen CatK-Fall, ist unverkennbar. In der zwitterionischen Simulation (Abbildung 4.33a) befinden sich mehrere Wasser-Cluster direkt bei Cys25 und halten den Warhead auf Abstand; die Distanz zwischen dem Nitril-Kohlenstoffatom-Cluster und dem Schwefelatom des Cysteins beträgt

5.1 Å, in guter Übereinstimmung mit dem Median in der Simulation (5.2 Å). Das Netzwerk in der ungeladenen Simulation ist wesentlich übersichtlicher, da nur der zentrale Wasser-Cluster zwischen Gln19, Cys25, His164 und Trp186 zu beobachten ist. Dadurch ist das Nitril wesentlich näher am reaktiven Cystein positioniert (3.3 Å). Diese Interaktionen sind äquivalent zu den in Abbildung 4.16a und Abbildung 4.16b dargestellten Netzwerken.

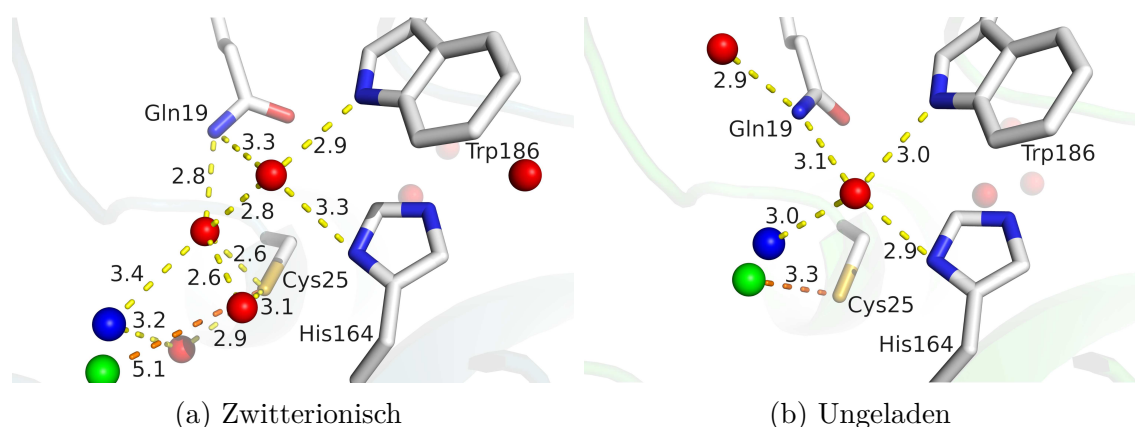


Abbildung 4.33.: Gitteranalyse der 1MS6-Simulationen. Wasser-Cluster sind rot, Nitril-Kohlenstoffatom-Cluster grün und Nitril-Stickstoffatom-Cluster blau dargestellt.

Erwähnenswert sind insbesondere die Simulationsergebnisse von 2G7Y. Technisch gesehen müsste diese Struktur aus dem Datensatz aussortiert werden, da keine kovalente Bindung von **C33** zum reaktiven Cystein in der Kristallstruktur modelliert ist und somit der Grenzwert für die C-S-Bindungslänge von  $1.8 \pm 0.1$  Å nicht erfüllt ist. Allerdings ist durchaus Elektronendichte in der 2Fo-Fc beobachtbar, die für eine derartige Bindung spräche. Aufgrund der Ähnlichkeit des Warheads zu der Reihe der 4X6-Strukturen (Kapitel 4.1.3.3) lag es nahe, von einem vergleichbaren Verhalten in MD-Simulationen auszugehen wie bei diesen Liganden, weshalb zu Vergleichszwecken die Kristallstruktur in den Datensatz aufgenommen wurde.

Abbildung 4.34 bestätigt diese Annahme. In der ungeladenen Simulation ist wieder ein alternierendes Verhalten beobachtbar, bei dem der Warhead zwischen zwei Orientierungen schwankt. Die eine entspricht der modellierten nichtkovalenten Orientierung, die andere ist analog zu einer kovalenten Bindung zum reaktiven Cystein, wie es die beobachtete Elektronendichte nahelegt. In den zwitterionischen Simulationen sind, wie bei den anderen Nitrilen, erhöhte Abstände aufgrund des koordinierenden Wassers beobachtbar.

Der für CatS simulierte Datensatz reproduziert somit die Ergebnisse der CatK-Simulationen, sogar für Spezialfälle wie die 4X6-Strukturen. Dies untermauert die These, dass aufgrund der hohen Ähnlichkeit der papainähnlichen Cysteinproteasen die in dieser Arbeit diskutierten

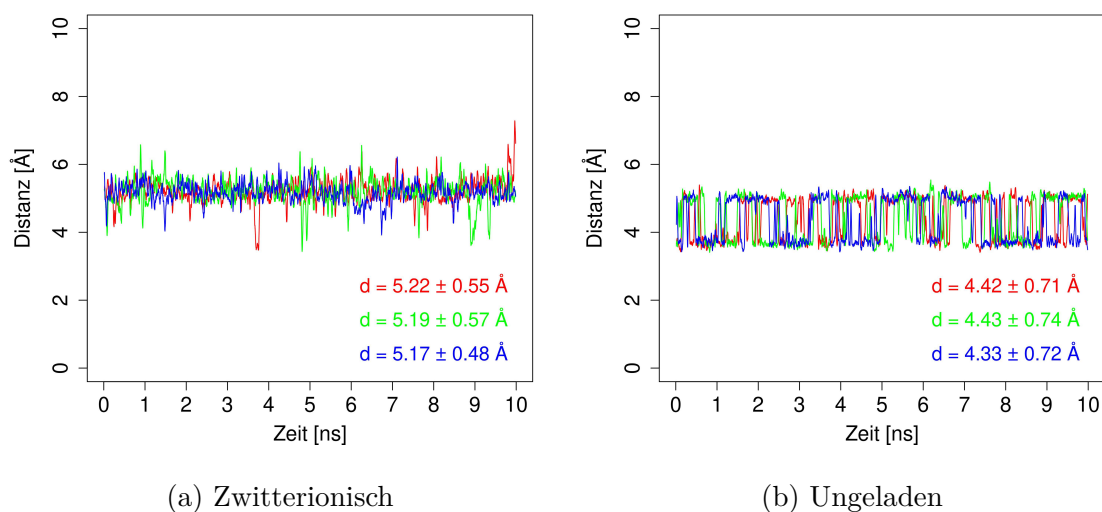


Abbildung 4.34.: Reaktiver Abstand in den 2G7Y-Simulationen.

Regelmäßigkeiten übertragbar sein könnten, was von einem Standpunkt des Wirkstoffdesigns von großem Wert wäre. So könnten, hypothetisch, Screenings für andere papainähnliche Cysteinproteasen von diesen Ergebnissen profitieren, um z. B. Pharmakophormodelle zu verbessern.

#### 4.1.6.2. SARS-CoV2-Main Protease

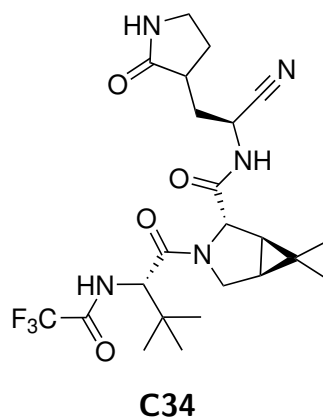


Abbildung 4.35.: Nirmatrelvir, der Ligand in 7VH8.

Um die Grenzen des hier beschriebenen Protokolls zur Erstellung eines nichtkovalenten Modells von kovalent bindenden Liganden zu testen, wurde die Kristallstruktur eines Proteins gewählt, das sich strukturell stark von den papainähnlichen Cysteinproteasen unterscheidet.

det. Hierfür wurde die Kristallstruktur von der SARS-CoV2-Main Protease im Komplex mit Nirmatrelvir (PDB-Code 7VH8, mit **C34**), aufgrund der aktuellen Relevanz, gewählt. Die Kristallstruktur wurde wie in Kapitel 4.1.2 beschrieben vorbereitet, d. h. nach Modellierung des nichtkovalenten Zustandes wurden beide Protonierungszustände für die katalytische Einheit mit drei 10 ns-Replikaten simuliert. Die Abstände im Verlauf der Simulation können in Abbildung 4.36 gesehen werden.

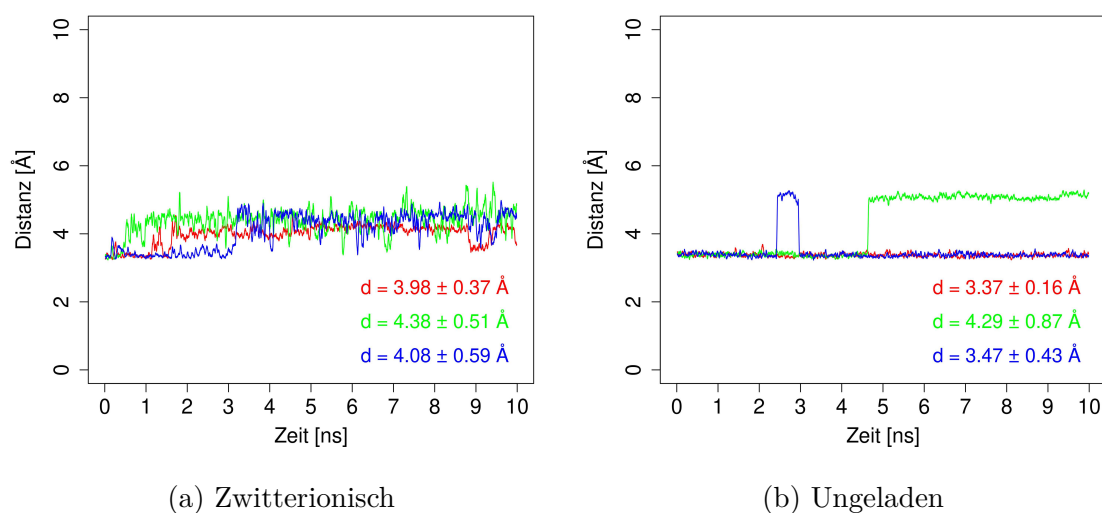
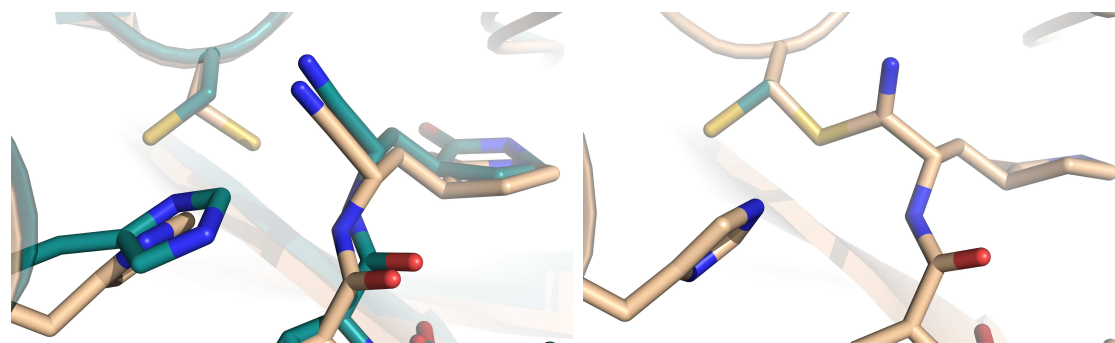


Abbildung 4.36.: Reaktiver Abstand in den 7VH8-Simulationen.

In der zwitterionischen Protonierung werden Abstände von ca. 4 Å beobachtet. In der ungeladenen Simulation sind zwei dominante Abstände beobachtbar, bei 3.4 Å und bei 5.1 Å. Abbildung 4.37a zeigt je eine Momentaufnahme für diese beiden Zustände. Diese Abbildung wurde erstellt, indem die Koordinaten aller Momentaufnahmen, deren reaktiver Abstand über oder unter 4 Å liegt, gemittelt wurden und anschließend die Aufnahme ausgewählt wurde, die zu diesen durchschnittlichen Koordinaten des Proteinrückgrats den geringsten RMSD aufwies. In Abbildung 4.37b ist zum Vergleich die Kristallstruktur gezeigt. Die aufgenommene Elektronendichte deutet auf zwei Orientierungen des Cysteins hin. 80% des Besetzungsgrades des Cysteins liegen in der kovalent gebundenen Form vor, während 20% nichtkovalent vorliegen und von dem Liganden wegdeuten. Die MD-Simulation mit ungeladener Bindetasche ist in der Lage, diese beiden Orientierungen zu reproduzieren. Vor diesem Hintergrund erscheint es durchaus denkbar, dass das hier vorgeschlagene Protokoll zur Modellierung von nichtkovalenten Strukturen durchaus breiter anwendbar ist als lediglich bei papainähnlichen Cysteinproteasen wie CatK oder CatS.



(a) Repräsentative Momentaufnahmen für die zwei Orientierungen des Cysteins innerhalb der Simulationen. (b) Die in der Kristallstruktur gezeigten Orientierungen des Cysteins. Eine von diesen ist kovalent mit dem Liganden verbunden, die andere nicht.

Abbildung 4.37.: Vergleich der zwei dominanten Zustände innerhalb der MD-Simulation mit ungeladener katalytischen Einheit mit der Kristallstruktur. Die Orientierung, die zu kleinere reaktiven Abständen führt, wurde beige eingefärbt, die für größere petrol.

## 4.1.7. Simulationen von im Docking vorhergesagter Bindemodi

### 4.1.7.1. Reproduktion der Ergebnisse der Kristallstrukturen durch Docking

Die bisherigen Untersuchungen basierten alle auf bereits veröffentlichten Kristallstrukturen mit kovalent gebundenen Inhibitoren, für die eine nichtkovalente Pose modelliert wurde. Häufig aber liegen zu Beginn eines Modelling-Projekts keine experimentellen Informationen zur Bindungspose des Liganden vor. Ein typischerweise genutztes Mittel in einer derartigen Situation ist Docking, um ein Modell eines plausiblen Bindungszustandes zu erhalten. Gerade für die Problematik des nichtkovalenten Zustandes eines kovalenten Inhibitors wäre ein Docking-Protokoll hilfreich, da für diesen Punkt des Bindungsprozesses kaum experimentelle Daten existieren. Zusätzlich legen Vergleichsstudien von kovalenten Docking-Programmen nahe, dass jene Programme eine bessere Performance erreichen, die den nichtkovalenten Zustand in ihrem Workflow berücksichtigen.<sup>23</sup>

Natalia Chen (Doktorandin des Instituts für Pharmazie an der Universität Würzburg) entwickelte freundlicherweise das folgende Docking-Protokoll und führte für die 20 bisher simulierten CatK-Komplexe das Docking aus. Ziel war es, zu überprüfen, ob die Kristallstrukturen bzw. die aus den Kristallstrukturen generierten nichtkovalenten Modelle reproduziert werden können, und ob sich die so erhaltenen Docking-Lösungen in der MD-Simulation ähnlich verhalten. Die zuverlässige Reproduktion der Kristallstruktur-Posen (bzw. derer nichtkovalenten Modelle) im Docking ist ein wichtiges Qualitätskriterium eines Docking-Protokolls, um in Abwesenheit von strukturellen Vergleichsdaten immer noch zuverlässige Modelle zu

generieren.

Das Docking wurde mit GOLD (*Genetic Optimisation for Ligand Docking*)<sup>22</sup> durchgeführt. Die grundlegende Idee des genetischen Algorithmus wird in Kapitel 3 erklärt. Drei unterschiedliche Scoring-Funktionen wurden als Fitness-Funktionen genutzt: GoldScore,<sup>22</sup> Chemscore<sup>90,169,170</sup> und ChemPLP.<sup>171</sup> Die Parameter „Populationsgröße“, „Nischengröße“ und maximale Anzahl an Operationen wurden jeweils von 100, 2 und 100 000 auf 500, 5 und 500 000 erhöht, während für die übrigen Parameter der Standard beibehalten wurde. Die Häufigkeit der genetischen Operationen wurde wie folgt gesetzt: Crossover auf 90, Mutation auf 95 und Migration auf 20. Der Suchradius betrug 10 Å um das Schwefelatom des Cys25. Pro Ligand wurden 100 Durchläufe des genetischen Algorithmus mit jedem Protonierungszustand der katalytischen Einheit und jeder Fitness-Funktion durchgeführt, was in insgesamt  $100 \times 2 \times 3 = 600$  Posen pro Ligand resultierte. Alle Posen wurde mit Hilfe der wissenschaftsbasierten Scoring-Funktion DrugScoreX (DSX)<sup>99</sup> erneut bewertet.

Tabelle 4.8 zeigt, wie viele Docking-Lösungen der 600 generierten Posen tatsächlich nahe an der Kristallstruktur sind (d. h. einen RMSD von kleiner als 2 Å aufweisen) und wie viele davon innerhalb der oberen 10 % des Rankings zu finden sind. In 16 von 20 Fällen kann innerhalb der 600 Posen mindestens eine Pose identifiziert werden, die kristallnah ist, allerdings sortiert das Rescoring mit DSX nur in 10 von 20 Fällen eine solche Pose in die Top 10 % der Docking-Lösungen. Darüber hinaus wird nur in 7 Fällen tatsächlich eine kristallnahe Pose als Top-Pose bewertet.

Diese Rescoring-Methode wäre also, in Abwesenheit von Informationen über die Kristallstruktur, unzureichend, um zuverlässig die korrekten Posen zu selektieren. Ein mögliches Vorgehen an dieser Stelle wäre daher gewesen, die Docking-Methode anzupassen oder das Rescoring zu verbessern, z. B. indem andere Docking-Programme getestet werden oder indem nach einem geeigneten Consensus-Scoring-Ansatz gesucht wird. Stattdessen wurde für die Definition eines zusätzlichen Auswahlkriteriums der Umstand genutzt, dass eine kovalente Reaktion möglich sein muss. Daher wurden nur solche Posen betrachtet, die einen Abstand von 3.7 Å oder weniger zwischen dem Schwefelatom des Cys25 und dem elektrophilen Kohlenstoffatom aufwiesen. Dieses Vorgehen wird auch bei anderen Docking-Protokollen wie z. B. CovDock<sup>24</sup> genutzt. Daten zu den ausgewählten Posen sind in Tabelle 4.9 zusammengefasst. Diese Auswahlmethode führt in 16 Fällen zur Auswahl einer Pose mit einem RMSD < 2.5 Å im Vergleich zur Kristallstruktur, in 15 Fällen sogar < 1.7 Å. In drei Fällen (1U9V, 3KWB und 3KWZ) weicht die ausgewählte Pose mit einem RMSD > 8 Å deutlich von der Kristallstruktur ab. Bei diesen drei Posen liegt der Ligand, im Vergleich zur Kristallstruktur-Pose, um 180° um die Achse des Nitril gedreht vor, sodass trotz des hohen RMSD-Werts noch



4.1. *CatK: Simulation von nichtkovalenten Modellen basierend auf kovalenten Protein-Ligand-Komplexen*

Komplex	Ligand	Posen mit RMSD < 2 Å	In den DSX-Top 10%	Bester DSX-Rang mit RMSD < 2 Å
1ATK	<b>C1</b>	34	0	161
1MEM	<b>C2</b>	112	0	116
1U9V	<b>C3</b>	0	0	-
1U9W	<b>C4</b>	26	0	199
1U9X	<b>C5</b>	3	0	557
2R6N	<b>C6</b>	71	47	1
3KW9	<b>C7</b>	0	0	-
3KWB	<b>C8</b>	0	0	-
3KWZ	<b>C9</b>	0	0	-
3O1G	<b>C10</b>	5	0	90
4DMX	<b>C11</b>	319	14	17
4DMY	<b>C12</b>	155	53	1
4X6H	<b>C13</b>	488	60	1
4X6I	<b>C14</b>	367	3	54
4X6J	<b>C15</b>	271	0	199
5TDI	<b>C16</b>	11	8	1
1BGO	<b>C17</b>	161	32	1
1NLJ	<b>C18</b>	384	49	1
1SNK	<b>C19</b>	91	8	12
6QBS	<b>C20</b>	66	53	1

Tabelle 4.8.: Auswertung des DSX-Rescorings. RMSD-Werte sind im Vergleich zur Kristallstruktur. Insgesamt wurden pro Ligand 600 Posen generiert und anschließend mit Hilfe von der Scoring-Funktion DSX bewertet.

niedrige reaktive Abstände erhalten werden.

In einem Fall (3KW9) kann keine Pose gefunden werden, die den Warhead nahe genug an Cys25 platziert. Es sollte betont werden, dass bei diesen vier Fällen (1U9V, 3KWB, 3KWZ und 3KW9) keine Pose unter den 600 generierten Docking-Lösung existiert, die einen RMSD < 2 Å im Vergleich zur Kristallstruktur aufzeigt (vergleiche Tabelle 4.8). Vor diesem Hintergrund wählt die gezeigte Auswahlmethode für den hier diskutierten Datensatz nahezu ideal kristallnahe Posen aus, sofern eine solche im Rahmen der Docking-Prozedur generiert wurde.

Die so ermittelten Strukturen wurden anschließend nach demselben Protokoll simuliert wie die aus der Kristallstruktur generierten nichtkovalenten Modelle. Es sollte betont werden, dass unabhängig von dem Protonierungszustand, der für die ausgewählte Pose genutzt wurde, sowohl CYM/HIP- als auch CYS/HIS-Simulationen durchgeführt wurden, um bei jedem Schritt des Protokolls agnostisch gegenüber jedem möglichen Protonierungszustand vorzu-

Komplex	Ligand	RMSD [ $\text{\AA}$ ]	DSX-Ranking
1ATK	<b>C1</b>	1.03	161
1MEM	<b>C2</b>	1.36	116
1U9V	<b>C3</b>	8.43	248
1U9W	<b>C4</b>	1.38	199
1U9X	<b>C5</b>	2.46	357
2R6N	<b>C6</b>	0.95	1
3KW9	<b>C7</b>	-	-
3KWB	<b>C8</b>	8.82	108
3KWZ	<b>C9</b>	8.07	232
3O1G	<b>C10</b>	1.56	90
4DMX	<b>C11</b>	0.74	25
4DMY	<b>C12</b>	0.75	3
4X6H	<b>C13</b>	1.64	6
4X6I	<b>C14</b>	1.54	113
4X6J	<b>C15</b>	0.92	199
5TDI	<b>C16</b>	0.87	1
1BGO	<b>C17</b>	1.62	12
1NLJ	<b>C18</b>	0.69	1
1SNK	<b>C19</b>	0.90	31
6QBS	<b>C20</b>	0.88	2

Tabelle 4.9.: Ausgewählte Posen für die MD-Simulation. Es wurde die beste Pose nach DSX-Score ausgewählt, die einen reaktiven Abstand von 3.7  $\text{\AA}$  oder kleiner aufweisen. RMSD-Werte sind im Vergleich zur entsprechenden Kristallstruktur-Pose des Liganden basierend auf den Schweratomen berechnet wurden.

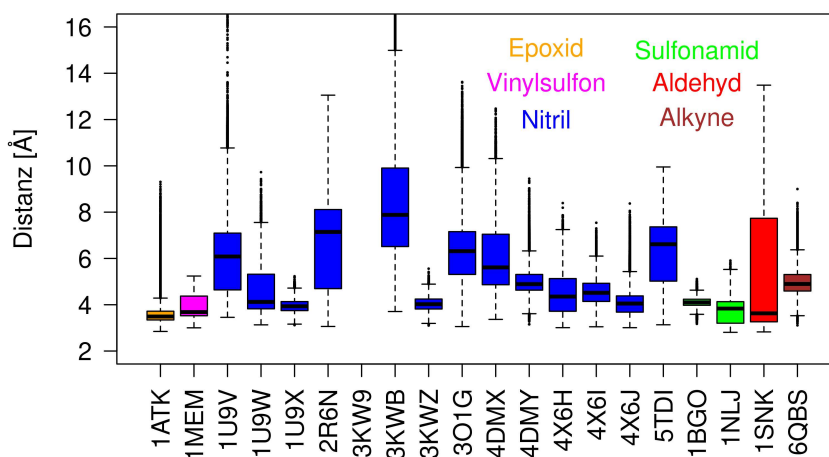
gehen.

Die Ergebnisse dieser Simulationen sind in Abbildung 4.38 gezeigt. Für die 16 Komplexe, bei denen durch Docking kristallnahe Posen modelliert wurden, können ähnliche Ergebnisse beobachtet werden wie in Abbildung 4.13. Erneut weisen die ungeladenen Simulationen geringere Interquartilabstände für den reaktiven Abstand auf als die vergleichbaren zwitterionischen Simulationen.

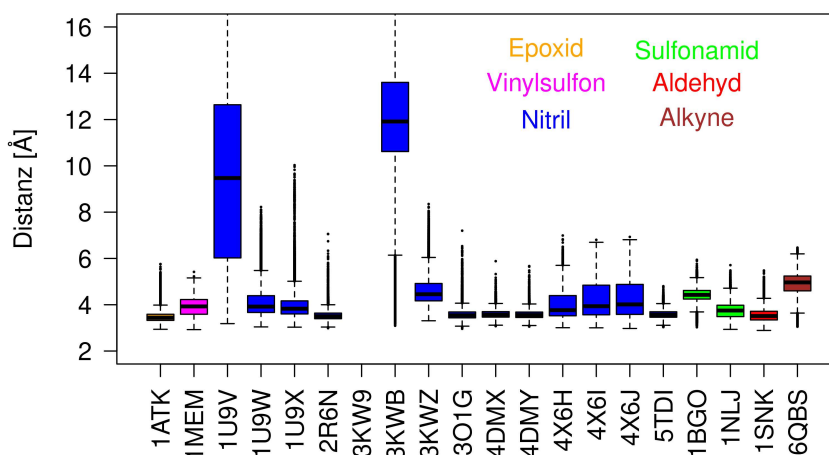
Deutlichere Abweichungen sind allerdings für 1U9V und 3KWB beobachtbar. Sowohl in den zwitterionischen als auch in den ungeladenen Simulationen ist für mehrere Replika eine Dissoziation des Liganden aus der Bindetasche beobachtbar. Vor dem Hintergrund, dass für diese beiden Komplexe keine kristallnahe Pose gefunden wurde, ist dieser Verlauf der Trajektorien durchaus nachvollziehbar. 3KWZ, dessen Docking-Lösung ebenfalls einen RMSD  $> 8 \text{\AA}$  aufwies, liegt allerdings konstant innerhalb der Bindetasche in allen Simulationen vor. Es ist denkbar, dass im Verlauf eines ausgiebigeren Samplings es ebenfalls zur Dissoziation

#### 4.1. CatK: Simulation von nichtkovalenten Modellen basierend auf kovalenten Protein-Ligand-Komplexen

kommen könnte, auch wenn sie im Rahmen dieser Trajektorien nicht beobachtet wurde. Im Wesentlichen kann aber zusammengefasst werden, dass die Simulationen basierend auf Docking-Posen die Ergebnisse der Kristallstruktur-Simulationen reproduzieren.



(a) Zwitterionisch



(b) Ungeladen

Abbildung 4.38.: Reaktiver Abstand der Simulation basierend auf den ausgewählten Docking-Lösungen.

#### 4.1.7.2. Simulation von Hits eines virtuellen Screenings

Der nächste logische Schritt aus den in Kapitel 4.1.3 und Kapitel 4.1.7.1 gewonnenen Erkenntnissen wäre, diese prospektiv anzuwenden im Rahmen eines Screenings. Kann das hier gefundene Kriterium eines im nichtkovalenten Zustand nahe an Cys25 platzierten Warheads

dabei helfen, falsch-positive Vorhersagen zu vermeiden?

Um diese Fragestellung zufriedenstellend beantworten zu können, wird ein großer Datensatz aus experimentell bestätigten *nicht* bindenden Liganden benötigt. Für CatK existiert vorteilhafterweise in der Literatur bereits ein virtuelles Screening, das hier als Vergleich genutzt werden kann. Schröder *et al.*<sup>172</sup> führten mit Hilfe von GOLD kovalente Dockings durch. Dadurch wurden aus einer fokussierten Datenbank aus 1247 elektrophilen Verbindungen 44 für IC<sub>50</sub>-Messungen ausgewählt. 39 dieser Substanzen wurden als Binder prognostiziert, während die übrigen fünf Verbindungen laut der Auswahlprozedur Nicht-Binder oder schwache Binder sein sollten. Allerdings bestätigte der Assay nur in 21 Fällen eine Inhibition, wobei einer dieser Hits aus den fünf negativ vorhergesagten Verbindungen stammt.

Allerdings können einige Kritikpunkte gegen diese Studie vorgebracht werden. Beispielsweise wurde die für viele kovalente Liganden typische Zeitabhängigkeit der Inhibition nicht genauer untersucht und in allen Assays lediglich mit einer Inkubationszeit von 30 min gearbeitet, sodass möglicherweise langsam bindende kovalente Liganden aussortiert wurden. Darüber hinaus erfolgte außer einem Warhead-basierten Filter keine Auswahl nach typischen, für ein Screening wichtigen Parametern wie z. B. Löslichkeit der Moleküle. Von der theoretischen Seite her ist fraglich, ob das Docking-Protokoll für kovalente Liganden, so wie es GOLD nutzt, sinnvoll für ein virtuelles Screening ist, da die kovalente Bindung hier *a priori* definiert wird, ohne zu testen, ob der Warhead im günstigsten nichtkovalenten Zustand korrekt orientiert ist. Falls letzteres nicht gegeben ist, wäre es denkbar, dass die Verbindung fälschlicherweise als kovalent bindende Substanz vorhergesagt wird, während sie tatsächlich nur nichtkovalent binden kann. An dieser Stelle könnten zusätzliche Analysen, wie sie im Rahmen dieser Doktorarbeit durchgeführt wurden, helfen, bessere Vorhersagen zu erzielen.

Daher wurde als *Proof-of-Concept* getestet, ob für ein Subset dieser 44 Verbindungen tatsächlich eine Verbesserung der Vorhersage erreicht werden kann, wenn der nichtkovalente Zustand aktiv berücksichtigt wird. Hierfür wurden erneut freundlicherweise von Natalia Chen Dockings im nichtkovalenten Modus von GOLD durchgeführt. Es wurde dasselbe Protokoll wie in Kapitel 4.1.7.1 genutzt. Es wurden bei dieser Auswertung nur die Nitrile innerhalb des Datensatzes betrachtet, die experimentell als inaktiv eingestuft wurden (d. h. einen IC<sub>50</sub>-Wert von über 30  $\mu$ M aufwiesen). Prinzipiell beinhaltet der Datensatz noch andere Warhead-Typen wie z. B. Thiosemicarbazone oder Vinylsulfone, für die allerdings nur wenige oder keine Kristallstrukturen (und somit keine Simulationen basierend auf diesen) als Referenz existieren. Für die so ausgewählten 9 Nitrile (siehe Tabelle 4.10) wurden erneut 100 Docking-Posen für jede Kombination der drei Scoring-Funktionen und zwei Protonierungszustände generiert, d. h. 600 Posen pro Ligand. Die beste Pose nach DSX mit einem reaktiven

Abstand kleiner als 3.7 Å wurde in beiden Protonierungszuständen simuliert. Analog zum publizierten Vorgehen<sup>172</sup> wurden alle Stereoisomere einer Verbindung generiert und gedockt, allerdings wurde für die anschließenden Simulationen nur dasjenige ausgewählt, welches den besten DSX-Score erhielt.

Für CP247470 konnte keine nichtkovalente Pose mit einem Abstand kleiner als 3.7 Å gefunden werden, weshalb hier keine Simulation durchgeführt wurde (siehe Tabelle 4.10). Im Rahmen des hier beschriebenen Protokolls würde diese Verbindung daher als inaktiv eingestuft und folglich in einem Screening eliminiert werden. Tabelle 4.10 fasst zusammen, wie häufig innerhalb des 3 x 10 ns Samplings es möglich war, den Warhead in einer Nähe von 3.7 Å, 4.0 Å und 4.3 Å unter Nutzung des jeweiligen Protonierungszustandes zu finden war. Als Vergleichswert werden vergleichbare Werte für die Simulation von 4DMX gezeigt (ausgehend vom nichtkovalenten Modell, das aus der Kristallstruktur generiert wurde). 4DMX wurde als Referenz ausgewählt, da es sowohl für den reaktiven Abstand (siehe Abbildung 4.13b) als auch für das Wasserstoffbrückennetzwerk (siehe Abbildung 4.16b) exemplarisch für die Nitrile ist. Der kurze Abstand von 3.7 Å orientiert sich dabei an den prinzipiell möglichen Abständen der Kristallstruktursimulationen (wie sie z. B. 4DMX erreicht), während die erweiterten Abstände von 4.0 Å und 4.3 Å sicherstellen sollen, dass nicht durch einen zu strengen Grenzwert, der sich an optimierten Liganden orientiert, potenzielle Binder ausgeschlossen werden. Zusätzlich können in Abbildung 4.39 Boxplots des reaktiven Abstandes gefunden werden.

Auffällig ist, dass erneut im zwitterionischen Zustand der Warhead nur sehr selten in der Nähe des katalytischen Cysteins beobachtet werden kann. Dies spiegelt sich ebenfalls in Abbildung 4.39a wider, wo (nahezu) alle Mediane über 5 Å liegen. CP150531 stellt die Ausnahme zu diesem Trend dar, da diese Verbindung in der zwitterionischen Simulation deutlich häufiger in der Nähe von Cys25 anzutreffen ist als die Kristallstruktur-Referenz 4DMX. Zusätzlich wird in den Simulationen mit ungeladener katalytischer Einheit für CP247445, CP247451 und CP247459 keine stabile Positionierung des Warheads in der Nähe von Cys25 beobachtet (siehe Abbildung 4.39b). Da diese Verbindungen somit in keinem Protonierungszustand stabil zu sein scheinen, sollte hier nicht von einer kovalenten Bindung ausgegangen werden. Diese Verbindungen würden daher in einem Screening ebenfalls eliminiert werden.

Die verbleibenden fünf Verbindungen CP79487, CP97899, CP99499, CP150531 und CP232977 zeigen hingegen über längere Abschnitte ihrer CYS/HIS-Simulation hinweg niedrige Distanzen, wenn auch nicht vergleichbar oft zu 4DMX (siehe Tabelle 4.10). Für diese fünf Verbindungen wurde daher zusätzlich die in Kapitel 4.1.4.1 beschriebene Gitteranalyse durchgeführt, um die primär auftretende Positionierung des Warheads zu beschreiben. Das Er-

	Verbindung	Vorhersage <sup>172</sup>	% < 3.7 Å	% < 4.0 Å	% < 4.3 Å
Zwitterionisch	CP79487	aktiv	0.1	0.4	1.3
	CP97899	aktiv	0.7	4.3	14.5
	CP99499	inaktiv	0.0	0.0	0.0
	CP150531	aktiv	5.1	29.9	69.5
	CP232977	inaktiv	0.0	0.0	0.1
	CP247445	aktiv	0.0	0.0	0.0
	CP247451	inaktiv	0.0	0.0	0.0
	CP247459	aktiv	0.0	0.0	0.0
	CP247470	aktiv	Keine Simulation		
	4DMX (Kristall)	-	0.5	1.5	4.2
Ungeladen	CP79487	aktiv	32.8	63.7	75.3
	CP97899	aktiv	13.7	38.5	61.6
	CP99499	inaktiv	0.1	4.0	23.0
	CP150531	aktiv	14.7	38.9	58.1
	CP232977	inaktiv	0.7	7.0	26.6
	CP247445	aktiv	0.0	0.0	0.0
	CP247451	inaktiv	0.0	0.0	0.0
	CP247459	aktiv	0.0	0.0	0.0
	CP247470	aktiv	Keine Simulation		
	4DMX (Kristall)	-	74.7	97.5	99.6

Tabelle 4.10.: Simulationsergebnisse für die Schröder-Verbindungen.

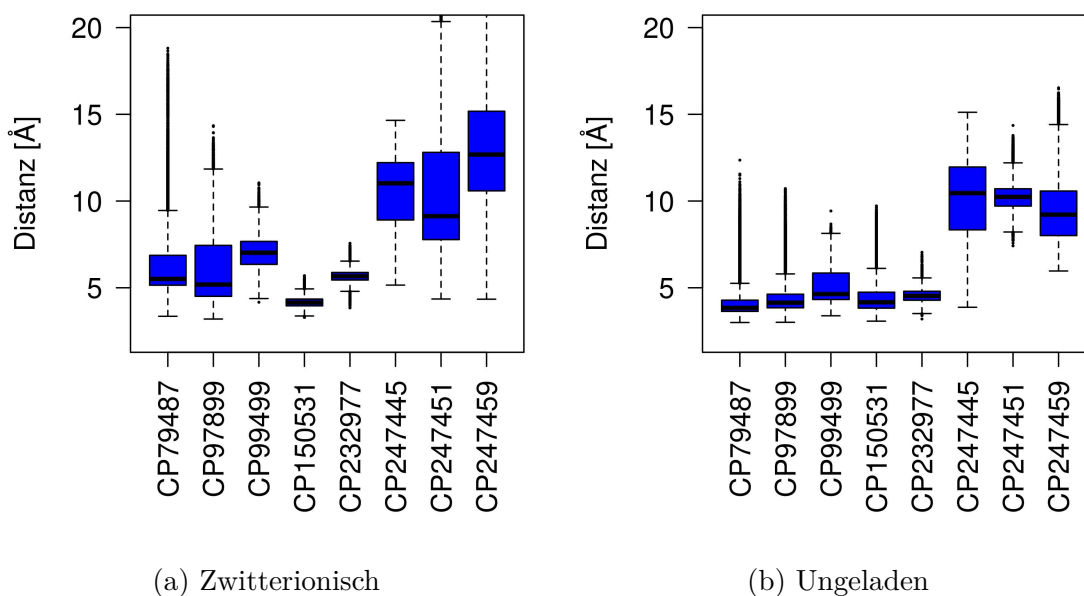


Abbildung 4.39.: Boxplots des Nitril-Subsets.

#### 4.1. *CatK*: Simulation von nichtkovalenten Modellen basierend auf kovalenten Protein-Ligand-Komplexen

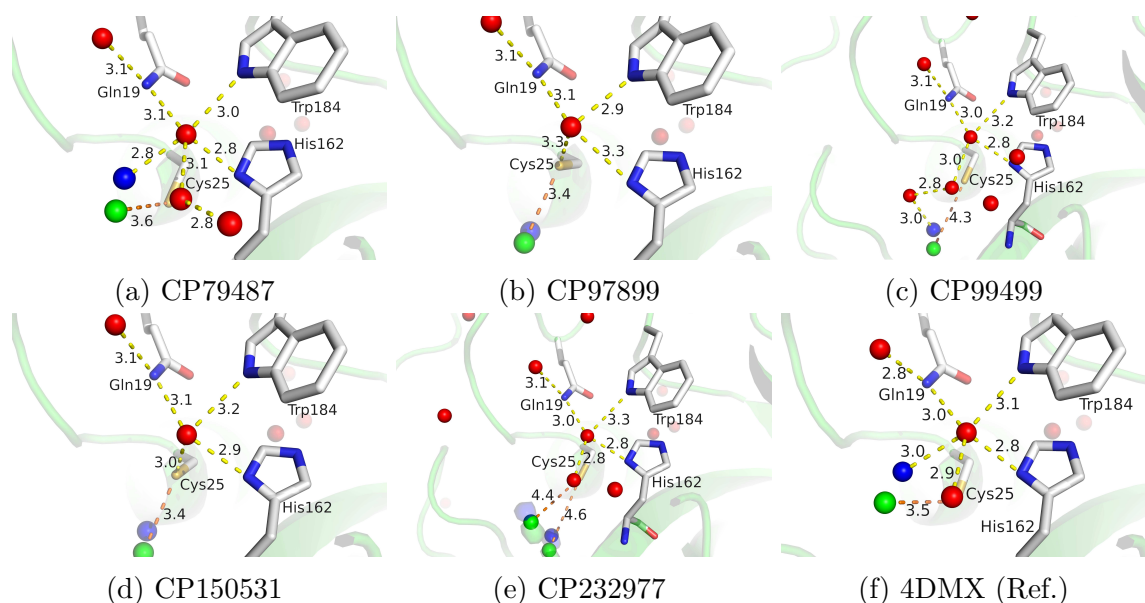


Abbildung 4.40.: Gitterbasierte Dichteanalyse des Subsets, 4DMX als Referenz. Für CP99499 und CP232977 wurden weitere Blickwinkel gewählt, um den Warhead ebenfalls abzubilden.

gebnis dieser Analyse ist in Abbildung 4.40 zu sehen. Lediglich bei einer einzigen Verbindung, CP79487, wird eine vergleichbare Positionierung des Warheads erzielt wie in der Kristallstruktur-basierten Simulation von 4DMX (vergleiche Abbildung 4.40a und Abbildung 4.40f). Bei CP97899 und CP150531 (Abbildung 4.40b und Abbildung 4.40d) ist das Nitril zwar noch vergleichbar nahe an Cys25, aber deutlich unterhalb der katalytischen Einheit platziert und nicht mehr im Kontakt mit dem zentralen Wasser-Cluster. Bei CP99499 und CP232977 ist das Nitril weiter entfernt von Cys25 als bei 4DMX (siehe Abbildung 4.40c und Abbildung 4.40e), wobei CP232977 zusätzlich zwischen zwei unterschiedlichen Positionierungen zu schwanken scheint.

Es wird davon ausgegangen, dass Gln19 zum katalytischen Mechanismus beiträgt, indem es die negative Ladung des Übergangszustandes des Substrats stabilisiert, vergleichbar zu dem Oxyanionenloch von Serinproteasen.<sup>33,34</sup> Wird davon ausgegangen, dass Nitrile für eine schnelle Reaktion von einer ähnlichen Stabilisierung profitieren müssen, sollte das Nitril so platziert sein, dass es in Richtung von Gln19 deutet. Dies wäre bei CP97899, CP99499, CP150531 und CP232977 nicht der Fall. (Auch in der zwitterionischen Simulation von CP150531 wäre diese ungünstige Orientierung gegeben.)

Mit Hilfe des hier beschriebenen Protokolls bestehend aus nichtkovalentem Docking mit einem abstands-basierten Auswahlkriterium, MD-Simulation für 3 x 10 ns und anschließender gitterbasierten Dichteanalyse hätten somit prinzipiell 8 von 9 Verbindungen als falsch-positiv

prognostiziert werden können, die im Rahmen eines bloßen kovalenten Dockings als bindende Liganden vorhergesagt wurden. Dies unterstreicht den Nutzen der Analyse des nichtkovalenten Zustandes von kovalent bindenden Molekülen. CP79487 zeigt in diesen drei Auswertungen ähnliche Ergebnisse, wie sie aus Kristallstrukturen von bestätigten Bindern gewonnen wurden, sodass es nicht als Nichtbinder identifiziert werden würde. Es wäre denkbar, dass eine andere Docking-Methode oder besseres Sampling im Rahmen der MD-Simulation hier die Richtigkeit der Vorhersage verbessern könnte. Allerdings wäre dann fraglich, ob dieser Aufwand gerechtfertigt wäre, um eine relativ geringe Verbesserung der in diesem Datensatz hohen Vorhersage-Genauigkeit zu erzielen. Um zu sehen, wie groß der Nutzen des hier beschriebenen Protokolls im Rahmen eines virtuellen Screenings wäre, wäre ein größerer Datensatz aus bestätigten Nichtbindern notwendig. Die hier durchgeführte Auswertung soll aber als *Proof-of-Concept* genügen.

#### 4.1.8. Kernaussagen der CatK-Studie

In dieser Untersuchung wurden Kristallstrukturen von kovalent inhibiertem CatK in einen nichtkovalenten Zustand überführt und anschließend klassischen MD-Simulationen unterzogen. Eine erste wichtige Erkenntnis dieser Simulationen war, dass die erhaltenen Modelle des nichtkovalenten Komplexes stabil genug für Simulationen im Nanosekunden-Bereich (vergleiche Kapitel 4.1.3) sind, zum Teil sogar für mehrere hundert Nanosekunden (vergleiche Kapitel 4.1.5). In diesen Simulationen ist, bei mindestens einem Protonierungszustand und unabhängig vom Typ des Warheads, ein Abstands-Median von ca. 3.7-4 Å zwischen dem elektrophilen Kohlenstoffatom des Warheads und dem nukleophilen Schwefelatom des Proteins beobachtbar (siehe Kapitel 4.1.3.5). Das bedeutet, dass der Warhead sich in diesen Modellen so nahe an die zu adressierende Aminosäure bewegen kann, wie es ein klassisches Kraftfeld zulässt, da dieser Abstand in etwa der Summe der Radien eines Schwefel- und eines Kohlenstoffatoms entspricht. Es sollte betont werden, dass diese Nähe anscheinend nicht permanent aufrecht erhalten werden muss, um eine kovalente Reaktion zu ermöglichen, wie die Simulationen von Kapitel 4.1.3.3 und Kapitel 4.1.3.4 nahelegen. Daher ist es vorstellbar, dass eine Art Grenzwert existiert, wie oft im Rahmen einer Simulation ein elektrophiles Atom in der Nähe seines Reaktionspartners beobachtet werden muss, sodass es realistisch wird, dass eine kovalente Bindung experimentell in einer Kristallstruktur beobachtet werden kann. Die genaue Definition dieses Grenzwertes ist aber im Rahmen der Ergebnisse dieser Arbeit nicht möglich. Die Analysen aus Kapitel 4.1.7.2 demonstrieren, dass ein Abstandskriterium als zusätzlicher Filter in virtuellen Screenings ein plausibles Vorgehen darstellen könnte, um



ungeeignete kovalente Liganden auszusortieren. Wie dieses Kriterium aber bestmöglich formuliert wird, müsste anhand eines größeren Datensatzes, der durch mehr experimentelle Ergebnisse gestützt ist, ermittelt werden.

Bemerkenswert ist darüber hinaus die Rolle des Protonierungszustandes, der das Solvationsmuster der katalytischen Einheit deutlich verändert. Hierfür wurde eine gitterbasierte Analyse durchgeführt, die basierend auf der erfassten Dichte innerhalb von Gitterpunkten und einem anschließenden Clustering Raumbereiche identifiziert, in denen Wassermoleküle, aber auch andere Atomtypen, wie z. B. Atome des Warheads, aufzufinden sind. Diese Methode erlaubte, bevorzugte Positionen von Solvensmolekülen zu erfassen, auch wenn diese aufgrund ihrer hohen Fluktuation nicht häufig genug Standard-Kriterien von Wasserstoffbrückenbindungen erfüllen. Diese Analyse zeigte, dass zwischen Gln19, His162 und Trp184 ein Wasser-Cluster vorhanden ist, der dicht an Cys25 angrenzt. Nitril- und Aldehydgruppen können diese Koordination im ungeladenen Zustand der katalytischen Einheit ergänzen, werden jedoch im zwitterionischen Zustand durch einen zusätzlichen Wasser-Cluster verdrängt und auf Abstand gehalten, sodass nur selten für eine Reaktion plausible Distanzen beobachtet werden. Häufiger kommt es in so einem Fall zu starken Fluktuationen des Nitrils oder Aldehyds. Dieses Ergebnis ließ sich qualitativ für die verwandte Cysteinprotease CatS reproduzieren (siehe Kapitel 4.1.6.1). Faszinierend dabei ist, dass unabhängig vom konkreten Liganden die bevorzugten Raumbereiche für das Nitril-Stickstoffatom und das Nitril-Kohlenstoffatom überlappen, wenn lediglich der ungeladene Zustand der katalytischen Einheit betrachtet wird (siehe Abbildung 4.41). Dieses Wissen könnte genutzt werden, um ein Pharmakophormodell zu entwickeln. Im Rahmen eines virtuellen Screenings für Inhibitoren von papainähnlichen Cysteinproteasen könnten die Nitril-Verbindungen selektiert werden, die ihren Warhead im nichtkovalenten Zustand so positionieren wie experimentell bestätigte kovalente Binder in ihren nichtkovalenten MD-Simulationen. Aus der Perspektive des Wirkstoffdesigns ist ebenfalls relevant, dass die aus der Kristallstruktur generierten nichtkovalenten Modelle sich prinzipiell auch durch Docking generieren ließen (siehe Kapitel 4.1.7.1), sofern im Rahmen dieser Prozedur eine kristallnahe Pose identifiziert wird.

Es sollte betont werden, dass die hier durchgeführten klassischen MD-Simulationen nicht dafür geeignet sind, eine definitive Aussage darüber zu treffen, welcher Protonierungszustand besser zur Beschreibung des katalytischen Zentrums von CatK geeignet ist. Auch wenn im zwitterionischen Zustand nur selten eine Reaktion möglich scheint, so könnte dies bereits ausreichend sein, da, wie zuvor erläutert, kein Cutoff bekannt ist, wie oft eine passende Koordination in der Simulation beobachtet werden muss, bis eine Reaktion experimentell erfasst werden könnte. Darüber hinaus wäre für die zwitterionische Protonierung eine niedrigere Ak-

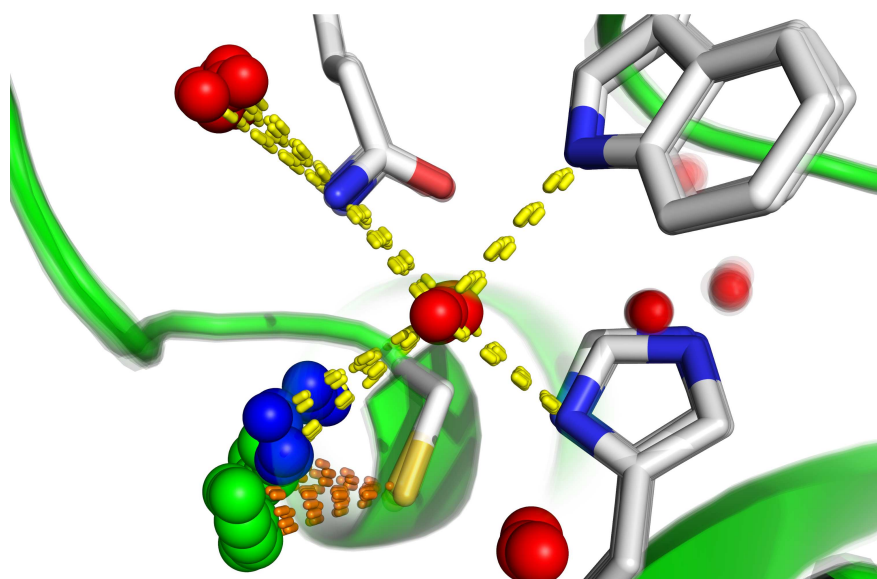


Abbildung 4.41.: Überlagerung der Gitter-Analyse jedes Nitrils (bis auf **C8** bzw. 3KWB), basierend auf den CYS/HIS-Simulationen. Wasser-Cluster sind als rote, Nitril-Kohlenstoff-Cluster als grüne und Nitril-Stickstoff-Cluster als blaue Kugeln gezeigt. **C8** bzw. 3KWB wurde dabei als einzige Simulation nicht überlagert, aus den in Kapitel 4.1.4.1 genannten Gründen. Einige vordergründige Wasser-Cluster wurden für bessere Sichtbarkeit nicht gezeigt. Die Cluster für die Kohlenstoffatome und die Stickstoffatome der Nitrile nehmen räumlich eine sehr ähnliche Position ein. Dies könnte im Rahmen eines virtuellen Screenings einen nützlichen Startpunkt für ein Pharmakophormodell darstellen.

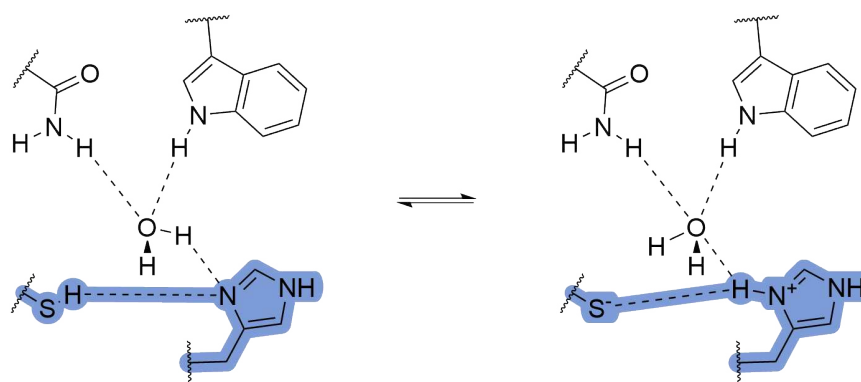
tivierungsenergie zu erwarten. Nichtsdestoweniger ergibt der ungeladene Zustand für Nitrile und Aldehyde ein konsistenteres Bild, das nicht eine Dissoziation des Liganden zur Folge hat (vergleiche Kapitel 4.1.5). In diesem Zusammenhang sind die Simulationen von Kapitel 4.1.3.3 und Kapitel 4.1.6.2 erwähnenswert. Hier war es möglich, mithilfe der Simulationen im ungeladenen Protonierungszustand die unterschiedlichen Orientierungen des Warheads oder des Cysteins zu beobachten, die auch durch die Kristallstrukturen experimentell erfasst wurden.

Diese Ergebnisse bieten eine neue Perspektive in der lang anhaltenden Debatte, welcher Protonierungszustand für die papainähnlichen Cysteinproteasen der korrekte ist. Wie bereits in Kapitel 1.2.1 dargestellt, werden basierend auf unterschiedlichen theoretischen Methoden zum Teil drastisch voneinander abweichende Ergebnisse erhalten, was darauf hindeutet, dass zumindest manche dieser Methoden systematische Fehler aufweisen.

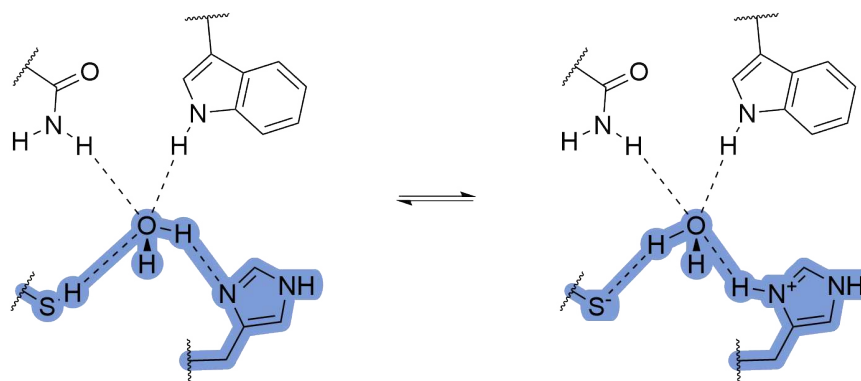
Das Übersehen des zuvor beschriebenen zentralen Wasser-Clusters könnte einen derartigen systematischen Fehler darstellen. In Methoden, die Quantenmechanik und Molekularmechanik kombinieren (QM/MM-Methoden) muss aus Gründen der Effizienz die QM-Region oft aus sehr wenigen Schweratomen zusammengesetzt werden. Es liegt daher nahe, nur Cys25 und His162 hierfür auszuwählen, sowie evtl. noch zusätzliche Atome des Liganden, falls vorhanden. Anschließend werden dann energetische Unterschiede zwischen dem zwitterionischen und dem ungeladenen Zustand berechnet, in dem das Proton direkt vom Cystein auf das Histidin übertragen wird (oder umgekehrt). Wassermoleküle, die dafür bekannt sind, derartige Protonenübertragungen zu ermöglichen und auch in quantenmechanischen Rechnungen schon dazu beigetragen haben, experimentelle Energieunterschiede zu replizieren,<sup>25, 173–175</sup> wären dann nur Teil der MM-Region und nicht Teil der QM-Region, obwohl es zumindest denkbar ist, dass die Übertragung des Protons nicht direkt, sondern über den zentralen Wasser-Cluster verläuft (siehe Abbildung 4.42). Technisch wäre ein derartiges Setup wahrscheinlich aufwendiger, da das zentrale Wassermolekül häufig durch ein anderes ersetzt wird (siehe Kapitel 4.1.4.2), d. h. es müsste wahrscheinlich eine flexible Definition der QM-Region genutzt werden, oder mehrere, nahe gelegene Wassermoleküle aufgenommen werden. Es erscheint allerdings plausibel, dass die Berücksichtigung des Wasser-Clusters innerhalb einer QM/MM-Rechnung dazu beitragen könnte, den korrekten Protonierungszustand der papainähnlichen Cysteinproteasen besser zu erfassen.<sup>175</sup>

Allerdings sollte nicht vergessen werden, dass sich Protonierungszustände und die dazugehörigen chemischen Gleichgewichte leicht unter neuen Bedingungen ändern können, zum Beispiel bei der Bindung eines Liganden.<sup>38</sup> Insbesondere kovalent bindende Liganden scheinen hierfür prädestiniert zu sein aufgrund des oft stark polaren Warheads, der in direkter

Nachbarschaft der katalytischen Einheit platziert wird. Dies zeigt sich in dieser Studie z. B. bei den Liganden mit einer Sulfonyl-Funktion, die durch die Protonierung von His162 eine zusätzliche Wasserstoffbrückenbindung ausbilden können. Dieser Umstand wirft die Frage auf, wie relevant der dominante Protonierungszustand des *apo*-Proteins ist oder ob eine Ligand-abhängige Analyse nicht prinzipiell aussagekräftiger ist für Wirkstoffdesignprojekte. Zusammenfassend zeigt die Analyse des nichtkovalenten Zustandes kovalenter Inhibitoren großes Potenzial, um das Design von kovalenten Wirkstoffen zu vereinfachen und zu verbessern.



(a) Minimales Setup der QM-Region. Hier sind keine Protonentransfere über das Wassermolekül möglich.



(b) Verbessertes Setup der QM-Region. Hier sind Protonentransfere über das Wassermolekül möglich.

Abbildung 4.42.: Mögliche Setups für eine QM/MM-Simulation zur Analyse des Protonentransfers vom Cystein zum Histidin in papainähnlichen Cysteinproteasen. Die QM-Region ist farbig unterlegt.

## 4.2. USP28: Strukturbasiertes Design eines kovalenten Inhibitors

Die selektive Adressierung von USP28 in Anwesenheit von dem strukturell ähnlichem USP25 stellt eine große Herausforderung dar. Um dies zu erreichen, erscheinen hoch affine, schnell bindende Liganden wie kovalente Inhibitoren erfolgversprechend, da diese die unterschiedliche Oligomerisierung der beiden Proteine ausnützen könnten (siehe Kapitel 1.2.3.1). Nachfolgend wird das Design eines Inhibitors für USP28 beschrieben. Grundlage für dieses Design bilden zum Zeitpunkt dieser Arbeit unpublizierte Kristallstrukturen von USP28-Konstrukten im Komplex mit AZ1 und Vismodegib (**5** und **6**, siehe 1.2.3.1 und Abbildung 4.43), die freundlicherweise von Prof. Dr. Caroline Kisker und Dr. Florian Sauer (Rudolf-Virchow-Zentrum) zur Verfügung gestellt wurden. Basierend auf einer Analyse dieser Strukturen (siehe Kapitel 4.2.1) können zwei unterschiedliche Strategien für ein Inhibitor-Design entwickelt werden. Die Entwicklung eines kovalenten Inhibitors wird im Rahmen dieser Arbeit hauptsächlich thematisiert (siehe Kapitel 4.2.4), die Möglichkeit einer Verlinkung von AZ1 und Vismodegib wird hier nur schematisch zur Ergänzung diskutiert (siehe Kapitel 4.2.5). Zusätzlich wird durch mehrere MD-Simulationen untersucht, ob sich die Dynamik von USP25 oder USP28 stark genug unterscheidet, und welche Auswirkungen Mutationen auf diese Dynamik haben (siehe Kapitel 4.2.6).

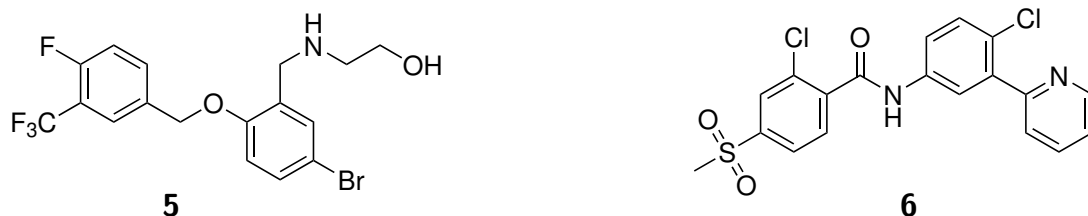


Abbildung 4.43.: AZ1 (**5**) und Vismodegib **6**.

### 4.2.1. Analyse des möglichen Vorgehens anhand von Kristallstrukturen

Die Grundlage für das strukturbasierte Design bilden frühe Refinements von unveröffentlichten Kristallstrukturen von USP28 im Komplex mit AZ1 und Vismodegib, die von Prof. Dr. Kisker und Dr. Florian Sauer zur Verfügung gestellt wurden. Zum Zeitpunkt der Zusammenfassung dieser Arbeit sind die Kristallstrukturen zwar finalisiert, aber noch nicht in der PDB hinterlegt. Bezüglich des Bindemodus der Liganden liegen allerdings keine signifikanten

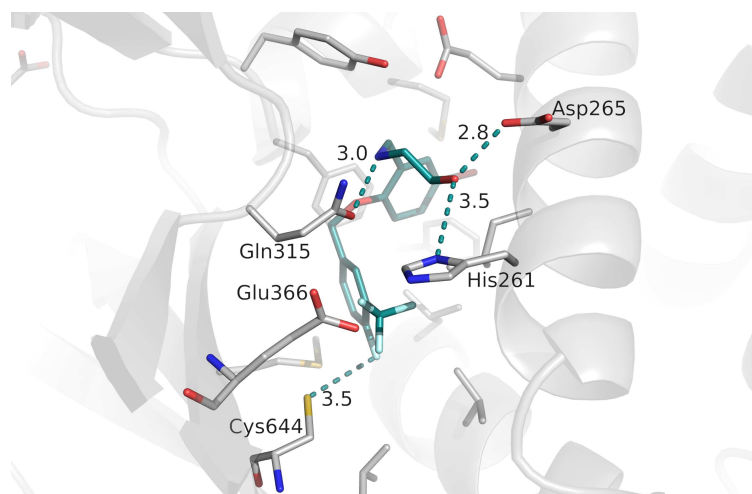
Unterschiede zwischen den finalen und den frühen Refinements vor. Daher werden hier die Ergebnisse basierend auf den ursprünglichen Refinements gezeigt.

Abbildung 4.44 zeigt ein frühes Refinement der (zum Zeitpunkt dieser Arbeit unveröffentlichten) Kristallstruktur von AZ1 im Komplex mit USP28 ohne die UCID-Spitze (USP28 $\Delta$ tip). Der Ligand befindet sich in einer Bindetasche zwischen der  $\alpha$ 5-Helix und den  $\beta$ -Faltblättern der „Handfläche“. Im Vergleich zu *apo*-Kristallstrukturen von USP28 ist  $\alpha$ 5 deutlich versetzt, um den Liganden zu akkommodieren. Besonders tief vergraben ist der Bromphenylring, der sich in einer lipophilen Tasche aus Phe186, Leu264, Leu267, Met288 und Phe292 (siehe Abbildung 4.44b) befindet. Nennenswert ist hierbei die Orientierung der Seitenkette von Met288, die vom Liganden wegdeutet. In *apo*-Kristallstrukturen (z. B. PDB 6HEJ) ist diese Aminosäure so orientiert, dass die lipophile Tasche zu klein für den Bromphenylring wäre.

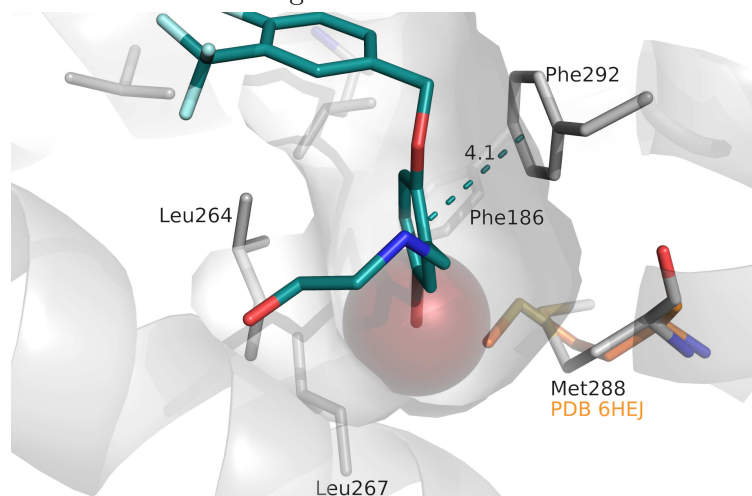
Die flexible Aminoethanolkette von AZ1 bildet mehrere Wasserstoffbrückenbindungen zu den umstehenden Aminosäuren aus. In diesem frühen Refinement ist für das vermutlich positiv geladene Amin eine Orientierung zu Gln315 modelliert, um der experimentell beobachteten Elektronendichte zu entsprechen. Die OH-Gruppe ist günstig zwischen His261 und Asp265 platziert. Der fluoridierte Phenylring des Moleküls ist parallel zur  $\alpha$ 5-Helix platziert. Dessen CF<sub>3</sub>-Gruppe ist unterhalb eines Bogens, der von Glu366 und His261 aufgespannt wird. In direkter Nachbarschaft zum F-Atom von AZ1 befindet sich Cys644 (Abstand 3.5 Å).

Abbildung 4.45 zeigt ein unveröffentlichtes frühes Refinement der Kristallstruktur von Vismodegib in USP28- $\Delta$ UCID. Zum Vergleich ist eine kurz vor dieser Arbeit publizierte Kristallstruktur von Vismodegib mit USP28, PDB 8HJE, zu sehen.<sup>80</sup> Beide Strukturen platzieren die chlorierte Sulfonylphenyl-Einheit in dieselbe Position zwischen Ser257 und Cys644 (siehe Abbildung 4.45a). So wird die Amid-Bindung direkt unter Glu366 positioniert. In der bereits publizierten Struktur weist das Amid einen Diederwinkel von 100.5° auf, sodass das Amid-Stickstoffatom auf Cys644 deutet (siehe Abbildung 4.45b). In der unveröffentlichten Struktur beträgt dieser Diederwinkel fast 180°, wie es aufgrund des partiellen Doppelbindungscharakters der Amid-Bindung erwartet wird. Durch diese physikalisch sinnvollere Geometrie kann außerdem eine Wasserstoffbrückenbindung zu Glu366 ausgebildet werden. Aufgrund dieses Unterschieds im Diederwinkel wird der Phenylring, der an dieses Amid anschließt und mit einem Pyridin-Ring in *meta*- und einem Chloratom in *para*-Position substituiert ist, in den beiden Strukturen genau entgegengesetzt orientiert: In PDB 8HJE befindet sich der Pyridin-Ring dort, wo sich das Chloratom in der hier diskutierten, unveröffentlichten Struktur befindet, und umgekehrt.

Abbildung 4.46 zeigt die beiden Kristallstrukturen aus Abbildung 4.44 und 4.45, durch PyMol<sup>1</sup> aligniert. Es wird deutlich, dass Teile des Vismodegib-Liganden und des AZ1-Liganden

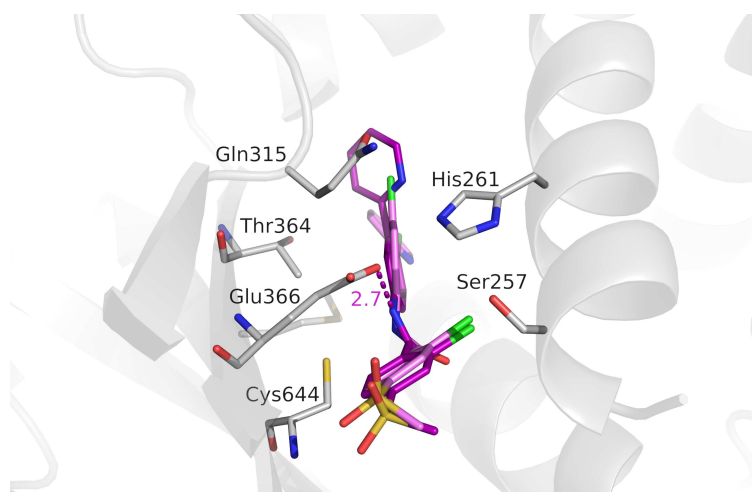


(a) Blick entlang der Bindungstasche und  $\alpha 5$ -Helix (auf der rechten Seite des Bildes). AZ1 ist hierbei in petrol gezeigt, das Protein in grau. Alle Abstände sind in  $\text{\AA}$  angegeben. Am unteren Bildrand ist Cys644 zu sehen, welches sich in unmittelbarer Nähe von AZ1 befindet (Abstand  $3.5 \text{ \AA}$ ). Die Abstände der flexiblen Aminoethanolkette zu unterschiedlichen Seitenketten des Proteins deuten allerdings auf Wasserstoffbrückenbindungen hin.

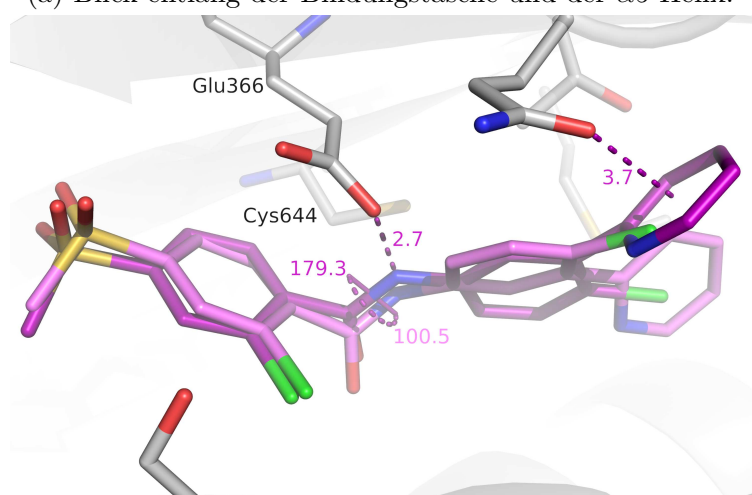


(b) Lipophile Tasche, die von AZ1 ausgefüllt wird. Die Größe der lipophilen Tasche wird durch die graue Oberfläche dargestellt. Zusätzlich wird die eigentliche Größe des Bromatoms (basierend auf dem van-der-Waals-Radius) von AZ1 in rotbraun angezeigt, das diese Tasche vollständig ausfüllt. In orange ist die Met288-Seitenkette von PDB 6HEJ gezeigt. Im Komplex mit AZ1 muss diese Seitenkette eine andere Orientierung annehmen, um nicht mit dem Bromatom von AZ1 zu kollidieren. Außerdem ist die Stacking-Interaktion zwischen AZ1 und Phe292 zu sehen.

Abbildung 4.44.: Unveröffentlichte Kristallstruktur von USP28 in Komplex mit AZ1, freundlicherweise zur Verfügung gestellt von Prof. Dr. Caroline Kisker und Dr. Florian Sauer. Es sind keine Protonen gezeigt, da diese nicht durch Röntgenbeugung aufgelöst werden.



(a) Blick entlang der Bindungstasche und der  $\alpha$ 5-Helix.



(b) Alternativer Blickwinkel mit Fokus auf die Amid-Bindung mit unterschiedlichem Diederwinkel. Nur für die unveröffentlichte Struktur ist der Winkel (mit einem Halbkreis angedeutet) explizit eingezeichnet, für PDB 8HJE ist nur der Wert angegeben.

Abbildung 4.45.: Kristallstrukturen von USP28 in Komplex mit Vismodegib. Die unveröffentlichte Struktur, die freundlicherweise von Prof. Dr. Caroline Kisker und Dr. Florian Sauer zur Verfügung gestellt wurde, ist in dunklem, die bereits publizierte Struktur PDB 8HJE in hellem Lila zu sehen.

denselben Bereich innerhalb der Bindetasche einnehmen. Der fluorierte Ring von AZ1 grenzt an den Pyridin-substituierten Phenylring von Vismodegib an.

Dieser Überlapp der AZ1- und Vismodegib-Bindestellen kann für das Design neuer Inhibitoren ausgenutzt werden, indem beide Moleküle miteinander verlinkt werden. Dieser „Mer-



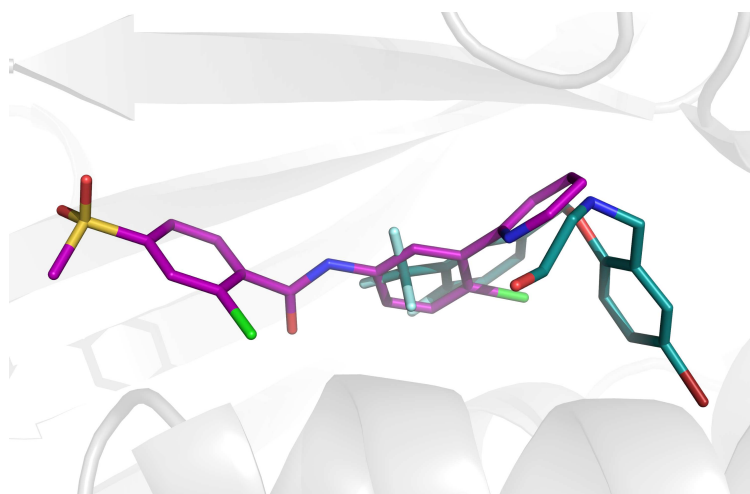


Abbildung 4.46.: Überlapp der AZ1- und Vismodegib-Bindestellen.

ging Approach“ findet breite Anwendung beim fragmentbasiertem Wirkstoffdesign<sup>176,177</sup> und versucht synergistische Effekte durch das gleichzeitige Adressieren von unterschiedlichen Taschen zu erzeugen. Di Zhou *et al.*<sup>80</sup> zeigen, dass ein derartiges Vorgehen vielversprechend sein könnte (auch wenn die Arbeitsgruppe vermutlich keine Kenntnisse von der AZ1-Struktur hatte). Wenn Vismodegib mit unterschiedlichen basischen Aminen verlängert, d. h. das Vismodegib-Gerüst in Richtung der AZ1-Bindestelle ausgedehnt wird, wird eine verbesserte Inhibition erreicht.<sup>80</sup> Für diese modifizierte Vismodegib-Verbindungen wurden Docking-Studien durchgeführt, die das Amin in der Nähe von Glu268 und Tyr293 platzieren. Diese Position ist ähnlich zu der des AZ1-Amins in der hier diskutierten, unveröffentlichten Kristallstruktur aus der Arbeitsgruppe um Prof. Dr. Caroline Kisker. Die Vermutung liegt nahe, dass derartige Vismodegib-basierte Liganden noch stärker an USP28 binden würden, wenn die tiefe lipophile Tasche zwischen Leu267, Met288 und Phe292 adressiert wird, die durch den Bromphenylring von AZ1 optimal ausgefüllt wird.

Alternativ dazu könnte Cys644 dafür genutzt werden, um eine kovalente Inhibitions-Strategie zu verfolgen. Der Abstand zwischen dem Fluoratom von AZ1 und dem Schwefelatom von Cys644 beträgt lediglich 3.5 Å (siehe Abbildung 4.44a). Dies ist vergleichbar zu den Abständen, die in Kapitel 4.1 für das nichtkovalente Modell eines Warheads und seines Targets beobachtet wurden. Es scheint daher plausibel, dass eine modifizierte AZ1-Struktur, die statt dem Fluoratom ein Elektrophil aufweist, in der Lage wäre, mit Cys644 zu reagieren. Wie in Kapitel 1.2.3.1 diskutiert könnte eine kovalente Inhibition von USP28 möglicherweise erlauben, *in vivo* zwischen USP25 und USP28 zu differenzieren und eine selektive Inhibition zu erreichen. Eine kovalente Inhibition bietet darüber hinaus eine Reihe an Vorteilen

wie erhöhte Affinität und verlangsamte Dissoziation vom Target im Vergleich zu klassischer nichtkovalenter Inhibition (siehe Kapitel 1.1.1 und Kapitel 1.1.2).

#### 4.2.2. Strukturvorschläge für mögliche kovalente Inhibitoren

Für das Design eines kovalenten Inhibitors wurde von der in Abbildung 4.44 gezeigten Kristallstruktur und dem Bindemodus von AZ1 ausgegangen. *Para*-Substitutionen am fluorierten Ring von AZ1 sollten hierbei den Warhead bereits günstig nahe an Cys644 platzieren. Um die sterischen Möglichkeiten diesbezüglich auszuloten wurde eine Reihe verschiedener Elektrophile gedockt (siehe Abbildung 4.47).

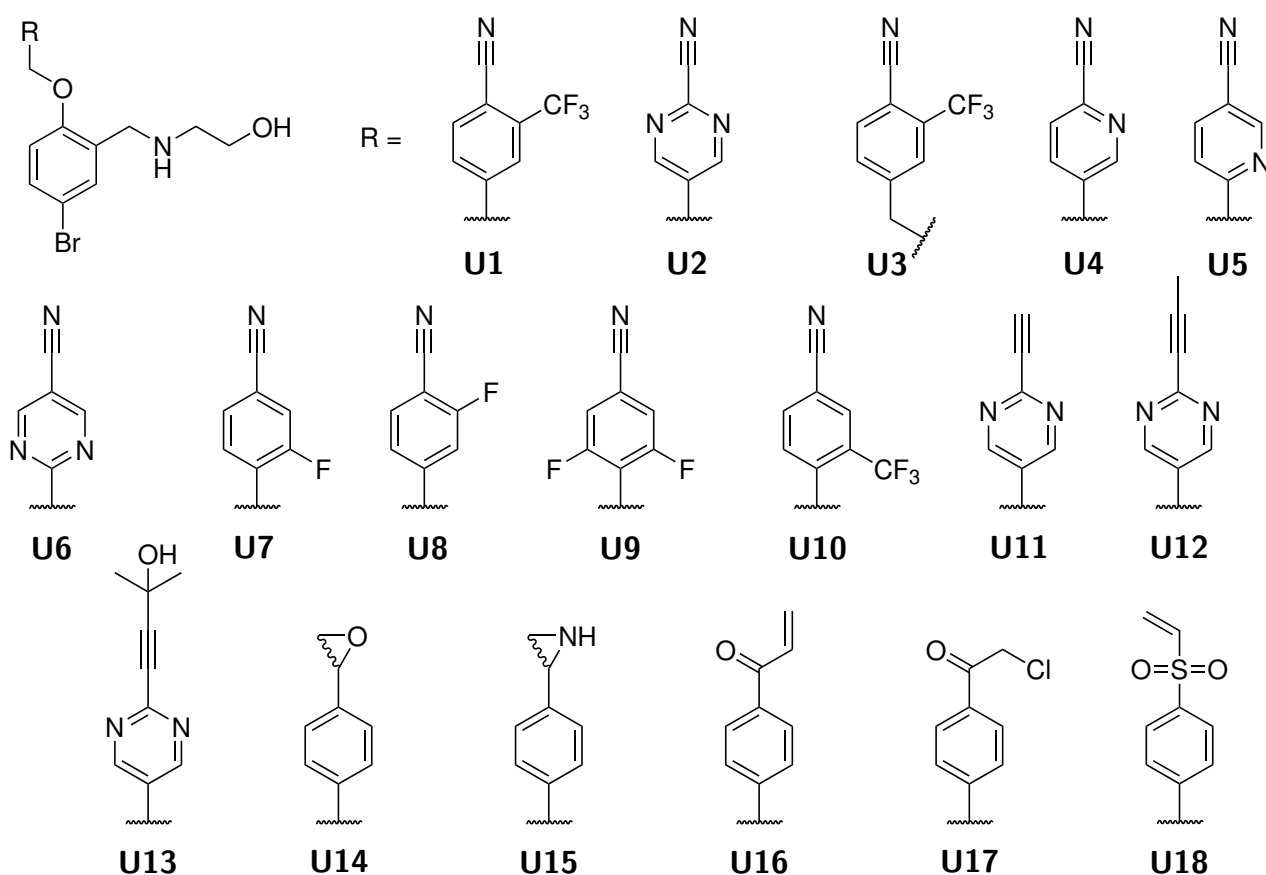


Abbildung 4.47.: AZ1-Variationen mit unterschiedlichen reaktiven Gruppen.

**U1** zeigt die höchste Ähnlichkeit zur ursprünglichen AZ1-Struktur **5**, da es die ursprüngliche CF<sub>3</sub>-Gruppe in *meta*-Position beibehält und die *para*-Fluorsubstitution durch ein Nitril ersetzt. In **U2** ist der fluorierte aromatische Ring durch ein Pyrimidin-System mit einem Nitril in *para*-Funktion ersetzt. Von allen gezeigten Nitril-Warheads sollte **U2** über das elek-

tronenärmste heteroaromatische System verfügen, das einen ähnlichen Raumbedarf aufweist wie ein Phenylring mit Nitrilfunktion. **U3** ist um eine CH<sub>2</sub>-Einheit länger als **U1**, um eine flexiblere Platzierung des Nitrils zu erlauben. Die Nitrile **U4** bis **U10** wurden in Zusammenarbeit mit Victor Olmos entworfen, da diese synthetisch leicht zugänglich wären und durch unterschiedliche Substitutionen unterschiedliche Reaktivitäten des Nitrils ermöglichen würden.

Alkine weisen einen sehr ähnlichen Raumbedarf zu Nitrilen auf, sind allerdings deutlich unreaktiver. In den Dockings wurden daher nur Alkine mit Pyrimidinring getestet. **U11** ist hierbei die einfachste Variante, während **U12** und **U13** schrittweise den endständigen Raumbedarf erhöhen. Für **U13** ist erwähnenswert, dass die endständige Isopropanol-Funktion in Alkin-Metathesen eine beliebte Schutzgruppe ist und unter basischen Bedingungen als Aceton abgespaltet werden kann (Retro-Favorski-Reaktion),<sup>178</sup> sodass **U11** entstehen würde.

**U14** und **U15** weisen sterisch kleine Warheads in der Form eines Epoxids und eines Aziridins auf, die irreversible Bindungen ausbilden. **U16** bis **U17** hingegen sind sterisch deutlich anspruchsvoller. Durch dieses breite Spektrum an Warheads werden unterschiedliche Kombinationen an Raumbedarf und Reaktivität ermöglicht.

### 4.2.3. Methodische Details der Docking-Protokolle

Mithilfe von Docking soll getestet werden, ob der Liganden mit Warhead die Pose von AZ1 reproduzieren können. Es wurden zwei unterschiedliche Docking-Ansätze verfolgt: Neben dem klassischen nichtkovalenten Docking von AutoDock3<sup>93</sup> und AutoDock4<sup>94</sup> wurde das WIDOCK-Protokoll (siehe Kapitel 3.1.2) getestet.<sup>102</sup> Dieses manipuliert die Lennard-Jones-Potenziale der beiden reaktive Atome so, dass diese sich bis auf 1.8 Å einander annähern können, was dem Abstand einer kovalenten Bindung entspricht. Diese Distanz wird allerdings nicht erzwungen, d. h. es ist immer noch möglich, dass im Rahmen des Dockings eine ungünstige Orientierung des Warheads generiert wird. Durch das WIDOCK-Protokoll kann hier insbesondere getestet werden, ob bei der Ausbildung der kovalenten Bindung signifikanten Änderungen der Pose entstehen würden.

Im Rahmen des Designs von kovalenten Inhibitoren wurden sowohl Dockings mit AutoDock3 als auch AutoDock4 durchgeführt. In dieser Arbeit werden allerdings lediglich die Dockings mit AutoDock4 diskutiert, da diese besser vergleichbar mit dem WIDOCK-Protokoll sind (welches eine Modifikation von AutoDock4 darstellt). Für die Inhibitoren, die eine AZ1- und eine Vismodegib-Struktur verlinken, werden in dieser Arbeit nur AutoDock3-Ergebnisse gezeigt.

Nach Aufbereitung des Proteins in MOE<sup>148</sup> (Ergänzen von fehlenden Seitenketten, Ergänzen von Endstücken an Loop-Enden, Entfernen von Ionen und Puffermolekülen aus der PDB-Datei) wurden für die Dockings mit AutoDock3 zwei *bash*-Tools genutzt: Mithilfe von *protonate* wurden alle nichtpolaren Protonen entfernt und alle polaren Protonen hinzugefügt, woraufhin anschließend mit *addchrgsol* Partiaalladungen gemäß des *AMBER united atom force fields*<sup>179</sup> und Solvatationsparameter<sup>93</sup> den Atomen des Proteins zugewiesen wurden. Problematisch war hierbei das in der Kristallstruktur vermutlich doppelt protonierte His261, da für dieses keine standardmäßigen Parameter in AutoDock hinterlegt sind. Partiaalladungen hierfür wurden daher, in Analogie mit der ursprünglichen Parametrisierung von AutoDock, aus dem *AMBER united atom force field*<sup>179</sup> übernommen, während Solvatationsparameter aus der Literatur gewählt wurden.<sup>180</sup> Zur Vorbereitung der Liganden wurden diese in MOE mit Hilfe des Kraftfeldes MMFF94x minimiert.<sup>150</sup> Protonierungszustände wurden ebenfalls in MOE mit Hilfe von *protonate3D*<sup>149</sup> bei pH=7 bestimmt. Dadurch lagen alle Amine protoniert vor, d. h. alle Liganden (bis auf die Aziridine) sind einfach positiv geladen. Die Aziridine sind zusätzlich am zyklischen Stickstoff protoniert. Auch wenn fraglich ist, ob diese doppelte Protonierung bei pH=7 wirklich korrekt ist, so ist sie zumindest praktikabler, da innerhalb des Docking-Algorithmus das Stickstoffatom nicht durch die Ringebene durchschwingen kann. Eine einfache Protonierung des Aziridin-Stickstoffatoms würde daher ein Pseudo-Stereozentrum erzeugen und den Docking-Aufwand verdoppeln. Jedem Atom wurden dann Gasteiger-Ladungen als Partiaalladungen zugewiesen.<sup>181</sup> Um die drehbaren Bindungen innerhalb des Liganden zu definieren, wurde das Kommandozeilen-Tool *autotors* genutzt.<sup>93</sup> In diesem Schritt werden ebenfalls alle nichtpolaren Wasserstoffatome entfernt, da diese implizit Teil der entsprechenden Kohlenstoffatome sind, in Übereinstimmung mit dem *united atom force field*. Es sollte betont werden, dass hierbei die Einfachbindung innerhalb des Michaelsystems von **U16** bzw. innerhalb des Vinylsulfons von **U18** als vollständig rotierbar angenommen wurde, auch wenn hier vermutlich ein partieller Doppelbindungscharakter vorliegt. Für die Dockings mit AutoDock4 wurde dieselbe Präparation mit Hilfe von *AutoDockTools* durchgeführt, um die notwendige Datei im erforderlichen *pdbqt*-Format zu erhalten.

Folgende Parameter wurden für den genetischen Algorithmus in AutoDock gewählt: Die Populationsgröße betrug 300, die maximale Anzahl an zu durchlaufenden Generationen 27 000 und die maximale Anzahl an Energieberechnungen  $12 \times 10^6$ . Die Raten für Mutation, Crossover und lokale Minimierung nach Solis und Wets<sup>182</sup> betragen 0.02, 0.8 und 0.06. 300 Iterationen wurden im Rahmen einer lokalen Minimierung durchgeführt, wobei nach vier konsekutiven Erfolgen oder Misserfolgen die Schrittweite der Minimierung angepasst wurde.

100 Durchläufe dieses genetischen Algorithmus wurden durchgeführt. Das Clustering der 100 erhaltenen Docking-Posen wurde mit Hilfe eines Cutoffs von 2 Å durchgeführt. Die am besten bewertete Pose sowie alle Posen, die einen RMSD von 2 Å oder weniger zu dieser aufwiesen, bildeten Cluster 1. Die nächstbeste Pose, die noch nicht gewählt wurde, bildete Cluster 2 (zusammen mit allen Posen, die einen RMSD von kleiner als 2 Å zu dieser aufwiesen). Dieses Vorgehen wird wiederholt, bis alle Posen einem Cluster zugewiesen sind. Wann immer möglich, wurde versucht, diese Parameter zwischen den unterschiedlichen Versionen und Protokollen von AutoDock gleich zu halten. Im Fall von hier nicht gelisteten Parametern wurde der Default-Wert übernommen.

Das Protein, bzw. dessen Wechselwirkung mit unterschiedlichen Atomtypen, wird durch unterschiedliche Gitter repräsentiert. Für jeden Atomtyp wird durch *AutoGrid*<sup>93</sup> ein Gitter generiert, und an jedem Gitterpunkt wird vorab der energetische Beitrag zum Score berechnet. Der Abstand zwischen Gitterpunkten betrug 0.375 Å; das Zentrum des Gitters wurde auf die gerundeten Koordinaten von Cys644 positioniert (das Gitter-Zentrum exakt auf Cys644 zu platzieren produzierte Fehlermeldungen). In x-, y- und z-Richtung wurde das zu betrachtende System durch jeweils 70, 90 und 60 Kantenlängen eines Gitterwürfels aufgespannt, d. h. die resultierende Box für das Docking war 26.25 Å x 33.75 Å x 22.5 Å groß.

Für das WIDOCK-Protokoll wurde Cys644 mit einer flexiblen Seitenkette definiert. Das modifizierte Potenzial zwischen den reaktiven Atomen wurde, in Analogie zur ursprünglichen Publikation,<sup>102</sup> mit einem 13-7-Potenzial anstelle des klassischen 12-6-Potenzials beschrieben. Allerdings wurde für alle Liganden die maximale Potenzialtopftiefe von 1 kcal/mol angenommen und nicht wie im originalen Protokoll anhand von Reaktivitätsdaten von 0.175 bis 1 skaliert,<sup>102</sup> da hier nur von Interesse war, ob prinzipiell die Warheads nahe genug am reaktiven Schwefelatom platziert werden können. Der Gleichgewichtsabstand der reaktiven Atome wurde auf die Länge einer C-S-Bindung (1.8 Å) gesetzt.

Die Alkine **U11**, **U12** und **U13**, die Epoxide (*R*)-**U14** und (*S*)-**U14** sowie die Aziridine (*R*)-**U15** und (*S*)-**U15** weisen prinzipiell innerhalb ihres Warheads zwei elektrophile Kohlenstoffatome auf, die prinzipiell in der Lage wären, von Cys644 nukleophil angegriffen zu werden. Für die Gruppe der heteroaromatischen Alkine existieren Kristallstrukturen, die darauf hindeuten, dass der Angriff auf das endständige Kohlenstoffatom bevorzugt ist, da so die Ladung des Intermediats an der Benzylposition befindlich ist und durch Mesomerie-Effekte und den elektronenziehenden Charakter des (Hetero-)Aromaten stabilisiert werden kann (vergleiche Abbildung 4.48).<sup>17</sup> Im Rahmen des WIDOCK-Protokolls wurde daher nur dieser Kohlenstoff als reaktives Atom untersucht. Für Aziridine und Epoxide ist bekannt, dass die Regioselektivität des nukleophilen Angriffs stark von den Bedingungen abhängt (sie-

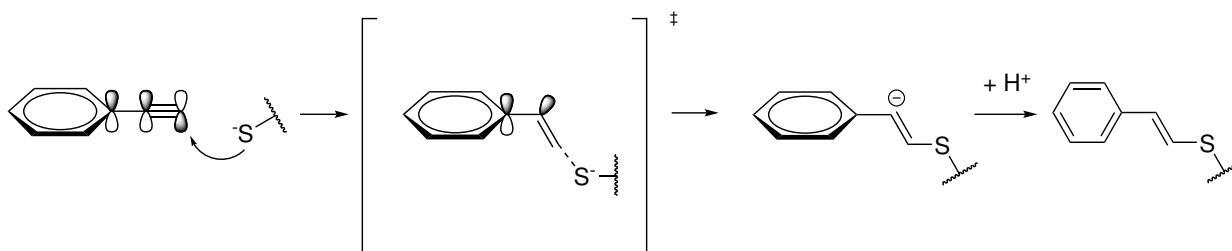


Abbildung 4.48.: Postulierter Mechanismus des Angriffs eines Thiolats an ein Alkin.

he Abbildung 4.49).<sup>183</sup> Für basische bis neutrale Bedingungen wird ein  $S_N2$ -Mechanismus angenommen, bei dem der Angriff am weniger substituierten Kohlenstoffatom erfolgt. Unter sauren Bedingungen kann ein  $S_N1$ -artiger Mechanismus vorliegen, sodass der Angriff am höher substituierten Kohlenstoffatom bevorzugt wird, da hier die intermediäre positive Ladung besser stabilisiert wird.<sup>183</sup> Dieser Effekt kann durch konjugierte  $\pi$ -Systeme verstärkt werden.<sup>184</sup> Kinetische Einflüsse, wie z. B. die sterisch leichtere Zugänglichkeit des geringer substituierten Kohlenstoffatoms, können aber auch unter sauren Bedingungen das  $S_N2$ -Produkt als Neben- oder Hauptprodukt begünstigen.<sup>183</sup> Es ist unklar, welche dieser möglichen Effekte überhaupt auf die Situation eines in einem Protein gebundenen Liganden übertragbar sind. Die sterisch leichtere Zugänglichkeit des  $S_N2$ -Kohlenstoffatoms könnte zum Beispiel irrelevant sein, wenn dieser nicht in der Nähe von Cys644 platziert werden kann. Daher wurden alle Epoxide und Aziridine sowohl mit jeweils dem  $S_N1$ - als auch dem  $S_N2$ -Kohlenstoffatom als reaktives Zentrum mit dem WIDOCK-Protokoll gedockt.

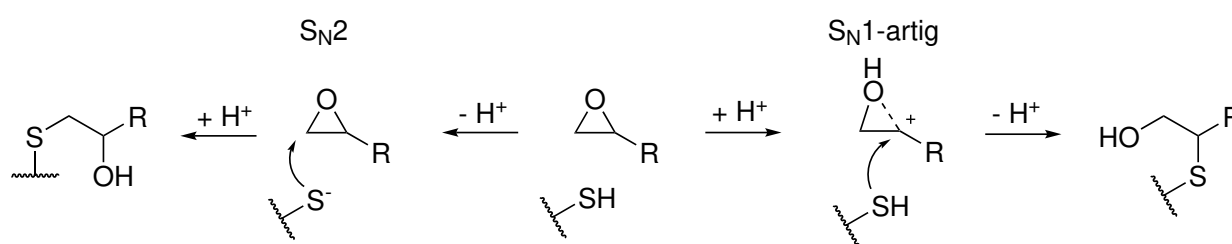


Abbildung 4.49.: Mögliche Mechanismen des Angriffs eines Thiols an ein Epoxid.

Zur Bewertung der erhaltenen Posen werden mehrere Kriterien herangezogen. Neben der *Estimated Free Enthalpy*, die AutoDock4 als Score benutzt, wurde mit Hilfe von *fconv*<sup>155</sup> ein Substruktur-RMSD im Vergleich zur Kristallstruktur von AZ1 berechnet. RMSD-Werte unter 2 Å gelten als eine ausreichend nahe Reproduktion der Kristallstruktur. Zusätzlich werden Abstandsmessungen zwischen dem Schwefelatom von Cys644 und dem elektrophilen Kohlenstoffatom der Verbindungen durchgeführt, um einschätzen zu können, ob die Ausbil-

dung einer kovalenten Bindung plausibel ist. Hierbei wird ein Cutoff von 4 Å verwendet, in Analogie zu den Ergebnissen in Kapitel 4.1. Diese Messung wird für alle Posen des Top-Clusters durchgeführt (da Posen, die zwar Teil desselben Clusters sind, prinzipiell einzelne Atome deutlich anders platzieren können). Für die WIDOCK-Ergebnisse wird ähnlich verfahren, allerdings liegt hier der Cutoff für die Abstandsmessung bei 2.2 Å. Es sollte betont werden, dass die erhaltene *Estimated Free Enthalpy* des WIDOCK-Protokolls aufgrund des veränderten Potenzials nicht mit den Ergebnissen aus dem unmodifizierten Autodock4 vergleichbar ist.

#### 4.2.4. Untersuchung der Möglichkeiten zur Entwicklung eines kovalenten Inhibitors basierend auf AZ1

##### 4.2.4.1. Docking-Ergebnisse der Nitril-Verbindungen

Nitrile sind zweiatomige Warheads, und daher im Vergleich mit anderen Warheads relativ klein.<sup>185</sup> Es müssen keine Freiheitsgrade innerhalb des Warheads berücksichtigt werden (wie z.B. bei Vinylsulfonen und Michael-Systemen). Die sterischen Ansprüche eines Liganden ändern sich bei der Inkorporation eines Nitrils nur geringfügig,<sup>185</sup> sodass die Affinität des AZ1-Gerüsts zum größten Teil erhalten bleiben sollte.

Ein wesentlicher Nachteil der Nitrile ist die Reversibilität ihrer Reaktion, was es experimentell anspruchsvoll macht, das Vorliegen einer kovalenten Bindung zweifelsfrei nachzuweisen. Für die Inhibition von USP28 wäre außerdem eine schnelle irreversible Inhibition von Vorteil, um Interaktionen mit USP25 zu vermeiden. Auf der anderen Seite sind die erhaltenen IC<sub>50</sub>-Werte von reversiblen Inhibitoren, anders als bei irreversiblen Inhibitoren, thermodynamisch sinnvoll und vergleichbar mit den IC<sub>50</sub>-Werten von nichtkovalenten Verbindungen (unter der Annahme, dass eine Gleichgewichtseinstellung gewährleistet ist).

Darüber hinaus sollte sich die Reaktivität eines Nitrils gut über den angrenzenden Aromaten steuern lassen, je nachdem, wie elektronenarm dieser aufgebaut ist<sup>186</sup> (vergleiche Kapitel 4.2.4.5). Außerdem sind Nitril-Bausteine für die Synthese leicht zugänglich, sodass unterschiedliche Nitrile mit unterschiedlichen Substitutionen und somit unterschiedlichen Reaktivitäten leicht getestet werden können sofern eine modulare Syntheseroute konzipiert wurde. Die Nitrile **U1** bis **U3** wurden aus sterischen und elektronischen Gründen ausgewählt, während die Nitrile **U4** bis **U10** in Zusammenarbeit mit Victor Olmos entworfen wurden (siehe Kapitel 4.2.2).

Die Ergebnisse des nichtkovalenten Dockings der Nitrile mit AutoDock4 sind in Tabelle 4.11 zusammengefasst. Alle Verbindungen weisen Substruktur-RMSDs unter 2.5 Å auf, in 9 von

10 Fällen sogar kleiner als 1.2 Å. Darüber hinaus weisen die Top-Posen aller Verbindungen außer **U9** einen reaktiven Abstand kleiner als 4 Å auf. Die Top-Pose von **U9** erfüllt diese Anforderung knapp nicht, andere Posen innerhalb des Top-Clusters allerdings durchaus. Das zeigt, dass es für diese Nitril-Liganden durchaus möglich ist, ihr AZ1-Grundgerüst wie in der Kristallstruktur zu platzieren und zusätzlich den Warhead so zu positionieren, dass er nahe an Cys644 platziert ist. Die Docking-Ergebnisse legen damit nahe, dass eine Reaktion möglich ist. Abbildung 4.50 zeigt exemplarisch die Top-Pose von **U4** im Vergleich zur Kristallstruktur.

Verbindung	Estim. Free Enthalpy, Top-Pose [kcal/mol]	RMSD Top-Pose [Å]	Distanz Top-Pose [Å]	Minimale Distanz Top-Cluster [Å]	Größe Top-Cluster
<b>U1</b> ( <i>o</i> -CF <sub>3</sub> )	-10.59	1.08	3.93	3.20	11
<b>U2</b> (2-Pyrimidin)	-10.37	0.93	3.56	3.36	26
<b>U3</b> ( <i>o</i> -CF <sub>3</sub> , Linker verlängert)	-10.87	2.37	3.83	3.64	7
<b>U4</b> (2-Pyridin)	-10.51	0.87	3.57	3.32	28
<b>U5</b> (3-Pyridin)	-10.62	0.97	3.45	3.42	28
<b>U6</b> (5-Pyrimidin)	-10.16	0.94	3.53	3.22	41
<b>U7</b> ( <i>m</i> -fluoro)	-10.87	0.82	3.52	3.44	28
<b>U8</b> ( <i>o</i> -fluoro)	-10.88	1.06	3.83	3.49	30
<b>U9</b> ( <i>m,m</i> -difluoro)	-10.40	1.10	4.17	3.57	12
<b>U10</b> ( <i>m</i> -CF <sub>3</sub> )	-11.22	1.09	3.68	3.68	4

Tabelle 4.11.: Nichtkovalente AutoDock4-Ergebnisse der AZ1-basierten Nitrile. Für die Verbindungen ist, in Relation zum Nitril, ein kurzer Deskriptor des strukturellen Unterschieds angegeben. Neben der *Estimated Free Enthalpy* wird der Substruktur-RMSD des AZ1-Gerüsts im Vergleich zur Kristallstruktur angegeben. Die angegebenen Distanzen wurden zwischen dem reaktiven Kohlenstoffatom und dem Schwefelatom von Cys644 gemessen.

Zusätzlich zu den klassischen Dockings wurde das WIDOCK-Protokoll für die Verbindungen **U1**, **U2** und **U3** getestet. Tabelle 4.12 fasst die Ergebnisse dieser Testläufe zusammen.

Im ursprünglichen WIDOCK-Protokoll wurde ein Ligand dann als kovalent bindender Hit eingestuft, wenn für die am besten bewertete Pose ein C-S-Abstand kleiner als 2.2 Å beobachtbar war. Die zusätzlichen 0.4 Å Abstand im Vergleich zu den idealen 1.8 Å Bindungslänge sollten repulsive van-der-Waals-Wechselwirkungen mit anderen Atomen ausgleichen.<sup>102</sup> Dieses streng formulierte Kriterium wird in Tabelle 4.12 von keiner Verbindung erfüllt. Selbst wenn der komplette Top-Cluster betrachtet wird, kann keine Pose gefunden werden, die einen C-S-Abstand von 2.2 Å oder kleiner aufweist. Im Vergleich zu Tabelle 4.11 wurden nur



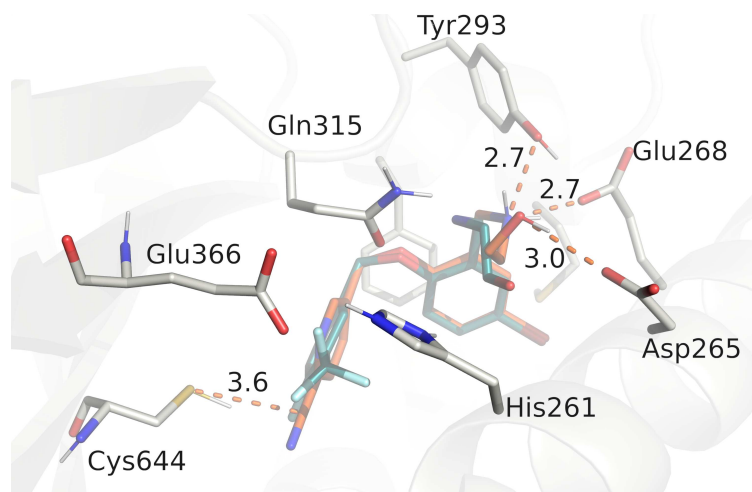


Abbildung 4.50.: Top-Pose von **U4** in orange, AZ1 (**5**, siehe Kapitel 1.2.3.1) als Referenz in petrol. Die eingezeichneten Abstände sind in Å angegeben und kennzeichnen Wasserstoffbrückenbindungen sowie den Abstand zwischen den reaktiven Atomen. Es werden die polaren Wasserstoffatome nur für das aufbereitete Protein und den gedockten Liganden gezeigt, nicht für die Kristallstruktur-Referenz.

minimale Annäherungen erreicht. Visuelle Inspektion der entsprechenden Top-Posen zeigte keine offensichtliche Kollision zwischen Ligand und Protein, die den Warhead auf Abstand halten müsste (siehe exemplarisch Abbildung 4.51 für **U2**). Erneut reproduziert die Top-Pose den Bindemodus der Kristallstruktur, auch wenn im Rahmen des genetischen Algorithmus ein etwas anderes Wasserstoffbrückennetzwerk für die flexible Aminoethanolkette gefunden wird. Allerdings liegt immer noch ein reaktiver Abstand von 3.15 Å vor.

Es liegt daher die Vermutung nahe, dass hier methodische Probleme mit dem WIDOCK-Protokoll vorliegen, sodass selbst mit den modifizierten Parametern für das elektrophile Kohlenstoffatom das Abstandskriterium von 2.2 Å nicht erreicht werden kann.

Daher wurde versucht, eine zweidimensionale Darstellung des Lennard-Jones-Potenzials eines Nitrils und eines Schwefelatoms zu erstellen, um so die Analyse der Problematik zu erleichtern (siehe Abbildung 4.52). Eine zweidimensionale Beschreibung ist für ein lineares System wie das Nitril aufgrund der Rotationssymmetrie ausreichend.

Im Rahmen dieses Modellsystems wird ein Nitril mit seinem elektrophilen Kohlenstoffatom im Ursprung eines Koordinatensystem platziert. Das Nitril-Stickstoffatom wird hierbei in die negative Richtung der x-Achse ausgerichtet; das Kohlenstoffatom, das die Verbindung zum restlichen Liganden darstellen soll, entlang der positiven x-Achse. Die Abstände zwischen den Atomen entsprechen denen einer C-N-Dreifachbindung (1.2 Å) und einer C-C-Einfachbindung (1.4 Å). Zusätzlich wird innerhalb dieses Koordinatensystems ein 12 Å x

Verbindung	Estim. Free Enthalpy, Top-Pose [kcal/mol]	RMSD Top-Pose [Å]	Distanz Top-Pose [Å]	Minimale Distanz Top-Cluster [Å]	Größe Top-Cluster
<b>U1</b> ( <i>o</i> -CF <sub>3</sub> )	-10.44	1.24	5.73	3.96	5
<b>U2</b> (2-Pyrimidin)	-11.00	1.09	3.15	3.02	29
<b>U3</b> ( <i>o</i> -CF <sub>3</sub> , Linker verlängert)	-11.14	2.40	3.62	3.62	5

Tabelle 4.12.: Docking-Ergebnisse des WIDOCK-Tests mit den Nitrilen **U1** - **U3**. Für die Verbindungen ist, in Relation zum Nitril, ein kurzer Deskriptor des strukturellen Unterschieds angegeben. Neben der *Estimated Free Enthalpy* wird der Substruktur-RMSD des AZ1-Gerüsts im Vergleich zur Kristallstruktur angegeben. Die angegebenen Distanzen wurden zwischen dem reaktiven Kohlenstoffatom und dem Schwefelatom von Cys644 gemessen.

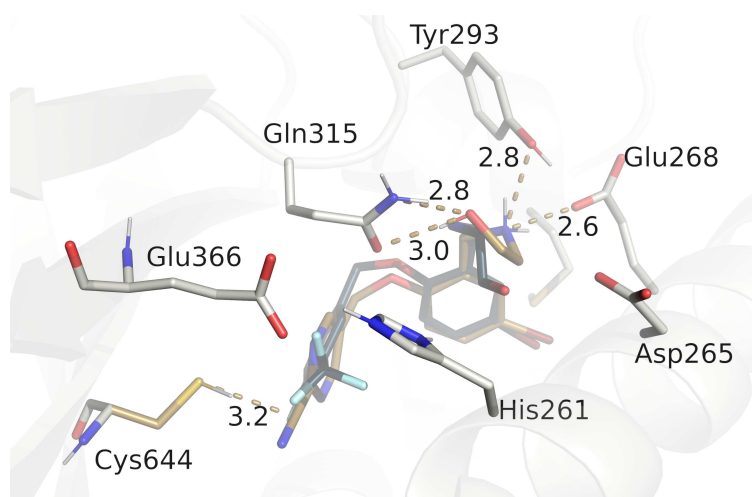


Abbildung 4.51.: WIDOCK-Ergebnis für **U2** mit linearen Warhead. In dunkeltürkis ist die Kristallstruktur von AZ1 (**5**) als Referenz gezeigt, während die Top-Pose in braun dargestellt ist. Die eingezeichneten Abstände sind in Å angegeben und kennzeichnen Wasserstoffbrückenbindungen sowie den Abstand zwischen den reaktiven Atomen. Es werden nur polare Wasserstoffatome dargestellt, allerdings nicht für die Kristallstruktur-Referenz.

12 Å großes Gitter aufgespannt, zentriert auf den Ursprung und mit einem Abstand von 0.1 Å zwischen jedem Gitterpunkt. Vor dem Hintergrund, dass der Standard-Gitterpunktabstand in AutoDock 0.375 Å beträgt, stellt dies ein vergleichsweise dichtes Gitter dar und soll für dieses Modellsystem genügen. Für jeden Gitterpunkt wird nun der Wert der van-der-Waals-Wechselwirkung berechnet, die entsteht, wenn ein Schwefelatom S<sub>Z</sub> sich an dieser Stelle befindet. AutoDock4 nutzt die in Gleichung 4.1 dargestellte Form für derartige Wechselwirkungen; der Nutzer kann dann selbst für jedes Paar aus Atomtypen *i* und *j* definieren,

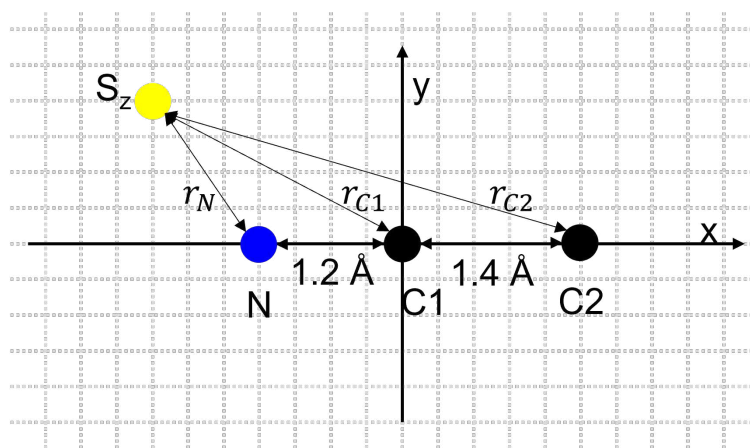


Abbildung 4.52.: Aufbau des Modellsystems zur Untersuchung der Problematik mit dem WIDOCK-Protokoll. Ein Stickstoffatom (N) und zwei Kohlenstoffatome (C1 und C2) sind linear angeordnet und weisen Abstände gemäß eines Nitrils zueinander auf. Für ein Schwefelatom  $S_Z$  werden die x- und y-Koordinaten variiert, die Abstände zu den Nitrilatomen gemessen ( $r_N$ ,  $r_{C1}$  und  $r_{C2}$ ) und dann die Summe der resultierenden van-der-Waals-Wechselwirkungen gebildet. Für das reaktive Kohlenstoffatom C1 werden die Parameter gemäß den unterschiedlichen Docking-Protokollen gewählt.

in welchem Abstand  $r_{ij}$  das Minimum des Potentials auftreten soll, welche Potentialtopftiefe  $\epsilon_{ij}$  vorliegen soll und mit welchen Potenzen  $n$  und  $m$  die repulsiven bzw. attraktiven Terme berechnet werden sollen. Die in Tabelle 4.13 gezeigten Parameter wurden für diese Berechnungen verwendet und sind an jene von AutoDock3 angelehnt. Die Gesamtwechselwirkung ergibt sich aus der Summe der einzelnen Interaktionen zwischen den Atompaaaren. Für das Nitril-Kohlenstoffatom C1 werden zwei unterschiedliche Atomtypen angenommen. Mit der Verwendung des (klassischen) aliphatischen Kohlenstoffatomtyps C wird das AutoDock-Potenzial  $V_{AutoDock}$  modelliert (Gleichung 4.2), für die Modellierung des WIDOCK-Potenzials werden die Parameter eines (maximal) reaktiven Kohlenstoffatoms  $C_Z$  verwendet (Gleichung 4.3).

$$V_{ij}(r) = \frac{\frac{m}{n-m}\epsilon_{ij}r_{ij}^n}{r^n} - \frac{\frac{n}{m-n}\epsilon_{ij}r_{ij}^m}{r^m} \quad (4.1)$$

Atomtyp-Paar	$r_{ij}$ [Å]	$\epsilon_{ij}$ [kcal/mol]	$n$	$m$
N- $S_Z$	3.75	0.179	12	6
C- $S_Z$	4.00	0.173	12	6
$C_Z$ - $S_Z$	1.80	1.000	13	7

Tabelle 4.13.: Relevante Parameter für das Modellsystem.

$$V_{\text{Autodock4}}(r_N, r_{C1}, r_{C2}) = V_{N-SZ}(r_N) + V_{C-SZ}(r_{C1}) + V_{C-SZ}(r_{C2}) \quad (4.2)$$

$$V_{\text{WIDOCK}}(r_N, r_{C1}, r_{C2}) = V_{N-SZ}(r_N) + V_{CZ-SZ}(r_{C1}) + V_{C-SZ}(r_{C2}) \quad (4.3)$$

Abbildung 4.53 zeigt die Heatmaps des oben beschriebenen Modellsystems. Abbildung 4.53a beschreibt den Standardfall in klassischen AutoDock-Dockings. Das Potenzialminimum befindet sich 3.8 Å entfernt vom elektrophilen Kohlenstoffatom. Da der repulsive Term des Lennard-Jones-Potenzials mit der zwölften Potenz des Abstandes skaliert, schlägt bei leicht kürzeren Abständen bereits das Vorzeichen des Potenzials um, und es werden abstoßende Wechselwirkungen festgestellt. In Dockings, wie in diesem Kapitel beschrieben, können auch kürzere Abstände beobachtet werden, da hier zusätzlich attraktive Interaktionen, wie z. B. Coulomb-Wechselwirkungen, eine Rolle spielen können. In Abbildung 4.53b ist im Vergleich hierzu das 2D-Lennard-Jones-Potenzial für das WIDOCK-Protokoll zu sehen. Das innerhalb des Gitters gefundene Minimum ist mit 3.6 Å nur marginal näher am Ursprung platziert. Innerhalb eines Radius von 2.2 Å um den Ursprung herum sind keinerlei negative Potenzialwerte auffindbar. Dies bedeutet, dass es energetisch ungünstig wäre, das Abstandskriterium des WIDOCK-Protokolls zu erfüllen. Da WIDOCK aber per Definition des modifizierten Potenzials der reaktiven Atome Abstände bis zu 1.8 Å tolerieren müsste, muss die in Abbildung 4.53b gezeigte Abstoßung auf einen sterischen Konflikt mit den beiden angrenzenden Atomen zurückzuführen sein. Der Zusammenstoß mit den anderen beiden Atomen scheint bei dem Cutoff-Wert von 2.2 Å so gravierend zu sein, dass der im modifizierten Potenzial definierte energetische Benefit dies nicht ausgleichen kann. Vor dem Hintergrund, dass die Repulsion mit der zwölften Potenz ansteigt, ist dies wenig verwunderlich.

Darüber hinaus ist das in Abbildung 4.53b beobachtete Minimum auch energetisch weniger stark ausgeprägt als im ursprünglichen AutoDock4-Protokoll (-0.37 kcal/mol im Vergleich zu -0.49 kcal/mol). Dies ist dadurch zu erklären, dass es in Abbildung 4.53b zu einer geringeren Überlappung der einzelnen Lennard-Jones-Minima kommt (da das reaktive Atom zu klein ist).

Offensichtlich ist das originale WIDOCK-Protokoll damit nicht geeignet, Nitrile zu beschreiben. Bei anderen Warheads existiert diese Problematik nicht oder in deutlich geringerem Maße, da das elektrophile Kohlenstoffatom entweder endständig ist oder nicht von benachbarten Atomen eingehüllt wird, sodass es sterisch leichter zugänglich ist.

Damit ergibt sich die Frage, wie das Protokoll modifiziert werden muss, um zu sinnvollen Ergebnissen zu führen. Die naheliegende Vorgehensweise hierfür wäre, mit der gewinkelten

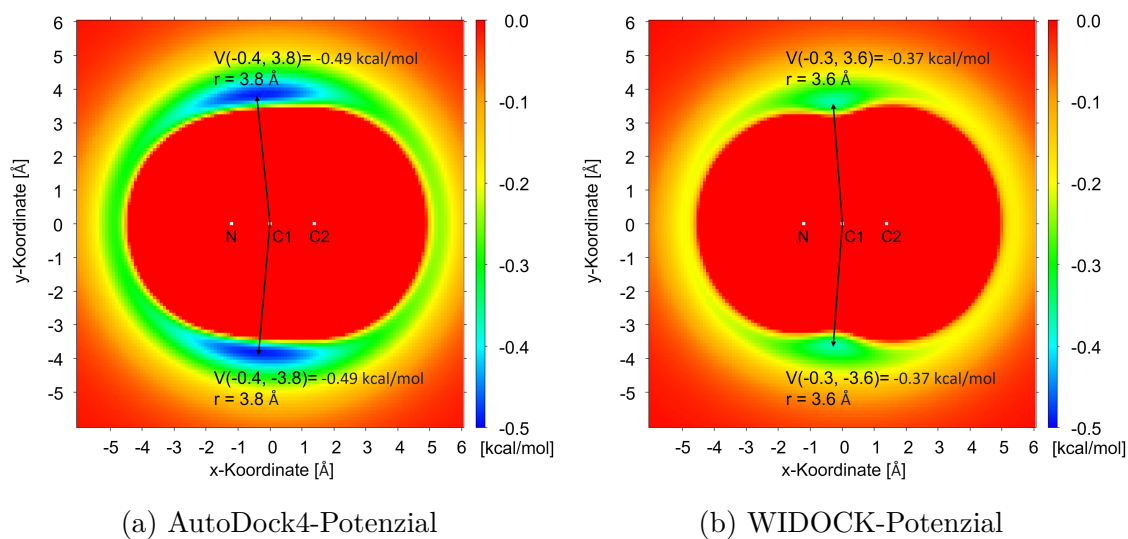


Abbildung 4.53.: Berechnete 2D-Lennard-Jones-Potenziale, die ein Schwefelatom in der Nähe eines Nitrils erfahren würde, entsprechend dem Aufbau des Gedankenexperiments. Die Farbskala der Heatmaps unterscheidet nur negative Werte, da die positiven (repulsiven) Werte mit der zwölften Potenz ansteigen. Letztere werden daher alle rot dargestellt.

Geometrie des Thioimidat-Produkts zu arbeiten, anstelle der linearen des Nitrils. Auf diese Art und Weise wird der Einfluss der benachbarten Atome verringert und das vergrabene zentrale Kohlenstoffatom zugänglicher. Andere Vorgehensweisen, wie z. B. die van-der-Waals-Parameter der benachbarten Atome ebenfalls noch zu modifizieren, oder den Potenzialtopf des WIDOCK-Potenzials noch tiefer zu senken, laufen Gefahr, chemisch nicht sinnvolle Lösungen zu produzieren. Alternativ könnte der Cutoff erhöht werden, ab dem ein WIDOCK-Docking als erfolgreich eingestuft wird; allerdings müsste dieser dann so hoch gesetzt werden, dass wenig Unterschied zu einem klassischen Docking bestünde.

Aus diesen Gründen wurde das WIDOCK-Docking aller Nitril-Verbindungen auch mit gewinkelten Warheads durchgeführt. Hierfür wurden die Liganden erneut in MOE geladen, analog zum Reaktionsprodukt modifiziert und minimiert. Das polare Proton, das sich nach der Reaktion am (ehemaligen Nitril-)Stickstoffatom befindet, wurde allerdings entfernt, da sich dieses im Rahmen des Dockings formal noch an Cys644 befindet. Außerdem wurde die Bindung zwischen Aromat und Nitril als rotierbar angenommen; dies war bei der zuvor linearen Geometrie des Nitrils vernachlässigbar. Eine vollständige Rotierbarkeit ist ebenfalls kritisch zu betrachten, da das entstehende Thioimidat als  $\pi$ -System mit dem Aromaten in Konjugation steht. Ohne Kenntnis des exakten Torsionspotenzials für diese Bindung erscheint es aber sinnvoller und praktikabler, schlicht eine komplette Drehbarkeit anzunehmen.

Die Resultate sind in Tabelle 4.14 zusammengefasst.

Verbindung	Estim. Free Enthalpy, Top-Pose [kcal/mol]	RMSD Top-Pose [Å]	Distanz Top-Pose [Å]	Minimale Distanz Top-Cluster [Å]	Größe Top-Cluster
<b>U1</b> ( <i>o</i> -CF <sub>3</sub> )	-10.53	0.77	3.96	3.89	2
<b>U2</b> (2-Pyrimidin)	-10.67	0.78	2.18	2.13	18
<b>U3</b> ( <i>o</i> -CF <sub>3</sub> , Linker verlängert)	-10.23	2.47	4.81	4.00	7
<b>U4</b> (2-Pyridin)	-10.48	1.29	4.07	2.11	11
<b>U5</b> (3-Pyridin)	-11.02	1.69	2.08	2.08	16
<b>U6</b> (5-Pyrimidin)	-10.69	0.94	2.13	2.10	11
<b>U7</b> ( <i>m</i> -fluoro)	-11.11	1.21	2.26	2.08	12
<b>U8</b> ( <i>o</i> -fluoro)	-10.75	1.70	2.14	2.11	10
<b>U9</b> ( <i>m,m</i> -difluoro)	-9.74	1.67	4.30	2.29	4
<b>U10</b> ( <i>m</i> -CF <sub>3</sub> )	-10.95	1.15	2.13	2.13	8

Tabelle 4.14.: WIDOCK-Ergebnisse für die Nitril-Verbindungen unter Nutzung eines gewinkelten Warheads. Für die Verbindungen ist, in Relation zum Nitril, ein kurzer Deskriptor des strukturellen Unterschieds angegeben. Neben der *Estimated Free Enthalpy* wird der Substruktur-RMSD des AZ1-Gerüsts im Vergleich zur Kristallstruktur angegeben. Die angegebenen Distanzen wurden zwischen dem reaktiven Kohlenstoffatom und dem Schwefelatom von Cys644 gemessen.

Die modifizierte Warhead-Geometrie erlaubt in vielen Fällen, das Abstandskriterium von 2.2 Å mit einer Pose des Top-Clusters zu erfüllen. **U9** erscheint hier leicht problematisch, da der Score der Top-Pose im Vergleich zu den anderen Verbindungen gesenkt ist und auch die Clustergröße des Top-Clusters abgenommen hat. Scheinbar wirkt sich der höhere sterische Anspruch der Verbindung aus. Exemplarisch ist hier die Top-Pose von **U6** in Abbildung 4.54 gezeigt.

Interessanterweise ist bei **U1** ein sehr niedriger RMSD-Wert im Vergleich zur Kristallstruktur beobachtbar, allerdings wird das WIDOCK-Kriterium auch mit gewinkeltem Warhead nicht erfüllt. Ebenso ist bei **U3** der Warhead zu weit entfernt. Ähnlich wie bei **U9** sind die Clustergrößen reduziert. Tabelle 4.15 zeigt für **U1** zusätzlich zur Top-Pose noch die Ergebnisse des zweitbesten Clusters. Dieser weist einen schlechteren Score und schlechteren RMSD-Wert sowie eine ähnlich geringe Clustergröße auf, erfüllt aber das Abstandskriterium, das für WIDOCK gesetzt ist. Um diesen Sachverhalt näher darzustellen zeigt Abbildung 4.55 die beiden Posen im Vergleich. Der Hauptunterschied zwischen den beiden Posen ist die Orientierung der CF<sub>3</sub>-Gruppe. Bei dem Top-Cluster (Abbildung 4.55a) ist diese, wie in der Kristallstruktur,

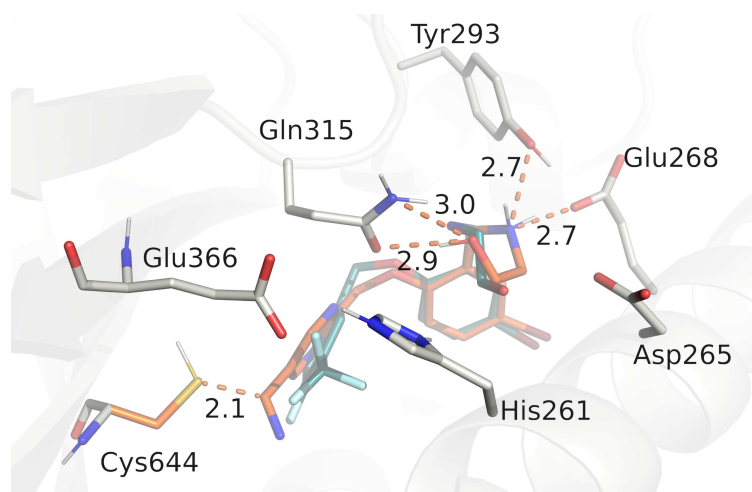
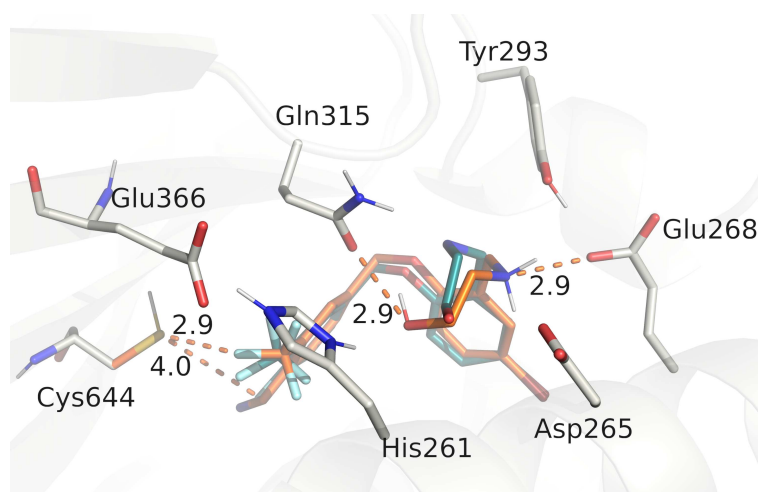


Abbildung 4.54.: Top-Pose von **U6** mit gewinkelten Warhead in orange, AZ1 (**5**) als Referenz in petrol. Die eingezeichneten Abstände sind in Å angegeben und kennzeichnen Wasserstoffbrückenbindungen sowie den Abstand zwischen den reaktiven Atomen. Es werden nur polare Wasserstoffatome dargestellt, allerdings nicht für die Kristallstruktur-Referenz.

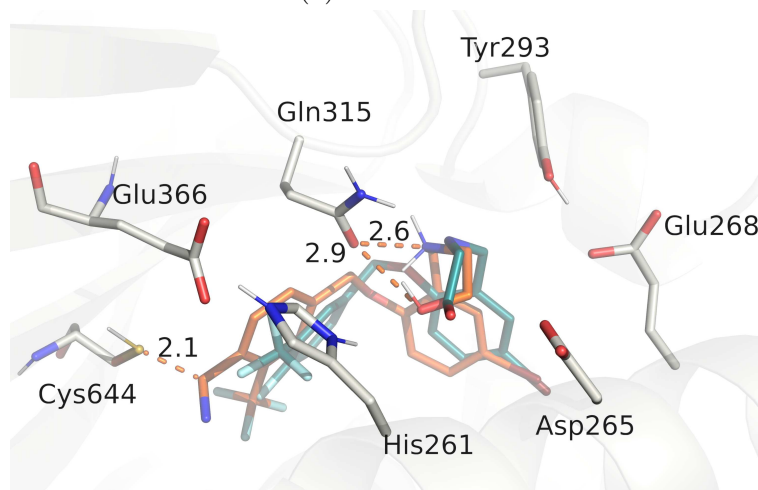
in Richtung His261 und Glu366 ausgerichtet. Zwischen dem Schwefelatom von Cys644 und dem nächsten Fluoratom von **U1** liegen nur 2.9 Å. Würde die Orientierung der CF<sub>3</sub>-Gruppe beibehalten, aber das aromatische System näher an Cys644 verschoben werden, bis das Nitril das Abstandskriterium erfüllt, käme es zur Kollision zwischen dem Fluoratom und Cys644. Für diese kristallnahe Pose scheint es daher nicht möglich, den Warhead nahe genug an der reaktiven Aminosäure zu platzieren. Abbildung 4.55b unterstreicht diesen Punkt. Für den zweitbesten Cluster liegt der Abstand zwischen Cys644 und dem reaktiven Kohlenstoffatom unter dem Cutoff, allerdings muss hier die CF<sub>3</sub>-Gruppe in die entgegengesetzte Richtung orientiert werden, was die komplette aromatische Einheit verschiebt. Zusätzlich muss der Bromphenyl-Ring von der Kristallstruktur-Positionierung abweichen. Insgesamt wird daher die Top- (und einzige) Pose von Cluster 2 um ca. 0.6 kcal/mol schlechter bewertet als die Top-Pose von Cluster 1.

<b>U1</b>	Docked Energy Top-Pose [kcal/mol]	RMSD Top-Pose [Å]	Distanz Top-Pose [Å]	Minimale Distanz Top-Cluster [Å]	Größe Top- Cluster
Top-Pose Cluster 1	-10.53	0.77	3.96	3.89	2
Top-Pose Cluster 2	-9.94	2.35	2.14	2.14	1

Tabelle 4.15.: WIDOCK-Ergebnisse der beiden Top-Cluster von **U1**. Cluster 1 weist einen sehr geringen RMSD-Wert zur Kristallstruktur auf, während Cluster 2 das Abstandskriterium erfüllt.



(a) Cluster 1



(b) Cluster 2

Abbildung 4.55.: Docking-Ergebnisse des WIDOCK-Dockings mit gewinkelten Warhead für **U1**. Gezeigt sind die Top-Posen der beiden bestbewerteten Cluster. Die eingezeichneten Abstände sind in Å und kennzeichnen Wasserstoffbrücken sowie relevante Abstände zu Cys644. Die Docking-Lösungen sind in orange gezeigt, die Kristallstruktur von AZ1 (**5**) in petrol. Es werden nur polare Wasserstoffatome dargestellt, allerdings nicht für die Kristallstruktur-Referenz.

Zusammenfassend lässt sich damit sagen, dass es im Rahmen des klassischen nichtkovalenten Dockings als auch im Rahmen des WIDOCK-Dockings möglich ist, Posen für AZ1-artige Nitrile zu generieren, die den Warhead nahe an Cys644 platzieren als auch einen niedrigen RMSD-Wert zur Kristallstruktur-Pose aufweisen. Steigt der sterische Anspruch der Substitution jedoch, insbesondere in der Nähe des Warheads, wird es schwierig, diese Kriterien zu erfüllen. Eine *ortho*-Substituierung, wie hier mit einer CF<sub>3</sub>-Gruppe, scheint bereits zu



groß zu sein. Damit wäre von den Warheads **U1** und **U3** abzuraten, während die übrigen Warheads geeignet erscheinen.

#### 4.2.4.2. Docking-Ergebnisse der Alkin-Verbindung

Das Alkin als Warhead weist viele Parallelen zu einem Nitril auf: Prinzipiell müssen nur zwei zusätzliche Schweratome in eine Ligand-Struktur eingefügt werden, sodass, bei geeigneter Positionierung des Warheads, der ursprüngliche Bindemodus des Liganden nur minimal adaptiert werden muss. Anders als Nitrile können Alkine zusätzlich substituiert werden, sodass von ihnen ausgehend der Ligand verlängert werden kann. Zusätzlich ist ihr Reaktionsmechanismus irreversibel, was, wie in Kapitel 1.2.3.1 diskutiert, günstiger für die Inhibition von USP28 wäre. Kritisch für Alkine ist allerdings die hohe Reaktionsbarriere (siehe Kapitel 4.2.4.5), was nicht nur die Reaktion mit dem unaktivierten Cys644 erschwert, sondern auch die Syntheserouten aufwendiger macht.

Verbindung	Estim. Free Enthalpy, Top-Pose [kcal/mol]	RMSD Top-Pose [Å]	Distanz Top-Pose [Å]	Minimale Distanz Top-Cluster [Å]	Größe Top-Cluster
<b>U11</b> (H)	-10.58	1.27	3.86	3.62	29
<b>U12</b> (Me)	-10.62	1.20	3.85	3.65	20
<b>U13</b> (2-Propanol)	-11.03	1.04	4.31	4.12	3

Tabelle 4.16.: Ergebnisse des nichtkovalenten Dockings mit AutoDock4 für die Alkin-Verbindungen **U11** - **U13**. Für die Verbindungen wird die endständige Substitution des Alkins angegeben. Neben der *Estimated Free Enthalpy* wird der Substruktur-RMSD des AZ1-Gerüsts im Vergleich zur Kristallstruktur angegeben. Die angegebenen Distanzen wurden zwischen dem reaktiven Kohlenstoffatom und dem Schwefelatom von Cys644 gemessen.

Die Ergebnisse der Alkine **U11**, **U12** und **U13** für das nichtkovalente Docking mit AutoDock4 sind in Tabelle 4.16 zusammengefasst. Die drei Alkine zeigen immer größere werdende Substitutionen auf, wobei **U13** mit der 2-Propanol-Schutzgruppe interessant war, da diese häufig im Rahmen der Sonogashira-Kupplung genutzt wird. **U13** könnte theoretisch durch Entschützung im Basischen zu **U11** umgesetzt werden.<sup>178</sup>

Alle drei Verbindungen weisen mit ihrer Top-Pose einen Substruktur-RMSD von kleiner als 1.5 Å und somit eine AZ1-ähnliche Pose auf. **U11** und **U12** sind zusätzlich mit ihrem elektrophilen Kohlenstoffatom nahe genug an Cys644 platziert, dass eine Reaktion plausibel erscheint. **U13** allerdings kommt innerhalb des Top-Clusters nicht näher als 4.12 Å an das

Schwefelatom. Die Clustergröße nimmt für diese Verbindung ebenfalls deutlich ab.

	Verbindung	Estim. Free Enthalpy, Top-Pose [kcal/mol]	RMSD Top-Pose [Å]	Distanz Top-Pose [Å]	Min. Dist. Top-Cluster [Å]	Größe Top-Cluster
Linear	<b>U11</b> (H)	-10.91	1.31	2.11	2.06	22
	<b>U12</b> (Me)	-10.22	0.95	3.97	2.88	13
	<b>U13</b> (2-Propanol)	-10.86	1.26	3.51	3.03	7
Gewinkelt	<b>U11</b> (H)	-11.01	1.14	2.07	2.04	15
	<b>U12</b> (Me)	-10.52	0.98	3.50	2.21	7
	<b>U13</b> (2-Propanol)	-11.36	1.62	3.29	3.23	3

Tabelle 4.17.: Docking-Ergebnisse der Alkine mit dem WIDOCK-Protokoll. Für die Verbindungen wird die endständige Substitution des Alkins angegeben. Neben der *Estimated Free Enthalpy* wird der Substruktur-RMSD des AZ1-Gerüsts im Vergleich zur Kristallstruktur angegeben. Die angegebenen Distanzen wurden zwischen dem reaktiven Kohlenstoffatom und dem Schwefelatom von Cys644 gemessen.

Für die Verbindungen wurde ebenfalls die WIDOCK-Prozedur angewendet (siehe Tabelle 4.17). Wie zu erwarten trifft die in Kapitel 4.2.4.1 diskutierte Problematik linearer Warheads hier ebenfalls zu, falls das Alkin substituiert ist. Folgerichtig kann nur **U11** das Abstandskriterium von 2.2 Å erfüllen (siehe Abbildung 4.56a), während für die beiden anderen Alkine kaum eine Annäherung beobachtbar ist. Daher wurde die WIDOCK-Prozedur hier ebenfalls mit einem gewinkelten Warhead wiederholt (siehe Abbildung 4.56b). Die geometrische Änderung erlaubt nun auch für **U12**, das elektrophile Kohlenstoff nahe bei Cys644 zu platzieren (siehe Tabelle 4.17). Es sollte allerdings betont werden, dass nur für eine einzige gefundene Pose das Abstandskriterium fast erfüllt werden kann. Für **U13** konnte überhaupt keine Pose gefunden werden, die das Abstandskriterium erfüllt, was aufgrund des sterischen Anspruchs der 2-Propanol-Gruppe wenig verwunderlich ist.

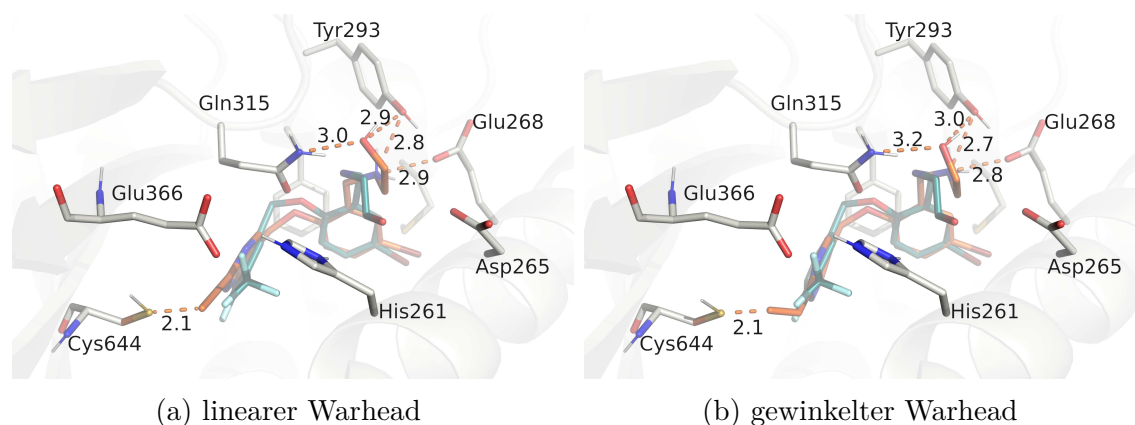


Abbildung 4.56.: WIDOCK-Ergebnisse von **U11** mit unterschiedlichen Warhead-Geometrien. In dunkeltürkis ist die Kristallstruktur von AZ1 (**5**) als Referenz gezeigt. Für beide Geometrien kann das reaktive Atom in der Nähe von Cys644 platziert werden. Die eingezeichneten Abstände sind in Å angegeben und kennzeichnen Wasserstoffbrückenbindungen sowie den Abstand zwischen den reaktiven Atomen. Es werden die polaren Wasserstoffatome nur für das aufbereitete Protein und den gedockten Liganden gezeigt, nicht für die Kristallstruktur-Referenz.

#### 4.2.4.3. Docking-Ergebnisse der Epoxide und Aziridine

Wie in Kapitel 4.2.3 diskutiert, zeichnen sich Epoxide und Aziridine dadurch aus, dass prinzipiell beide Kohlenstoffatome des Dreirings für eine Reaktion zur Verfügung stehen. Die Docking-Ergebnisse können daher helfen, aufzuklären, ob ein Atom günstiger positioniert ist als das andere. Es muss allerdings beachtet werden, dass **U14** und **U15** ein Stereozentrum aufweisen. Für jedes Enantiomer muss ein eigenes Docking durchgeführt werden. Ebenso ist es möglich, dass abhängig vom Stereoisomer ein anderes Kohlenstoffatom besser für die Reaktion platziert wird.

Die AutoDock4-Ergebnisse von (*R*)-**U14**, (*S*)-**U14**, (*R*)-**U15** und (*S*)-**U15** sind in Tabelle 4.18 zu sehen. Innerhalb der Top-Cluster können für beide Mechanismen Posen gefunden werden, die das jeweilige Kohlenstoffatom näher als 4 Å an Cys644 platzieren. Bei den Top-Posen von (*R*)-**U14**, (*S*)-**U14** und (*S*)-**U15** wird das Kohlenstoffatom für den  $S_N1$ -artigen Mechanismus näher an der zu adressierenden Aminosäure platziert, bei der Top-Pose von (*R*)-**U15** das für den  $S_N2$ -Mechanismus. Sehr auffällig ist, dass unabhängig vom postulierten Mechanismus und der Stereochemie das Aziridin knapp 1.5 kcal/mol besser bewertet wird als das Epoxid. Dies kann auf die Wasserstoffbrückenbindungen zurückgeführt werden, die zwischen dem protonierten Aziridin und dem deprotonierten Glu366 ausgebildet werden können. Alle Top-Posen der Aziridine weisen diese Wechselwirkungen auf, während kein

Epoxid diese Interaktion ausbilden kann.

Verbindung	Estim. Free Enthalpy Top-Pose [kcal/mol]	RMSD Top-Pose [Å]	Distanz Top-Pose [Å]		Min. Dist. Top-Clust. [Å]		Größe Top-Cluster
			S <sub>N</sub> 1	S <sub>N</sub> 2	S <sub>N</sub> 1	S <sub>N</sub> 2	
			<i>(R)</i> - <b>U14</b>	-11.03	1.19	3.50	
<i>(S)</i> - <b>U14</b>	-10.96	1.22	3.73	4.50	3.68	3.50	30
<i>(R)</i> - <b>U15</b>	-12.51	1.28	4.24	3.55	3.89	3.46	12
<i>(S)</i> - <b>U15</b>	-12.49	1.22	3.51	4.64	3.27	3.86	30

Tabelle 4.18.: Docking-Ergebnisse des nichtkovalenten Dockings mit AutoDock4. Neben der *Estimated Free Enthalpy* wird der Substruktur-RMSD des AZ1-Gerüsts im Vergleich zur Kristallstruktur angegeben. Die angegebenen Distanzen wurden zwischen dem reaktiven Kohlenstoffatom und dem Schwefelatom von Cys644 gemessen.

Tabelle 4.19 zeigt die WIDOCK-Ergebnisse der Epoxide und Aziridine *(R)*-**U14**, *(S)*-**U14**, *(R)*-**U15** und *(S)*-**U15**. Basierend auf der Diskussion der zweidimensionalen Lennard-Jones-Potenziale in Kapitel 4.2.4.1 erscheint es sinnvoller, nicht beide Kohlenstoffatome gleichzeitig innerhalb desselben Dockings als reaktive Zentren zu parametrisieren, um das chemisch falsche Szenario zu vermeiden, dass das Schwefelatom zu beiden Kohlenstoffatomen zeitgleich die Bindung ausbildet (und folglich auch zweimal energetisch von der Ausbildung einer Bindung profitieren würde, obwohl nur eine Bindung ausgebildet werden kann). Dementsprechend wurde hier für jeden Mechanismus ein Docking durchgeführt, bei dem das jeweilige Kohlenstoffatom zu einem reaktiven Atom modifiziert wurde.

Die Trends, die bereits in Tabelle 4.18 beobachtbar waren, setzen sich in Tabelle 4.19 fort. Die Top-Posen beider Aziridin-Enantiomere bilden die Wasserstoffbrückenbindungen zu Glu366 aus (siehe Abbildung 4.57). Abhängig vom Enantiomer und Mechanismus kann sich dann zusätzlich ein reaktives Kohlenstoffatom nahe genug an Cys644 platzieren, um in den attraktiven Bereich des WIDOCK-Potenzials zu gelangen. So wird z. B. im Rahmen eines hypothetischen S<sub>N</sub>1-artigen Mechanismus die Top-Pose von *(S)*-**U15** etwa 1 kcal/mol besser bewertet als die Top-Pose von *(R)*-**U15**. Es sollte betont werden, dass auch für *(R)*-**U15** prinzipiell innerhalb des Top-Clusters Posen gefunden werden können, die das WIDOCK-Kriterium erfüllen, nur dass diese nicht simultan die Wasserstoffbrückenbindungen ausbilden können und daher schlechter bewertet werden. Die Situation kehrt sich um, wenn von einem S<sub>N</sub>2-Mechanismus ausgegangen wird. Hier ist nur das *(R)*-Enantiomer in der Lage, die Wasserstoffbrückenbindungen zu Glu366 auszubilden und zeitgleich den niedrigen Abstand zu Cys644 einzunehmen, *(S)*-**U15** kann jedoch immer nur eine dieser Bedingungen erfül-

	Verbindung	Estim. Free Enthalpy, Top-Pose [kcal/mol]	RMSD Top-Pose [Å]	Distanz Top-Pose [Å]	Min. Dist. Top-Cluster [Å]	Größe Top-Cluster
$S_N1$	<i>(R)</i> -U14	-11.45	1.19	3.90	2.13	17
	<i>(S)</i> -U14	-11.48	1.08	2.27	2.26	6
	<i>(R)</i> -U15	-12.48	1.37	4.05	2.14	15
	<i>(S)</i> -U15	-13.61	1.28	2.14	2.09	23
$S_N2$	<i>(R)</i> -U14	-11.61	0.87	2.08	2.01	16
	<i>(S)</i> -U14	-11.53	1.18	2.08	1.97	12
	<i>(R)</i> -U15	-13.73	1.30	2.09	2.04	10
	<i>(S)</i> -U15	-12.78	1.19	4.06	2.01	19

Tabelle 4.19.: Ergebnisse des WIDOCK-Protokolls für U14 und U15.

len. Interessanterweise erhalten die beiden günstigen Kombinationen (*(S)*-Enantiomer mit  $S_N1$ -artigem Mechanismus und *(R)*-Enantiomer mit  $S_N2$ -Mechanismus) sehr ähnliche energetische Bewertungen. Diese Ergebnisse legen daher nahe, dass abhängig vom Stereoisomer unterschiedliche Mechanismen für ein Aziridin begünstigt werden würden, da andere Kohlenstoffatome günstig in Relation zu Cys644 positioniert werden. Jedoch ist es kritisch zu betrachten, ob die hier modellierten Ergebnisse direkt übertragbar auf ein Experiment wären, da von identischen Reaktivitäten für die beiden Kohlenstoffatome ausgegangen wurde. So wäre zum Beispiel denkbar, dass auch wenn für das *(S)*-Enantiomer der  $S_N1$ -artige Mechanismus sterisch günstiger wäre, die entsprechende Reaktionsbarriere zu hoch ist und daher nur das  $S_N2$ -Produkt beobachtet werden kann. Der nächste Schritt, um diese Fragestellung weiter zu untersuchen, wäre daher, über quantenmechanische Methoden die Mechanismen möglichst genau zu modellieren, um Aktivierungsenergien zu bestimmen. Diese könnten dann genutzt werden, um die Tiefe des Minimums des modifizierten Lennard-Jones-Potenzials zu parametrisieren (so, wie es im originalen WIDOCK-Protokoll eigentlich vorgesehen war<sup>102</sup>). Aufgrund des Umfangs der Rechnungen, der für die Ermittlung des korrekten Übergangszustandes notwendig wären, wurde im Rahmen dieser Arbeit davon abgesehen.

Anders als die Aziridine besitzen die Epoxide anstelle eines Wasserstoffbrückendonors einen (schwachen) Akzeptor in der Form des Epoxid-Sauerstoffatoms. Dieser ist potentiell in der Lage, eine Interaktion mit der Thiolgruppe des Cys644 auszubilden. Die Docking-Ergebnisse von *(R)*-U14 und *(S)*-U14 spiegeln das wider. Innerhalb der Top-Cluster können neben den Warhead-Orientierungen, die für eine kovalente Wechselwirkung sprechen, auch solche gefunden werden, die eine Wasserstoffbrückenbindung zwischen Epoxid-Sauerstoff und Cys644 zeigen. Für *(R)*-U14 im  $S_N1$ -artigen WIDOCK-Protokoll wird eine derartige Pose sogar als

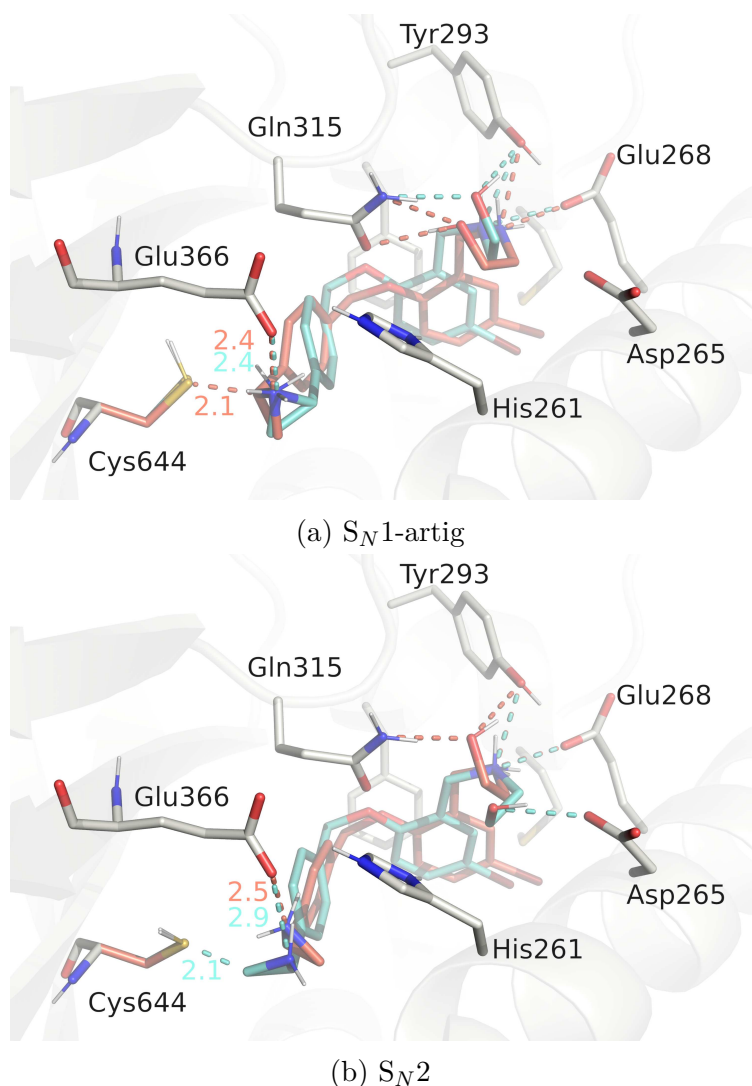


Abbildung 4.57.: Ergebnisse des WIDOCK-Dockings der beiden Aziridin-Enantiomere (*R*)-**U15** (türkis) und (*S*)-**U15** (korallfarben). Die eingezeichneten Abstände sind in Å angegeben und kennzeichnen Wasserstoffbrückenbindungen sowie den Abstand zwischen den reaktiven Atomen. Es werden nur die polaren Wasserstoffatome für das aufbereitete Protein und den gedockten Liganden gezeigt.

Top-Pose bewertet (siehe Abbildung 4.58). Anders als bei den Aziridinen ist es allerdings bei keiner der Kombinationen aus Mechanismus und Stereoisomer möglich, sowohl die kovalente Wechselwirkung als auch die Wasserstoffbrückenbindung auszubilden. Daher kann an dieser Stelle nicht so eindeutig ein Zusammenhang zwischen Stereoisomer und Mechanismus formuliert werden wie für die Aziridine.

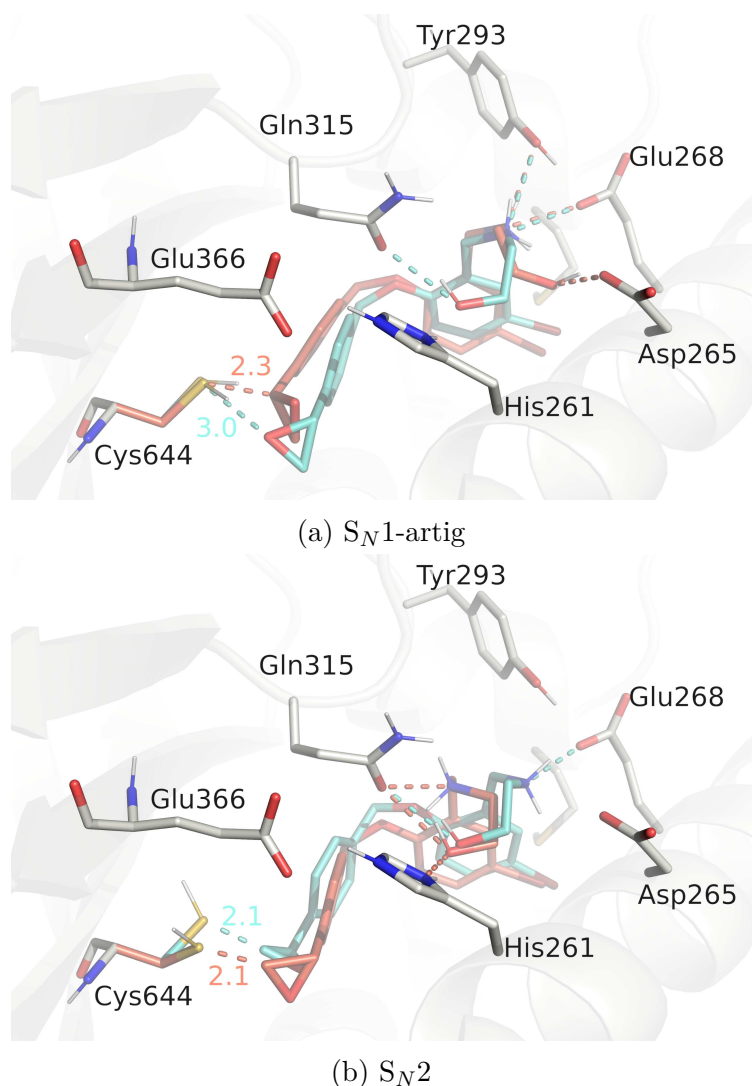


Abbildung 4.58.: Ergebnisse des WIDOCK-Dockings der beiden Epoxid-Enantiomere (*R*)-**U14** (türkis) und (*S*)-**U14** (korallfarben). Die eingezeichneten Abstände sind in Å angegeben und kennzeichnen Wasserstoffbrückenbindungen sowie den Abstand zwischen den reaktiven Atomen. Es werden nur die polaren Wasserstoffatome für das aufbereitete Protein und den gedockten Liganden gezeigt.

#### 4.2.4.4. Docking-Ergebnisse der sterisch anspruchsvollen Warheads

Die Verbindungen **U16**, **U18** und **U17** weisen im Vergleich zu den bisher besprochenen Molekülen große Warheads auf (Michael-System, Vinylsulfon und  $\alpha$ -halogeniertes Keton). Auch für diese sterisch anspruchsvollen Verbindungen können zum Teil noch AZ1-ähnliche nichtkovalente Posen im Rahmen eines unmodifizierten AutoDock4-Dockings gefunden werden, siehe Tabelle 4.20. Lediglich **U17** erfüllt nicht mehr die gesetzten Abstandskriterien.

Die Clustergrößen allerdings nehmen bei allen Verbindungen ab.

Verbindung	Estim. Free Enthalpy, Top-Pose [kcal/mol]	RMSD Top-Pose [Å]	Distanz Top-Pose [Å]	Min. Dist. Top-Cluster [Å]	Größe Top-Cluster
<b>U16</b>	-11.49	1.21	4.61	3.75	13
<b>U17</b>	-11.40	0.86	4.54	4.21	6
<b>U18</b>	-11.51	1.04	3.42	3.42	4

Tabelle 4.20.: Docking-Ergebnisse des nichtkovalenten Dockings für die sterisch großen Warheads. Neben der *Estimated Free Enthalpy* wird der Substruktur-RMSD des AZ1-Gerüsts im Vergleich zur Kristallstruktur angegeben. Die angegebenen Distanzen wurden zwischen dem reaktiven Kohlenstoffatom und dem Schwefelatom von Cys644 gemessen.

Alle drei Verbindungen weisen innerhalb ihres Warheads eine in dieser Docking-Studie drehbare Bindung auf. Dadurch ist das im WIDOCK-Protokoll reaktive Atom flexibel genug, um in der Nähe von Cys644 platziert zu werden, sodass alle diese Verbindungen im Top-Cluster das Abstandskriterium von 2.2 Å erfüllen (siehe Tabelle 4.21). Für **U18** ist allerdings auffällig, dass die Top-Pose einen Abstand von 6.04 Å zwischen reaktivem Kohlenstoffatom und Cys644-Schwefelatom aufweist. Das bedeutet, dass kein energetischer Benefit über das WIDOCK-Potenzial erreicht wird, und trotzdem diese Verbindung besser bewertet wird als andere im selben Cluster, die das WIDOCK-Kriterium erfüllen.

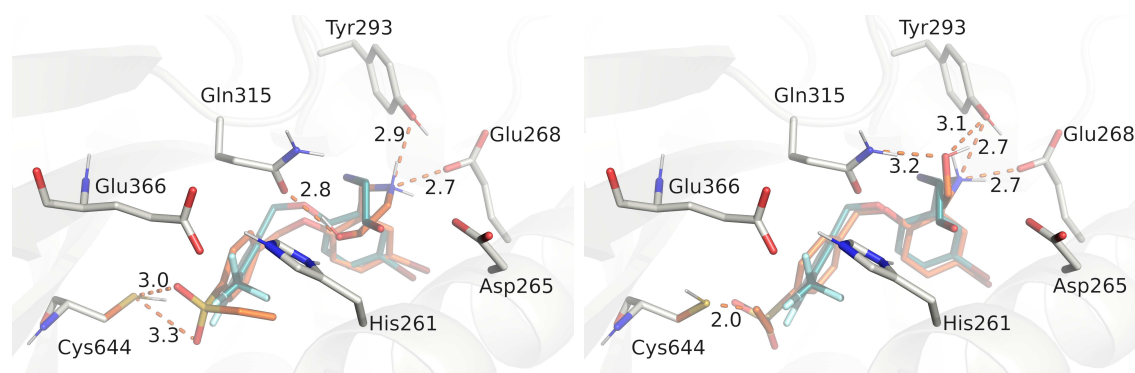
Verbindung	Estim. Free Enthalpy, Top-Pose [kcal/mol]	RMSD Top-Pose [Å]	Distanz Top-Pose [Å]	Min. Dist. Top-Cluster [Å]	Größe Top-Cluster
<b>U16</b>	-12.04	0.77	2.07	2.07	10
<b>U18</b>	-12.93	1.03	6.04	2.05	9
<b>U17</b>	-12.30	1.16	2.09	2.09	4

Tabelle 4.21.: Docking-Ergebnisse des WIDOCK-Protokolls für die sterisch anspruchsvollen Warheads. Neben der *Estimated Free Enthalpy* wird der Substruktur-RMSD des AZ1-Gerüsts im Vergleich zur Kristallstruktur angegeben. Die angegebenen Distanzen wurden zwischen dem reaktiven Kohlenstoffatom und dem Schwefelatom von Cys644 gemessen.

Abbildung 4.59 zeigt die Top-Pose und die Pose mit der minimalen reaktiven Distanz innerhalb des Clusters 1 von **U18**. In der bestbewerteten Pose (Abbildung 4.59a) sind die Sulfonylsauerstoffatome auf Cys644 ausgerichtet, sodass Wasserstoffbrückenbindungen zu dem Thiol ausgebildet werden können. Bei der Pose mit der minimalen reaktiven Distanz (Abbildung 4.59b) liegt der Warhead entgegengesetzt vor, sodass das reaktive Kohlenstoffatom



direkt auf Cys644 deutet. Wie bereits in Kapitel 4.2.4.3 für die Epoxid-Verbindungen gesehen, scheint eine Wasserstoffbrückenbindung energetisch ähnlich günstig zu dem WIDOCK-Potenzial zu sein, sodass beide Orientierungen im Rahmen des Dockings gefunden werden. Da die Warhead-Atome nicht Teil der Substruktur-RMSD-Berechnung sind, werden diese beide Posen demselben Cluster zugeordnet. Prinzipiell unterstreicht diese Beobachtung allerdings, dass an dieser Stelle der Bindetasche genügend Raum vorhanden ist, um verschiedene Orientierungen eines sterisch anspruchsvollen Warheads zu erlauben. Folglich sollten andere, hier nicht explizit gedockte Warheads in *para*-Position ebenfalls sterisch toleriert werden.



(a) Top-Pose mit Wasserstoffbrückenbindungen zu Cys644

(b) Pose mit minimaler reaktiver Distanz im Top-Cluster

Abbildung 4.59.: Ergebnisse des WIDOCK-Dockings von **U18**. Die eingezeichneten Abstände sind in Å angegeben und kennzeichnen Wasserstoffbrückenbindungen sowie den Abstand zwischen den reaktiven Atomen. Die Top-Pose ist in orange, die Kristallstruktur-Referenz in petrol gezeigt. Es werden die polaren Wasserstoffatome nur für das aufbereitete Protein und den gedockten Liganden gezeigt, nicht für die Kristallstruktur-Referenz.

#### 4.2.4.5. DFT-basierte Rechnungen zur Reaktivitätsvorhersage

Es ist davon auszugehen, dass Modifikationen in der Nähe des Warheads nicht nur die nicht-kovalenten Interaktionen des Liganden mit der Bindetasche beeinflussen, sondern auch die Reaktivität des Warheads selbst. Daher wurde im Rahmen dieser Arbeit ein in der Literatur bereits beschriebenes QM-Protokoll adaptiert,<sup>186</sup> um Abschätzungen der Reaktivität von Warheads zu ermöglichen.

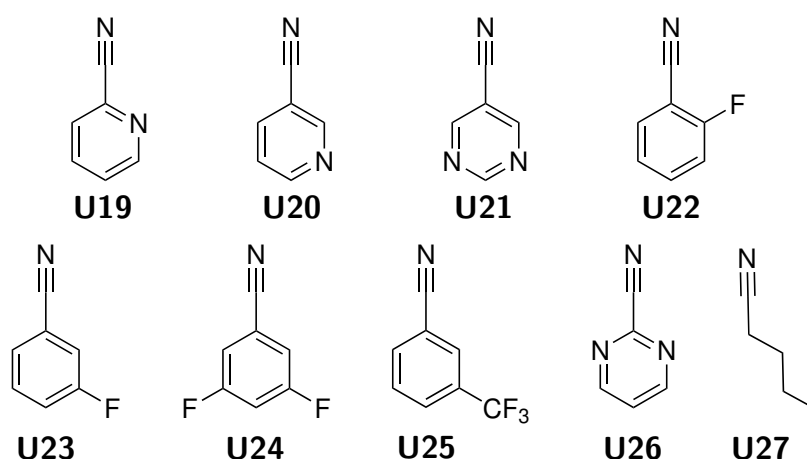


Abbildung 4.60.: Nitril-Warheads. Die hier diskutierten Strukturen orientieren sich dabei an den von Victor Olmos synthetisierbaren Liganden **U4** - **U10**. Zusätzlich sind einige Verbindungen aus dem ursprünglichen Protokoll<sup>186</sup> übernommen.

Dieses Protokoll konzentriert sich auf die Reaktion von Nitrilen mit Cysteinen. Cystein wird hierbei über ein zwitterionisches Cysteamin angenähert, da der Einfluss der Carboxygruppe als konstanter Parameter angenommen wird und so der Rechenaufwand reduziert werden kann. Als Modell-Mechanismus wird eine konzertierte Reaktion mit zyklischem Übergangszustand angenommen, bei dem der Angriff des negativ geladenen Schwefels an das Nitril-Kohlenstoffatom und die Übertragung des Protons der Ammoniumgruppe auf das Nitril-Stickstoffatom etwa zeitgleich erfolgen (siehe Abbildung 4.61). Berteotti *et al.*<sup>186</sup> konnten zeigen, dass die für diesen Mechanismus berechneten Aktivierungsenergien mit experimentell bestimmten Geschwindigkeitskonstanten für die Reaktion von Cystein mit den entsprechenden Nitrilen korrelieren.

Alle Berechnungen wurden mit Gaussian09, Version C.01, auf dem B3LYP/6-311++G\*\* - Level durchgeführt.<sup>122</sup> Alle Rechnungen nutzten das CPCM-Solvensmodell,<sup>187,188</sup> um implizit Wasser als Medium zu modellieren. Das Integral für die DFT-Rechnungen wurde auf *ultrafine* und die Konvergenzkriterien des SCF-Algorithmus auf *tight* gestellt, in Analogie zu

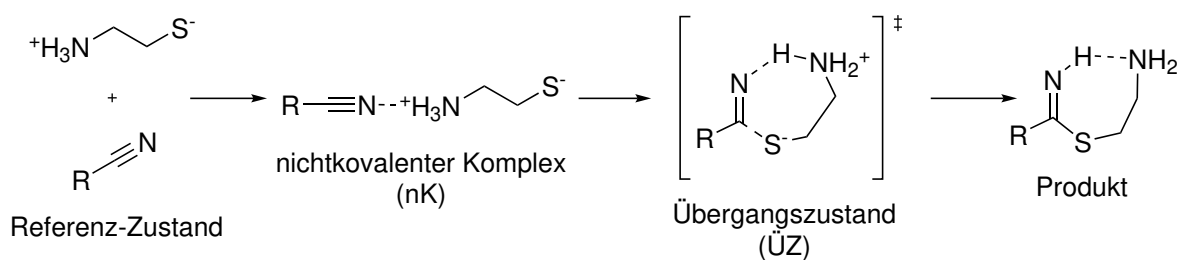


Abbildung 4.61.: Modell-Mechanismus für die Reaktion eines Nitrils mit Cysteamin. Basierend auf diesem Mechanismus können Aktivierungsenergien berechnet werden, die mit experimentell bestimmten Geschwindigkeitskonstanten von Cystein mit den entsprechenden Nitrilen korrelieren.<sup>186</sup>

neueren Gaussian-Versionen. Es sollte betont werden, dass keine Dispersionskorrektur verwendet wurde, um in bestmöglicher Analogie zur ursprünglichen Publikation zu bleiben.<sup>186</sup>

Um den Optimierungsprozess zu erleichtern, wurde für den ersten Schritt der Geometrieoptimierung explizit die Hesse-Matrix des Systems über die Option *Geom=CalcFC* berechnet. Als Startpunkt wurde der von Berteotti *et al.* publizierte Übergangszustand genutzt,<sup>186</sup> der mit GaussView modifiziert wurde, um den in dieser Arbeit diskutierten Molekülen zu entsprechen. Allerdings wurde lediglich das aromatische System in direkter Konjugation zum Nitril abgebildet, um die Rechenzeit zu verringern. Die Natur aller stationären Punkte wurde mit Hilfe von Frequenzberechnungen bestätigt; Energieminima weisen keine, Übergangszustände eine imaginäre Frequenz im berechneten Vibrationsspektrum auf (vgl. Kapitel 3.3.3). Ausgehend von den berechneten Übergangszuständen wurde die Reaktion mit Hilfe von IRC-Berechnungen (siehe Kapitel 3.3.4) in beide Richtungen verfolgt, um sowohl das Produkt als auch die Reaktanden zu erhalten und zu bestätigen, dass der modellierte Übergangszustand die korrekten Minima miteinander verbindet. Da die einzelnen Schritte entlang der Reaktionskoordinate nicht genauer untersucht wurden, wurde zur Vereinfachung der Rechnungen die Option *Recorrect=Never* gewählt. Dadurch sind allerdings kleinere Abweichungen vom idealen Verlauf entlang der Reaktionskoordinate beobachtbar. Der Energieunterschied (ohne Nullpunktsenergie-Korrektur) zwischen Reaktanden und Übergangszustand ergab die Aktivierungsenergie, die Energie der Edukte wurde als Referenz auf 0 kcal/mol gesetzt. In diesem Kapitel werden nur die Aktivierungsenergien basierend auf den SCF-Energien gezeigt. Durch die Frequenzberechnungen können allerdings auch thermodynamische Korrekturen für jeden Zustand berechnet werden, sodass auch die freien Aktivierungs-Enthalpien angegeben werden können. Die entsprechenden Werte sind im Anhang zu finden (siehe Tabelle A.1). Dort sind ebenfalls sowohl die Energiedifferenzen zwischen Referenz-Zustand und Produkt (basie-

rend auf den SCF-Energien) als auch die freien Reaktions-Enthalpien zu finden.

Der IRC-Algorithmus von Gaussian ist nicht dafür ausgelegt, die „wahren“ Energieminima innerhalb einer flachen Potenzialfläche zu identifizieren, weswegen die so erhaltenen Strukturen zusätzlich mit Hilfe einer standardmäßigen Energieminimierung optimiert werden müssen.<sup>25</sup> Für das System aus Reaktanden ist diese Berechnung sehr zeitintensiv, da nur schwache Wechselwirkungen im nichtkovalenten Komplex der beiden Moleküle bestehen und somit auf der sehr flachen Potenzialfläche viele Geometrieoptimierungsschritte erfolgen müssen, bis ein Minimum nach den Gaussian09-Kriterien identifiziert wird (falls ein derartiges Minimum im Rahmen von DFT-Methoden überhaupt existiert, da Standard-DFT-Methoden häufig Dispersionskräfte nicht korrekt und zum Teil repulsiv abbilden<sup>25,129</sup>). Testweise wurde für alle Nitrile diese Geometrieoptimierung bis zu einem Minimum verfolgt (entspricht in Abbildung 4.62 dem Zustand „nK“), auch wenn dafür mehrere Neustarts der Rechnung notwendig waren, da das Limit an Optimierungsschritten mehrfach überstiegen wurde.

Berteotti *et al.* umgingen dieses Problem, indem sie die Distanz zwischen den beiden reaktiven Atomen, die direkt am Ende der IRC-Berechnung vorlag, in der nachfolgenden Minimierung einfroren. Dadurch sind deutlich weniger Orientierungen des Cysteamins in Relation zum Nitril möglich, was das Optimierungsproblem deutlich vereinfacht. Eine derartige Einschränkung der Geometrieoptimierung ist allerdings artifizuell und kann ungünstige Geometrien des Warheads hervorrufen, wie z. B. nichtlineare Nitrile. Derartig gespannte Systeme als Referenzpunkt würden die berechneten Aktivierungsenergien deutlich geringer erscheinen lassen, als sie tatsächlich sind.

Im Rahmen dieser Arbeit daher wurde ein dritter Weg gewählt, um die Energie der Edukte als Referenz zu berechnen. Die beiden Reaktanden wurden nicht in einer gemeinsamen Geometrieoptimierung minimiert, sondern in getrennten und anschließend die Summe der beiden Energien betrachtet, siehe Abbildung 4.61, „Referenz-Zustand“. Dies erlaubt deutlich schneller ablaufende Optimierungen, die zeitgleich einen physikalisch relevanten und nachvollziehbaren Referenzpunkt darstellen (nämlich beide Moleküle bei unendlicher Entfernung).

Die Ergebnisse für die heteroaromatischen Systeme sind in Abbildung 4.62a, für die fluorierten Systeme in Abbildung 4.62b zusammengefasst. Zusätzlich ist ein Ausschnitt konzentriert auf den Übergangszustand gezeigt, siehe jeweils Abbildung 4.62c und Abbildung 4.62d. Qualitativ deckt sich das Ergebnis der Rechnungen mit dem, was die „chemische Intuition“ nahelegt: Die Pyrimidin-Nitrile **U21** und **U26** mit dem elektronenärmeren aromatischen System weisen niedrigere Aktivierungsenergien auf als die Pyridin-Varianten **U19** und **U20**, während das Alkyl-Nitril **U27** die höchste Aktivierungsenergie zeigt. Eine *ortho*-Positionierung

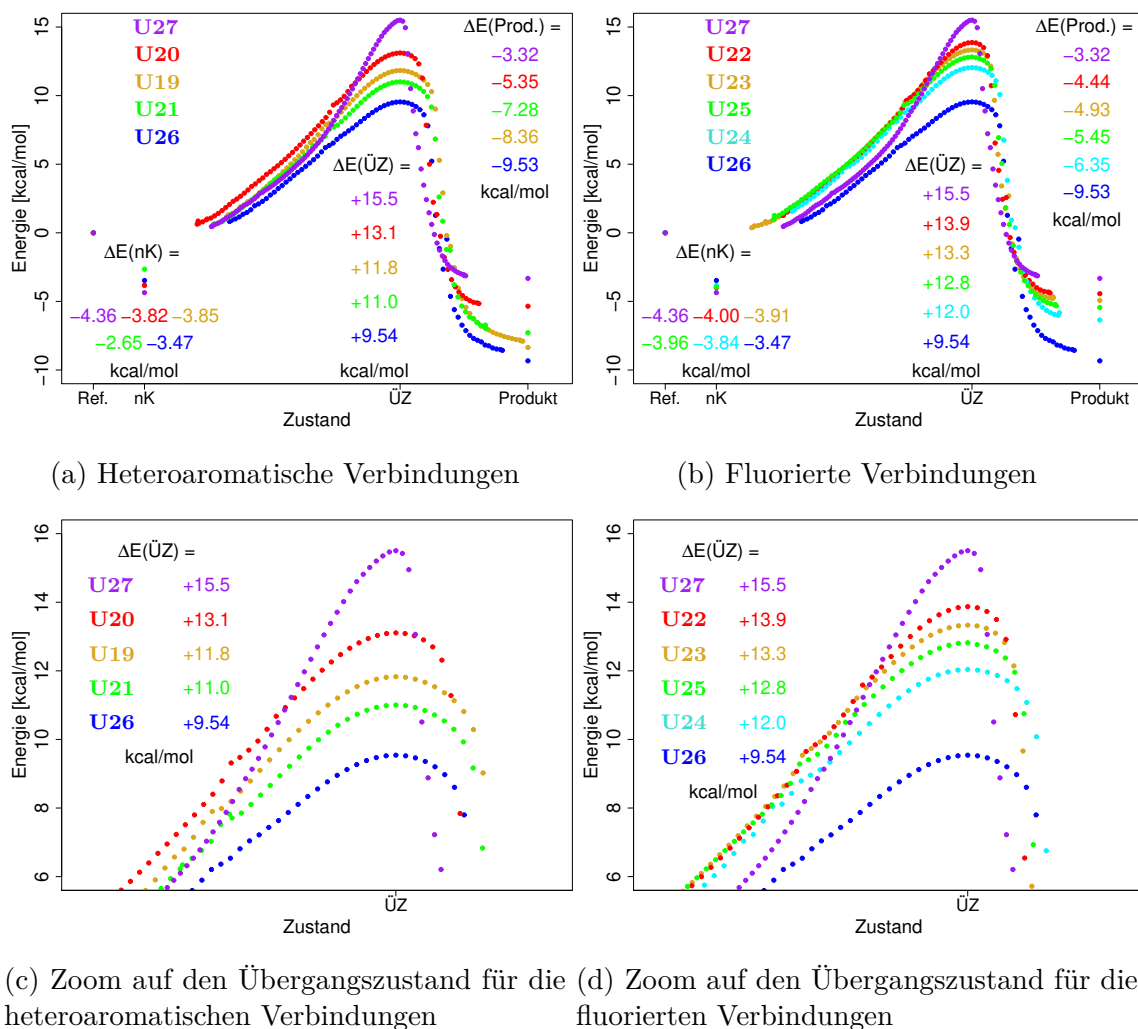


Abbildung 4.62.: Übersicht über die berechneten Energien für die unterschiedlichen Zustände entlang des Modell-Mechanismus (siehe 4.61).

des Heteroatoms (z. B. **U19**) scheint ebenfalls günstiger als eine *meta*-Positionierung (z. B. **U20**) (in Relation zum Nitril). Die Fluorierung scheint allgemein weniger ausschlaggebend zu sein als die Verwendung eines Heteroatoms innerhalb des Aromaten. Auffällig ist, dass für den nichtkovalenten Komplex aus Nitril und Cysteamin für alle Verbindungen vergleichbare, negative Energien berechnet wurden (etwa -4 kcal/mol). **U21** und **U26** weichen am deutlichsten hiervon ab. Für diese wird in Abbildung 4.63 exemplarisch der nichtkovalente Komplex als 3D-Modell gezeigt, in Abbildung 4.61 ist der nichtkovalente Komplex schematisch als Strukturformel bereits dargestellt. Für **U26** wird der nichtkovalente Komplex über eine lineare Wasserstoffbrückenbindung zwischen dem Nitril und der Ammoniumgruppe des Cysteamins definiert (siehe Abbildung 4.63a). Dies ist ebenfalls bei den anderen Verbindun-

gen der Fall, mit Ausnahme von **U21**, bei der der Winkel der Wasserstoffbrückenbindung deutlich von  $180^\circ$  abweicht (siehe Abbildung 4.63b); stattdessen wird das negativ geladene Schwefelatom in Nachbarschaft eines aromatischen Wasserstoffatoms platziert.

Die in Abbildung 4.62 gezeigte Stabilisierung von ca. 4 kcal/mol der nichtkovalenten Komplexe ist demnach auf eine ausgebildete Wasserstoffbrückenbindung zurückzuführen und ist vergleichbar mit Literaturwerten für die Stärke einer Wasserstoffbrückenbindung.<sup>97,135</sup> **U26** bildet hierbei die schwächste Bindung aus, da hier das Nitril im Vergleich zu den anderen Verbindungen am elektronenärmsten ist und daher den schlechtesten Akzeptor in dieser Reihe darstellt (aus demselben Grund zeigt es die niedrigste Aktivierungsenergie). Für **U21** wurde im Rahmen der Optimierung ein anderes lokales Minimum gefunden, das eine schwächere Wasserstoffbrückenbindung aufweist.

Würde dieser nichtkovalente Komplex nun als Referenz genutzt werden, würde die Aktivierungsenergie implizit zusätzlich den Bruch einer Wasserstoffbrückenbindung beinhalten, der abhängig vom Molekül und abhängig vom gefundenen lokalen Minimum unterschiedlich viel Energie benötigen würde, was die Vergleichbarkeit der Rechnungen eingrenzt. Dies unterstreicht, warum es sinnvoller ist, den Referenzpunkt der Energieberechnungen anhand von separierten Geometrieoptimierungen festzumachen.

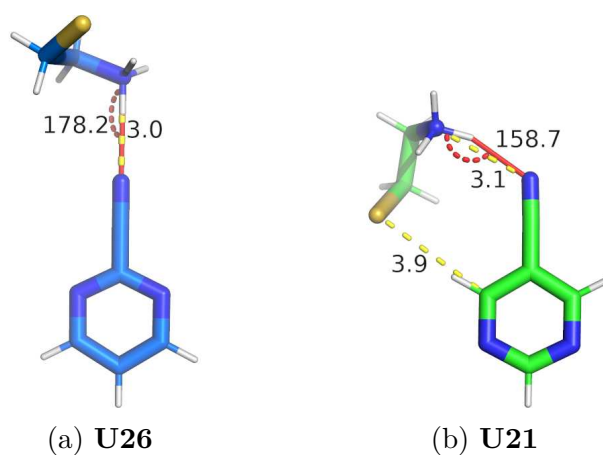


Abbildung 4.63.: Nichtkovalente Anlagerungskomplexe von zwei ausgewählten Verbindungen. Winkel der Wasserstoffbrückenbindungen wurden zwischen den Heteroatomen mit dem Wasserstoffatom als Scheitel gemessen und in Grad angegeben, während die Abstände zwischen Heteroatomen gemessen wurden und in Å angegeben sind.

Auffällig ist ebenfalls, dass die Energiedifferenz zwischen Produkt und Referenz-Zustand oder dem Komplex relativ gering ist. Dies ist qualitativ mit der experimentell bestätigten Reversibilität der Reaktion in Übereinstimmung. Im Falle von **U27** scheint es energetisch günstiger, eine Wasserstoffbrückenbindung auszubilden als die Reaktion zu durchlaufen. Die-

ser Umstand unterstreicht, dass eine unvorhergesehene Wasserstoffbrückenbindung ausreicht, um die Thermodynamik zugunsten der Edukte zu beeinflussen und die Reaktion zu bremsen, insbesondere bei schwachen Elektrophilen (so, wie es die Analyse der in Kapitel 4.1.3.3 diskutierten Kristallstrukturen nahelegt).

Ein analoges Vorgehen wurde für die Alkin-Varianten der in Abbildung 4.60 gezeigten Warheads entwickelt (siehe Abbildung 4.64). Alkine sind bekanntermaßen deutlich weniger reaktiv als Nitrile, reagieren dafür allerdings irreversibel und sind, zumindest theoretisch, zusätzlich funktionalisierbar.

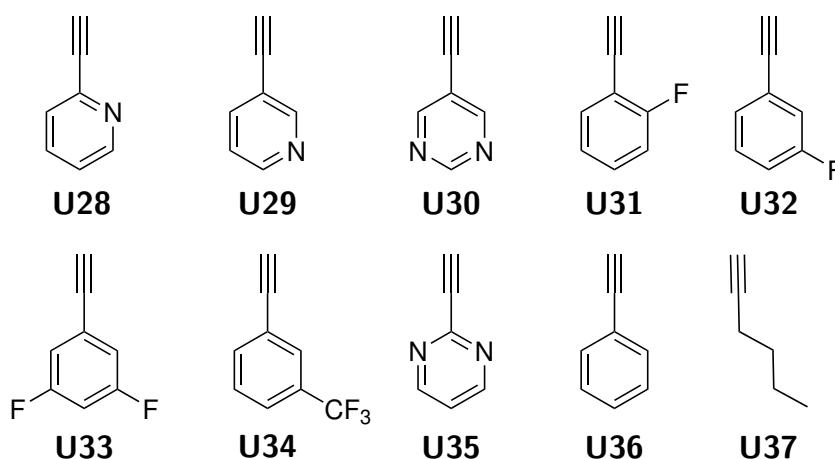


Abbildung 4.64.: Alkin-Warheads. Die Strukturen sind analog zu 4.60, mit einer zusätzlichen Referenzverbindung.

In der Literatur wird diskutiert, wie die Reaktion zwischen Alkin und Cystein mechanistisch abläuft.<sup>17,189</sup> Als Modell wird hier ein zweistufiger Mechanismus angenommen, bei dem ein negativ geladenes Schwefelatom (hier erneut repräsentiert durch zwitterionisches Cysteamin) am endständigen Kohlenstoffatom angreift, sodass die intermediär gebildete negative Ladung in Konjugation zum aromatischen System steht und stabilisiert wird (siehe Abbildung 4.65). Der zweite Schritt in diesem Modell wäre die Übertragung eines Protons auf das formal negativ geladene Kohlenstoffatom. Dieser Schritt sollte, im Vergleich zum ersten Schritt, eine deutlich kleinere Aktivierungsenergie benötigen (falls überhaupt), weswegen sich hier auf den Angriff des Schwefelatoms als geschwindigkeitsbestimmenden Schritt konzentriert wird.

Es sollte der Modellcharakter dieser Reaktion betont werden. Das hier vorgestellte Modell eines zweischrittigen Mechanismus ist in sich konsistent und erlaubt einen relativen Vergleich der Aktivierungsenergien der verschiedenen Alkine untereinander, ist allerdings nicht (wie bei den Nitrilen) durch Korrelation mit experimentell bestimmten Geschwindigkeitskonstanten

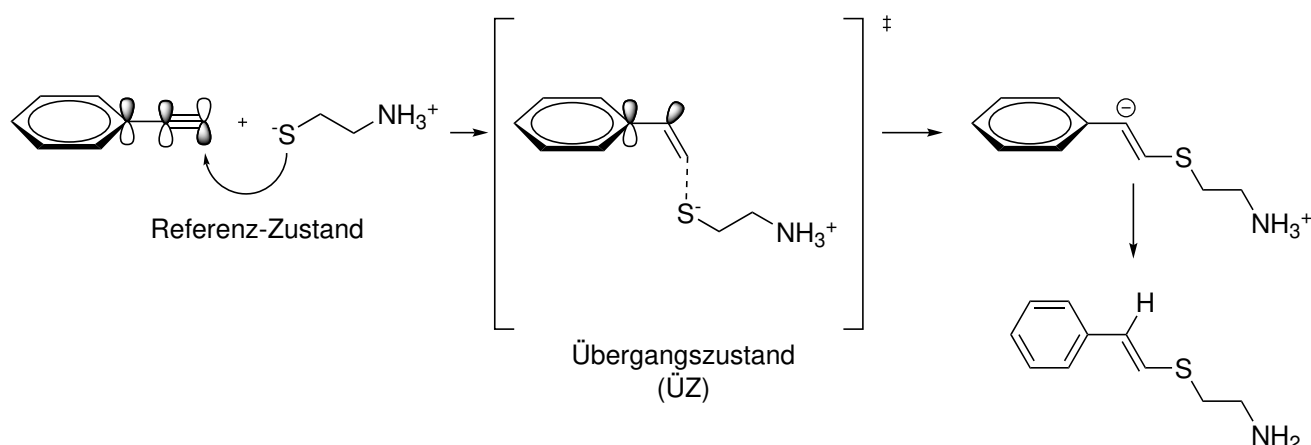
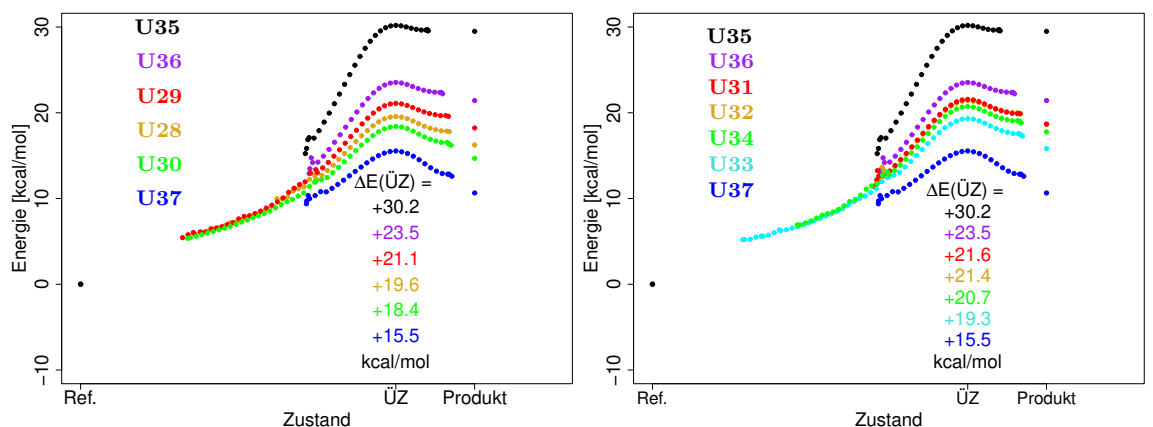


Abbildung 4.65.: Modell-Mechanismus für die Reaktion eines Alkins mit Cysteamin. In diesem Modell wird intermediär eine negative Ladung generiert, die in Konjugation zum aromatischen System steht. In einem hier nicht betrachteten zweiten Schritt erfolgt eine Protonierung des entsprechenden Kohlenstoffatoms.

gedeckt. Es ist daher vorstellbar, dass andere Mechanismen für Vorhersagen besser geeignet wären. Zum Beispiel wäre als alternativer Mechanismus die konzertierte Übertragung des Protons bei zeitgleichem Angriff des Schwefels denkbar, sodass nicht energetisch ungünstigere Intermediate gebildet werden müssen. Dies könnte entweder über eine ringförmige Konformation ermöglicht werden (vergleichbar zu der Reaktion der Nitrile zuvor), sodass die Ammoniumgruppe bereits korrekt positioniert ist, um das Proton zu übertragen, oder ein zusätzlicher Protonendonator wie Wasser müsste sich im System befinden, um das Alkin abzusättigen (was bei einer in Wasser stattfindenden Reaktion plausibel wäre). Versuche, derartige Übergangszustände zu optimieren, sind bisher erfolglos gewesen, da keine Konvergenz innerhalb der Optimierung eintrat. Im Falle des zyklischen Konstrukts traten Wechselwirkungen der Ammoniumgruppe mit z. B. den Heteroatomen des Aromaten ein, die den Übergangszustand verformten, sodass keine Übertragung des Protons mehr möglich war. Bei den Modellen, die ein oder mehrere Wassermoleküle beinhalteten, war die hohe Beweglichkeit des nur schwach assoziierten Wassers problematisch, sodass häufig kein Übergangszustand, der den Default-Standards von Gaussian09 entspricht,<sup>122</sup> identifiziert werden konnte. Dies bedeutet nicht, dass derartige Übergangszustände prinzipiell nicht existieren, sondern dass es mit dieser Methode hier nicht möglich war, sie zu beschreiben. Die Verwendung einer Dispersionskorrektur könnte z. B. helfen, die zuvor beschriebenen Übergangszustände zusätzlich zu stabilisieren und die Rechnungen zur Konvergenz zu bringen.

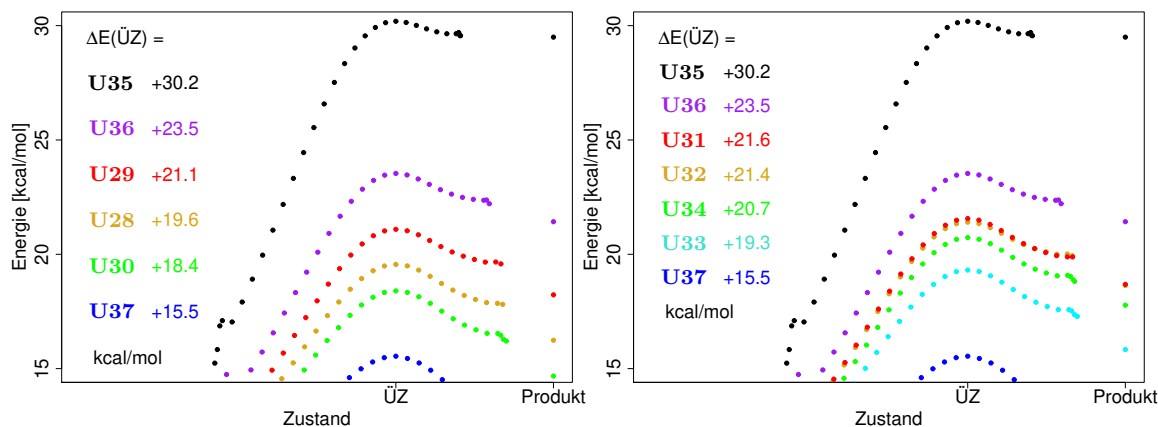
Die Ergebnisse sind in Abbildung 4.66 zu sehen. Es sind dieselben Trends und Abstufungen





(a) Heteroaromatische Alkin-Verbindungen

(b) Fluorierte Alkin-Verbindungen



(c) Zoom auf den Übergangszustand für die heteroaromatischen Alkin-Verbindungen

(d) Zoom auf den Übergangszustand für die fluorierten Alkin-Verbindungen

Abbildung 4.66.: Übersicht über die berechneten Energien für die unterschiedlichen Zustände entlang des Modell-Mechanismus für die Alkine (siehe Abbildung 4.65).

wie bei den Nitrilen erkennbar, jedoch sind die berechneten Aktivierungsenergien für diesen (theoretischen) Übergangszustand deutlich höher als bei den Nitrilen. Selbst das reaktivste Alkin besitzt eine höhere Aktivierungsenergie als das unreaktivste Nitril (vergleiche Abbildung 4.66 mit Abbildung 4.62). Dies ist wenig überraschend, da Alkine typischerweise als bioorthogonal betrachtet werden, d. h. dass sie für ihre Inertheit in biologischen Systemen bekannt sind. Dazu in Übereinstimmung hat das Alkin **U35** eine Aktivierungsenergie von 30.5 kcal/mol, was unter Raumtemperatur-Bedingungen nicht erreichbar wäre.

Nichtsdestoweniger war es für Wood *et al.*<sup>17</sup> möglich, nach genügend langer Inkubationszeit einen signifikanten Anteil an Alkylierung eines katalytisch inaktiven Cysteins eines *epidermal growth factor receptors*, kurz EGFRs, mit einem heteroaromatischen Alkin zu beobachten.

Diese Reaktion konnte sogar katalysiert werden, indem eine basische Funktion in das Alkin-Gerüst inkorporiert wurde. Es erscheint daher plausibel, dass ähnliche Versuche für USP28 ebenfalls funktionieren können. Mit dem hier erarbeiteten Protokoll werden nun energetische Vergleiche zwischen unterschiedlichen Derivaten der Grundstruktur möglich (ob z. B. Substitutionen am aromatischen Ring, die günstige nichtkovalente Wechselwirkungen ausbilden, es wert sind, die Aktivierungsenergie zu erhöhen), was das weitere Design von kovalenten Inhibitoren erleichtern sollte.

#### 4.2.4.6. Experimentelle Ergebnisse und Ausblick auf mögliche weitere Vorgehensweisen

Auch wenn die Ergebnisse in den vorhergehenden *in-silico*-Studien erfolversprechend sind, so ist am Ende dennoch ein Experiment nötig, um die aufgestellten Modelle zu überprüfen. Verbindung **U4** wurde freundlicherweise von Victor Olmos synthetisiert, während der dazugehörige  $IC_{50}$ -Wert der USP28-Inhibition von Jonathan Patzke bestimmt wurde. Für letzteres wurden 20 nM des Proteins vorgelegt und unterschiedlich lange mit variierenden Inhibitorkonzentrationen (von 5 nM bis 100  $\mu$ M) inkubiert. Als Substrat lag mit Rhodamin110 markiertes Ubiquitin in hohem Überschuss vor.

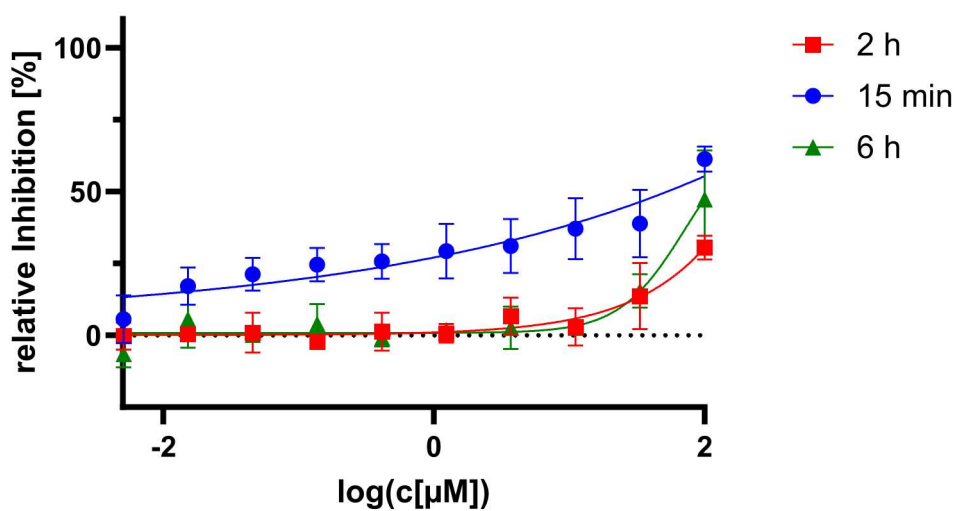


Abbildung 4.67.: Inhibition von USP28 durch **U4**. Experiment und Abbildung wurden freundlicherweise von Jonathan Patzke zur Verfügung gestellt.

Die Ergebnisse sind in Abbildung 4.67 zusammengefasst, die freundlicherweise von Jonathan Patzke erstellt wurde. Bei kurzen Inkubationszeiten, wie etwa 15 min, wird eine Inhibition

von 50 % erst bei 100  $\mu\text{M}$  erreicht. Für einen vollständigen Verlauf der Kurve müssten somit Konzentrationen außerhalb des Screening-Bereichs verwendet werden, sodass eine exakte Bestimmung des  $\text{IC}_{50}$ -Werts nicht möglich ist. Die ähnliche literaturbekannte Verbindung AZ3, **U38**, (siehe Abbildung 4.68), die zusammen mit AZ1 in einem Screening von AstraZeneca gefunden wurde, weist einen  $K_i$  von 59.6  $\mu\text{M}$  auf.<sup>70</sup> **U4** scheint daher prinzipiell schlechter zu binden.

Bei längeren Inhibitionszeiten wird außerdem keine Verbesserung der Inhibition beobachtet, wie man es typischerweise bei langsam bindenden, kovalenten Inhibitoren sieht. Dies könnte bedeuten, dass die Geschwindigkeitskonstante für die kovalente Reaktion  $k_1$  zu klein ist, um selbst bei 6 h einen Effekt zu zeigen.

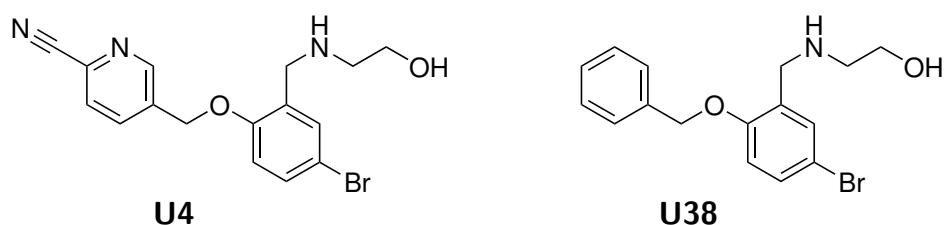


Abbildung 4.68.: **U4** im Vergleich zu **U38**.

Die Ergebnisse, wie sie in Kapitel 4.2.4.5 diskutiert wurden, geben eine klare Vorgehensweise vor, wie **U4** modifiziert werden kann, um dieser Problematik entgegenzuwirken: Die für **U4** assoziierte Aktivierungsenergie  $E_{A1}$  könnte um etwa 2.3 kcal/mol gesenkt werden, wenn stattdessen ein anderes heteroaromatisches System wie **U26** genutzt wird. Diese energetische Verbesserung steigert sich auf 2.8 kcal/mol, wenn thermodynamische Korrekturen berücksichtigt werden (siehe Anhang, Tabelle A.1). Der Arrhenius-Gleichung folgend (bzw. genauer: der Interpretation der Arrhenius-Gleichung basierend auf der Theorie des Übergangszustandes nach Eyring) sollte die neue Geschwindigkeitskonstante für die kovalente Reaktion  $k_2$  dann im Vergleich zur vorherigen Geschwindigkeitskonstante  $k_1$

$$k_2 = A \cdot e^{-\frac{E_{A2}}{RT}} = A \cdot e^{-\frac{E_{A1} - 2.8 \frac{\text{kcal}}{\text{mol}}}{RT}} = A \cdot e^{-\frac{E_{A1}}{RT}} \cdot e^{\frac{2.8 \frac{\text{kcal}}{\text{mol}}}{8.314 \frac{\text{J}}{\text{mol}\cdot\text{K}} \cdot 298.15\text{K}}} \approx k_1 \cdot 113 \quad (4.4)$$

etwa um einen Faktor von 100 größer werden (siehe Gleichung 4.4), vorausgesetzt, dass der präexponentielle Faktor  $A$  bei beiden Liganden (ungefähr) gleich ist. Alternativ wurde bereits in Kapitel 4.2.4.4 gezeigt, dass prinzipiell ein breites Spektrum an Warheads sterisch möglich ist, auch solche, die deutlich reaktiver sind als Nitrile (z. B. Epoxide und Michael-Systeme). Möglicherweise ist es aber auch hilfreich, einen neuen Blickwinkel einzunehmen und die fehlende Reaktivität nicht auf der Seite des Liganden, sondern auf der Seite des Proteins zu

bekämpfen. Cys644 hat kein Histidin in seiner direkten Umgebung. Folglich sollte dessen Reaktivität deutlich niedriger sein als die eines katalytisch aktiven Cysteins; möglicherweise zu niedrig, um ohne Aktivierung mit einem Warhead zu reagieren. Daher wäre eine mögliche Vorgehensweise, zusätzlich zu einem Warhead eine Base in direkter Nachbarschaft von Cys644 zu platzieren, um so das Proton des Thiols zu abstrahieren und die Nucleophilie zu erhöhen. Diese Vorgehensweise hat bereits bei anderen Projekten Erfolge gezeigt: So konnten z. B. Alkine, welche aufgrund ihrer geringen Reaktivität häufig als bioorthogonal bezeichnet werden, durch eine zusätzliche basische Funktion als kovalente Inhibitoren von EGFR-Kinasen genutzt werden, die ein unaktiviertes Cystein adressieren.<sup>17</sup>

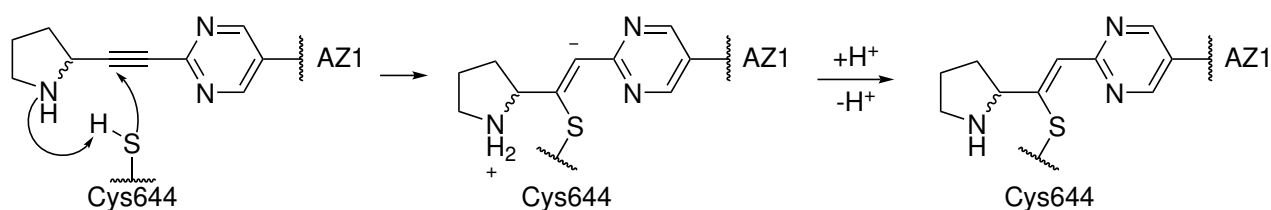


Abbildung 4.69.: Hypothetisches Alkin mit intrinsisch vorhandener Base.

Abbildung 4.69 zeigt schematisch, wie ein entsprechendes AZ1-basiertes Alkin aussehen und wie mechanistisch die Reaktion ablaufen könnte. Der hier gezeigte Pyrrolidin-Ring, der als weiterer Substituent eingeführt wurde, wäre hierbei prinzipiell durch andere basische Einheiten ersetzbar. Dieser Ring würde in der Bindetasche ungefähr dort platziert werden, wo sich in Vismodegib der sulfonierte Phenylring befinden würde, d. h. es wäre denkbar, sowohl die Vismodegib- als auch die AZ1-Bindestelle zu adressieren und an Cys644 kovalent zu binden. Eine andere mögliche Interpretation der Inhibitionsdaten ist, dass das Standard-Protokoll zur IC<sub>50</sub>-Bestimmung für reversibel-kovalente Inhibitoren weiter angepasst werden muss. So wurde zum Beispiel die Inkubation mit Inhibitor bei etwa 0 °C durchgeführt, da es üblich ist, Proteine wann immer möglich gekühlt zu lagern, um Denaturierung zu vermeiden. Senkung der Temperatur verlangsamt allerdings auch die Reaktion des Inhibitors mit dem Enzym, was kontraproduktiv ist. Idealerweise müsste also, bevor weitere Untersuchungen zu kovalenten Inhibitoren stattfinden, getestet werden, bis zu welcher Temperatur das Protein eine ausreichend lange Lebenszeit aufweist, um mehrere Stunden inkubiert zu werden. Dann könnte anschließend U4 erneut bei höheren Temperaturen vermessen werden, die förderlicher für eine Reaktion sind.

Zusammengefasst kann also gesagt werden, dass auch wenn die ersten experimentellen Versuche noch keine ausreichende Inhibition von USP28 zeigen, der kovalente Ansatz des Ligand-Designs durchaus großes Verbesserungspotenzial zeigt.

#### 4.2.5. Untersuchung der Möglichkeiten zur Entwicklung eines verlinkten Inhibitors aus AZ1 und Vismodegib

Auch wenn der „*Merging Approach*“ keinen Bezug zu kovalenten Inhibitoren aufweist, soll der Vollständigkeit halber hier kurz gezeigt werden, wie AZ1 und Vismodegib (**5** und **6**, siehe Abbildung 4.43) zu einer Verbindung verschmolzen werden könnten. Hierfür wurden Dockings mit dem zuvor beschriebenen Docking-Protokoll für AutoDock3 durchgeführt.

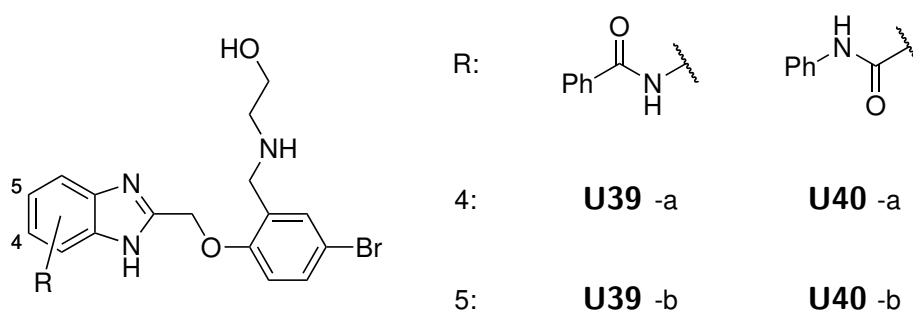


Abbildung 4.70.: Strukturvorschläge, die AZ1- und Vismodegib-Struktur miteinander verschmelzen.

Um die Komplexität des Docking-Problems möglichst gering zu halten, wurde der Vismodegib-Teil des Moleküls stark reduziert. Publierte *Structure-Activity-Relationship*-Studien (SAR-Studien) legen nahe, dass Variationen an diesem Part nur geringe Änderungen in der Affinität bewirken.<sup>80</sup> Die naheliegendste Verschmelzung von AZ1 und Vismodegib würde über ein Naphthalin-System erfolgen, da die Überlagerung der jeweiligen Kristallstrukturen einen Phenylring von AZ1 direkt neben einen Phenylring von Vismodegib platziert (siehe 4.46). Ein derartiger Naphthalin-Baustein würde jedoch die Lipophilie des Moleküls deutlich steigern, was prinzipiell während der Ligandoptimierung vermieden werden sollte.<sup>190</sup> Daher wurde ein Benzimidazol verwendet, welches einen ähnlichen Raumbedarf aufweist. Aus diesen Überlegungen wurde die Struktur **U39** abgeleitet. Bisher unveröffentlichte Mutationsstudien der Arbeitsgruppe Kisker legen nahe, dass die Wasserstoffbrückenbindung zwischen Glu366 und dem Vismodegib-Amid einen essentiellen Beitrag zur Bindungsaffinität darstellt. Daher wurde hier zusätzlich getestet, ob auch für ein invertiertes Amid (**U40**) sinnvolle Docking-Lösungen generiert werden können und ob diese besser oder schlechter bewertet werden. Die in Abbildung 4.70 dargestellten Substitutionspositionen 4 und 5 wären (vermutlich) in einem experimentellen Kontext äquivalent, da der Imidazolring über Protonierung und Deprotonierung tautomerisieren kann. Im Rahmen dieser Docking-Studie müssen aber die entsprechenden Tautomere getrennt betrachtet werden (**U39-a/b** bzw. **U40-a/b**).

Um die hier vorgeschlagenen Moleküle besser bewerten zu können, wurden verschiedene Metriken herangezogen: Da größere Moleküle *per se* größere Scores erhalten, da sie mehr Atome zum Wechselwirken besitzen, aber nicht zwangsläufig effizienter binden, wird hier ein *per-atom-score* (PAS) berechnet, welcher ein Quotient aus der in AutoDock3 berechneten *Estimated Free Binding Energy*  $\Delta G_{estim.}$  und der Anzahl an Schweratomen  $HA$  ist (siehe Gleichung 4.5). Konzeptionell ist der PAS mit der experimentellen *ligand efficiency* verwandt, für die typischerweise Werte von 0.3 kcal/mol erwartet werden.

$$PAS = \frac{-\Delta G_{estim.}}{HA} \quad (4.5)$$

$$LLE = pK_{i,estim.} - SlogP \quad (4.6)$$

$$\Delta G_{estim.} = RT \ln K_{i,estim.} \quad (4.7)$$

Da, wie zuvor bereits angedeutet, die Affinität des Moleküls gesteigert werden soll, ohne dessen Löslichkeit zu verschlechtern, wird für alle Moleküle ebenfalls eine *lipophilic ligand efficiency* (LLE) berechnet (siehe Gleichung 4.6).<sup>190,191</sup> Diese wiegt den negativen dekadischen Logarithmus der Inhibitionskonstante  $K_i$  gegen den Octanol-Wasser-Verteilungskoeffizienten  $\log P$  auf. Letzterer wurde mit Hilfe des MOE-Deskriptors *SlogP* abgeschätzt, während die Inhibitionskonstante über die Gleichung 4.7 aus der berechneten freien Bindungsenthalpie der entsprechenden Docking-Lösung bestimmt werden kann. In der Literatur werden für die LLE Werte von 5 bis 7 empfohlen.<sup>190</sup> Es wird zusätzlich ein Substruktur-RMSD-Wert im Vergleich zur AZ1-Kristallstruktur berechnet. Eine analoge Berechnung im Vergleich zur Vismodegib-Struktur ist wenig sinnvoll, da die hier betrachteten Liganden nur einen geringen Teil des ursprünglichen Vismodegibs aufweisen.

Verbindung	$\Delta G_{estim.}$ [kcal/mol]	$pK_{i,estim.}$	RMSD [Å]	PAS [kcal/mol]	LLE
<b>5</b>	-10.45	7.66	0.90	-0.42	3.58
<b>U39-a</b>	-13.06	9.58	1.13	-0.41	5.83
<b>U39-b</b>	-12.63	9.26	1.85	-0.39	5.51
<b>U40-a</b>	-13.14	9.63	0.85	-0.41	5.89
<b>U40-b</b>	-11.46	8.40	1.49	-0.36	4.66

Tabelle 4.22.: Docking-Ergebnisse für die Top-Posen von **U39** und **U40**. Auswahl der Top-Pose erfolgte anhand der berechneten freien Bindungsenthalpie  $\Delta G_{estim.}$

Tabelle 4.22 fasst die Docking-Ergebnisse für die beiden Verbindungen zusammen. Alle Top-Posen zeigen PAS- und LLE-Werte in akzeptablen Bereichen auf und reproduzieren für den AZ1-Part die Kristallstruktur mit einem RMSD-Wert kleiner 2 Å.

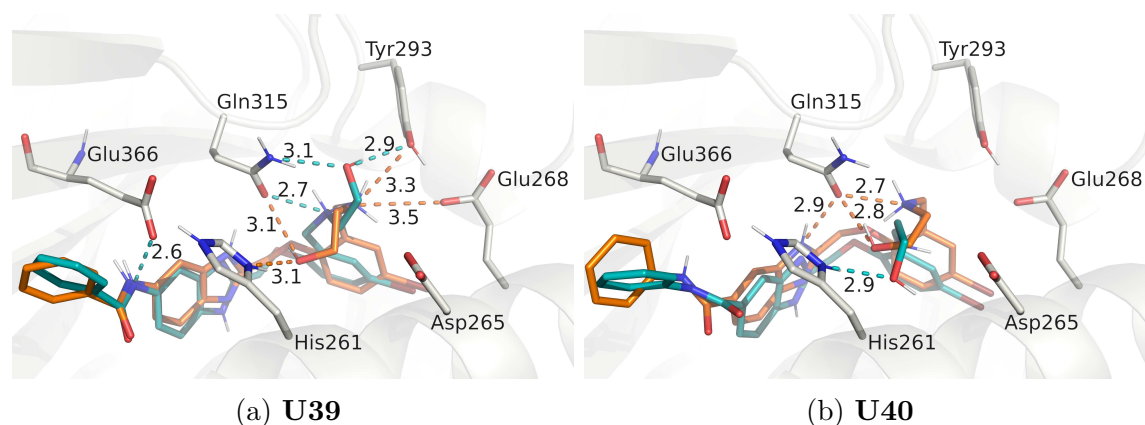


Abbildung 4.71.: Die erhaltenen Docking-Lösungen für die Verbindungen **U39** und **U40**. In orange sind deren a-Varianten, in türkis deren b-Varianten gezeigt. Es werden nur polare Protonen gezeigt.

**U39-a** wird leicht besser bewertet als **U39-b** (-13.06 vs. -12.63 kcal/mol). Das polare Benzimidazolproton ist bei **U39-a** auf das Carbonyl-Sauerstoffatom der Gln315-Seitenkette ausgerichtet (siehe Abbildung 4.71a), auch wenn der Abstand mit 3.6 Å zu groß ist, um von einer wirklichen Wasserstoffbrückenbindung auszugehen. Zusätzlich ist der RMSD-Wert im Vergleich zur Kristallstruktur der Top-**U39-a**-Pose niedriger als bei der Top-**U39-b**-Pose. Beide dieser Beobachtungen könnten für den geringen energetischen Unterschied verantwortlich sein.

Die Top-Pose von **U40-b** wird etwas schlechter bewertet als die übrigen Posen. Dies kann über die ungünstige Orientierung des protonierten Amins erklärt werden (siehe Abbildung 4.71b). Für dieses wurde im Rahmen des Suchalgorithmus keine Wasserstoffbrückenbindungen ausgebildet, weswegen diese Top-Pose schlechter bewertet wird im Vergleich zu den anderen, die derartige Wechselwirkungen aufweisen.

Im Vergleich zu einem Redocking von AZ1 (**5**) weisen die Verbindungen lediglich einen leicht schlechteren PAS auf, allerdings eine deutlich verbesserte LLE, da die Trifluormethylgruppe entfernt und der Phenylring durch das besser lösliche Benzimidazol ausgetauscht wurde. Da die hier gezeigten Verbindungen bereits akzeptable Parameter aufweisen, ist es durchaus vorstellbar, dass es mit einem ausführlicherem Design durchaus möglich wäre, hoch aktive USP28/25-Liganden über einen Merging-Ansatz zu entwickeln. Die hier durchgeführte Demonstration soll aber an dieser Stelle genügen.

## 4.2.6. MD-Simulation von USP25- und USP28-Konstrukten

Im Rahmen eines Nebenprojekts wurde ebenfalls die Beweglichkeit der beiden Proteine USP25 und USP28 durch MD-Simulationen miteinander verglichen. Für die Kooperationspartner Prof. Dr. Caroline Kisker, Dr. Florian Sauer und Jonathan Patzke waren hierbei Mutanten von Interesse, bei denen das zentrale Glutamat der Bindestelle durch Alanin ersetzt wurde, da diese Varianten schneller andere Proteine ubiquitinierten. Falls signifikante Unterschiede in der Dynamik zwischen USP25 und USP28 existieren, wäre dies ebenfalls ein möglicher Ansatzpunkt, um Selektivität zu erreichen.

### 4.2.6.1. Methodische Details der MD-Simulationen

Zwei Kristallstrukturen wurden als Startpunkt für die MD-Simulationen genutzt: PDB 6H4J für USP25 und eine (zum Zeitpunkt der Entstehung der Arbeit) unveröffentlichte Kristallstruktur von monomeren USP28- $\Delta$ UCID. Da für USP25 bisher keine monomere USP25- $\Delta$ UCID-Struktur existiert, wurde ein Homologie-Modell basierend auf 6H4J mit MOE konstruiert.<sup>148</sup> Hierfür wurden die Aminosäuren von Arg407 bis Met586 entfernt und stattdessen ein Motiv aus sechs Aminosäuren (abwechselnd Glycin und Serin) eingefügt. 100 mögliche Konformationen des Protein-Rückgrats wurden *de novo* für diesen neuen Loop His406-Gly-Ser-Gly-Ser-Gly-Ser-Ile587 generiert. Für die fünf besten Modelle wurden Seitenketten hinzugefügt. Anschließend wurde das energetisch am besten bewertete Modell ausgewählt. Andere, kleinere Loops, die nicht in 6H4J oder der USP28- $\Delta$ UCID-Struktur aufgelöst waren, wurden automatisch durch das *Structure Preparation Module* von MOE hinzugefügt. Für die Modelle der Mutanten wurde dann Glu373 (USP25) bzw. Glu366 (USP28) zu Alanin umgewandelt. Protonierungszustände und die Orientierung der Wasserstoffatome wurden mit Hilfe von MOEs *protonate3D*-Tool bestimmt.<sup>149</sup> Dadurch kamen Standard-Protonierungszustände für alle Aminosäuren zustande, d. h. alle Histidine sind einfach protoniert, alle Aspartate und Glutamate deprotoniert und alle Lysine positiv geladen. Es sollte betont werden, dass hierbei eine einfache Protonierung von His268/261 (USP25/28) vorliegt. Aufgrund der zahlreichen sauren Aminosäuren in der näheren Umgebung dieses Histidins wäre es ebenfalls plausibel, hier auch bei pH=7 von einer doppelten Protonierung auszugehen. In einigen Kristallstrukturen deutet die Seitenkette von Glu373/366 direkt auf den Imidazol-Ring von His268/261, da der pH-Wert bei der Kristallisation sauer eingestellt war. Der Einfachheit halber wird sich in dieser Arbeit auf einen Protonierungszustand konzentriert.

Die erhaltenen Konstrukte wurden zuerst mit Hilfe von *pmemd* von Amber22 für 2000 Schritte mit einem impliziten Solvensmodell minimiert, um sie an das Kraftfeld ff14SB anzupas-



sen.<sup>108,151</sup> Anschließend wurden die Proteine in einer Box aus TIP3P-Wassermolekülen solvatisiert,<sup>111</sup> die in jede Raumrichtung 10 Å von der Proteinoberfläche bis zum Rand misst. Vor dem Hintergrund, dass sowohl USP25 als auch USP28 stark mehrfach negativ geladen sind, wurde besondere Sorgfalt darauf verwendet, die korrekte Anzahl an Natrium- und Chloridionen zu der Box hinzuzufügen, um die Salzkonzentration von 0.15 M widerzuspiegeln und gleichzeitig ein neutrales System zu erhalten. Das von Machado *et al.*<sup>192</sup> vorgeschlagene Protokoll wurde daher befolgt. Für das 16-fach negativ geladene USP25 wurden 66 Natriumionen und 50 Chloridionen hinzugefügt, für das 18-fach negativ geladene USP28 62 Natriumionen und 44 Chloridionen (da USP28 in einer geringfügig kleineren Box solvatisiert ist werden hier weniger Ionen benötigt). Für die mutierten Versionen der Proteine wurde ein Chloridion weniger benötigt. Anschließend wurde die Box für weitere 2000 Schritte minimiert, wobei das Protein festgehalten wurde.

Für die so vollständig solvatisierten Systeme wurden jeweils sechs unabhängige Equilibrierungen und Simulationen durchgeführt. Hierbei wurde immer *pmemd* von Amber22 mit einem Zeitschritt von 2 fs, Langevin-Dynamik, einer Kollisionsfrequenz von 3 ps<sup>-1</sup> und einem Cutoff von 10 Å für die *particle-mesh*-Ewald Summation verwendet.<sup>118</sup>

Für das Equilibrierungsprotokoll wurde ein schrittweises Vorgehen gewählt, vergleichbar zu dem in Kapitel 4.1.3. Zuerst wurde das System von 100 K auf 300 K innerhalb von 500 ps erhitzt, während das Protein mit einem harmonischen Potenzial mit einer Kraftkonstante von 100 kcal/(Å mol) festgehalten wurde. Anschließend wurde das Protein wieder auf 100 K innerhalb von 500 ps abgekühlt, aber die Kraftkonstante zum Fixieren auf 10 kcal/(Å mol) gesenkt. Auf diese Art und Weise wird die Flexibilität des Proteins erhöht, ohne es zeitgleich zu erhitzen. Zuletzt wird erneut auf 300 K innerhalb von 500 ps erhitzt, während die Kraftkonstante auf 1 kcal/(Å mol) gesenkt wird. Alle diese Schritte erfolgten im NVT-Ensemble. Anschließend erfolgte 1 ns an Simulation im NPT-Ensemble (ohne ein harmonisches Potenzial zur Fixierung auf dem Protein), um die Dichte der Box anzupassen. Nach dieser Equilibrierung erfolgten 250 ns Simulation für jedes Replikat, wobei alle 10 ps die Koordinaten der Trajektorie gespeichert wurden. Insgesamt wurde so für jedes System 6 x 250 ns = 1.5 µs simuliert.

*cpptraj*<sup>154</sup> wurde zur Auswertung der Trajektorien verwendet. Hierfür wurden die Trajektorien auf das Rückgrat zweier benachbarter β-Faltblätter aligniert (Leu370/363 - Phe377/370 und Tyr650/643 - Gln656/649). Basierend auf diesem Alignment wurde ein hierarchisches agglomeratives Clustering unter Nutzung der RMSD-Werte des Rückgrats der zehn N-terminalen Aminosäuren der α5-Helix durchgeführt.

Zusätzlich zu diesem Clustering wurde eine Hauptkomponentenanalyse (*Principal Compo-*

ment Analysis, PCA) mit dem Package *bio3d* von RStudio durchgeführt.<sup>193</sup> Die zuvor beschriebene Alignierung und Auswahl an Atomen wurde beibehalten. Hierbei ist zu beachten, dass die erhaltenen Hauptkomponenten für den jeweiligen Satz an Trajektorien spezifisch sind und nicht analog zu den Hauptkomponenten anderer Systeme aufgebaut sein müssen. Um daher eine Vergleichbarkeit zwischen den Simulationen zu ermöglichen, wurden die Hauptkomponenten des USP28-Mutanten-Systems genutzt, um ein Koordinatensystem aufzuspannen, in dem die betrachteten Koordinaten der anderen Systeme ausgedrückt wurden.

#### 4.2.6.2. MD-Simulationen der unmutierten USP-Konstrukte

Nachdem die beiden USP-Proteine wie oben beschrieben vorbereitet wurden, wurden MD-Simulationen der unmutierten USPs durchgeführt.

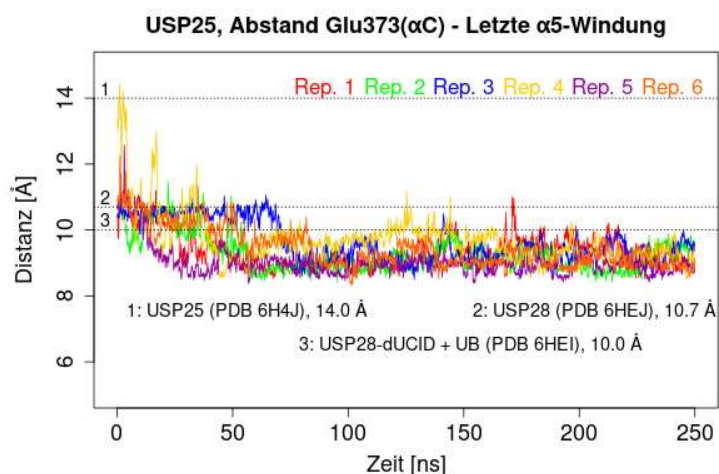
Abbildung 4.72 zeigt die Abstände zwischen dem  $\alpha$ -Kohlenstoffatom von Glu373/366 und dem Massenschwerpunkt des Backbones ( $\alpha$ -Kohlenstoffatom, Amid-Kohlenstoff und Amid-Stickstoff) der letzten Windung der  $\alpha$ 5-Helix, definiert über Val263/256, Ser264/257, Glu265/258 und Phe260/259 im Verlauf der USP25- $\Delta$ UCID und USP28- $\Delta$ UCID-Simulationen. Dieser Abstand wurde gewählt, um als Maß für die Öffnung der Bindetasche zu dienen (siehe Kapitel 1.2.3.1). Zur Orientierung sind ebenfalls Vergleichswerte von verschiedenen Kristallstrukturen angegeben.

Die PDB-Struktur 6H4J beinhaltet USP25 im tetrameren Zustand, d. h. in dieser Kristallstruktur liegt diese Furche weit geöffnet vor, da das AIM von USP25 in der Bindetasche platziert ist. Ausgehend von 6H4J wurde auch das hier simulierte System von monomeren USP25- $\delta$ UCID konstruiert. Wie aus Abbildung 4.72a deutlich wird, werden in der Simulation deutlich geringere Abstände als die in der tetrameren Kristallstruktur beobachteten 14.0 Å gemessen, was bedeutet, dass sich die Bindetasche schließt. Dies ist in allen sechs Replikaten beobachtbar. Die simulierten Abstände sind vergleichbarer zu den Abständen von USP28-Kristallstrukturen, z. B. USP28 mit UCID (6HEJ, 10.7 Å) oder USP28- $\delta$ UCID im Komplex mit Ubiquitin (6HEI, 10.0 Å), wenn auch etwas niedriger.

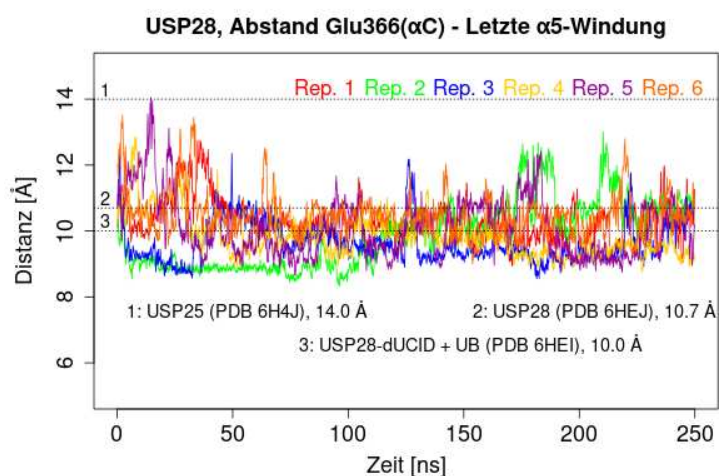
Die analogen Abstände der USP28-Simulationen sind in Abbildung 4.72b zu sehen. Auch hier werden Abstände vergleichbar zu den beiden USP28-Referenzen gefunden, allerdings sind hier große Schwankungen beobachtbar, bis hin zu 14 Å.

Ausgehend von diesen Simulationen kann also vermutet werden, dass monomeres USP25 in Lösung einen ähnlichen Öffnungsgrad aufweist wie monomeres bzw. dimeres USP28, und nicht den geöffneten Zustand aufweist, der in der tetrameren Struktur beobachtet wurde.

Es wurde außerdem vermutet, dass das in 6H4J modellierte  $\pi$ - $\pi$ -Stacking zwischen Phe253 und Phe266 dafür verantwortlich ist, dass USP25 überhaupt so weit geöffnet vorliegt.<sup>78</sup>



(a) USP25



(b) USP28

Abbildung 4.72.: Distanz zwischen Glu373/366 und dem Massenschwerpunkt des Backbones der letzten  $\alpha$ 5-Helixwindung.

Abbildung 4.73 zeigt die Distanz zwischen den Phenylring-Mittelpunkten im Verlauf der Simulationen. Die in der Kristallstruktur 6H4J beobachtete Stacking-Distanz (4.4 Å) ist als Referenz miteingezeichnet. Für mehrere Replikate kann dieselbe Distanz konstant über die gesamte Simulationszeit beobachtet werden, und zeitgleich eine kleine Öffnung der Bindetasche, vergleiche mit Abbildung 4.72a. Für Replik 1 wird ab ca. 150 ns ein leicht erhöhter Abstand beobachtet. In der Simulation ist ab diesem Zeitpunkt ein T-förmiges Stacking zu beobachten, anstelle des planaren Stackings der Kristallstruktur. In Replik 3 ist das Stacking aufgelöst, sodass deutliche Fluktuationen des Abstandes beobachtet werden kön-

nen.

Die Simulationen können die These, dass das  $\pi$ - $\pi$ -Stacking zwischen Phe253 und Phe266 für die Öffnung verantwortlich ist, also nicht unterstützen, da es prinzipiell möglich scheint, das Stacking aufrecht zu halten und die Bindetasche zu schließen.

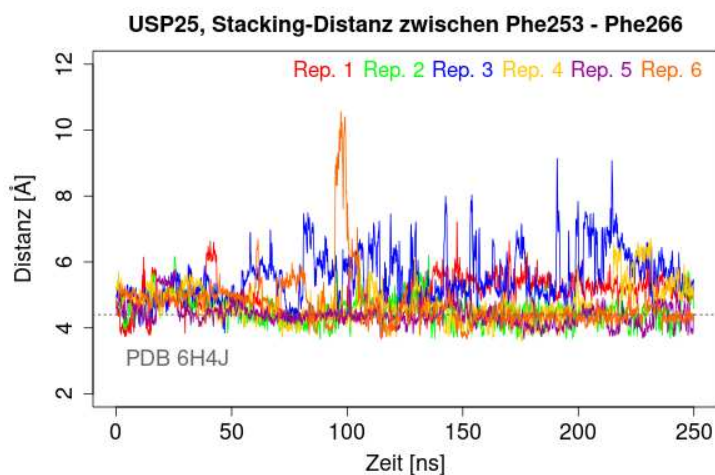


Abbildung 4.73.: Stacking zwischen Phe253 und Phe266 in der USP25-Simulation. Dieses Stacking wird in der Literatur für die Öffnung von USP25 verantwortlich gemacht.<sup>78</sup>

Zusätzlich wurde zur Charakterisierung der Simulationen ein Clustering durchgeführt. Hierfür wurden die Simulationen, wie in Kapitel 4.2.6.1 beschrieben, auf zwei benachbarte  $\beta$ -Faltblätter, die zur Bindetasche gehören, aligniert, und die zehn N-endständigen Aminosäuren der  $\alpha$ 5-Helix genutzt (Ser264/257-Trp273/266). Abbildung 4.74 zeigt das zentrale Frame der Hauptcluster der USP25- und USP28-Simulationen. Für die USP25-Simulationen können 71.7 % der simulierten Frames diesem Cluster zugeordnet werden, dessen Zentroid-Frame in Abbildung 4.74a zu sehen ist. Natriumionen aus dem Solvens koordinieren hier an nahe beieinander befindliche negativ geladene Aminosäuren und bilden ein Netzwerk aus. Für das analoge USP28-Frame ist eine erstaunliche Ähnlichkeit zu PDB 6HEJ (*apo*-USP28) zu beobachten. Hier sind direkte und über Wasser vermittelte Wasserstoffbrücken zwischen den  $\beta$ -Faltblättern und der  $\alpha$ -Helix beobachtbar. 59.2 % der simulierten Frames können diesem Hauptcluster zugeordnet werden.

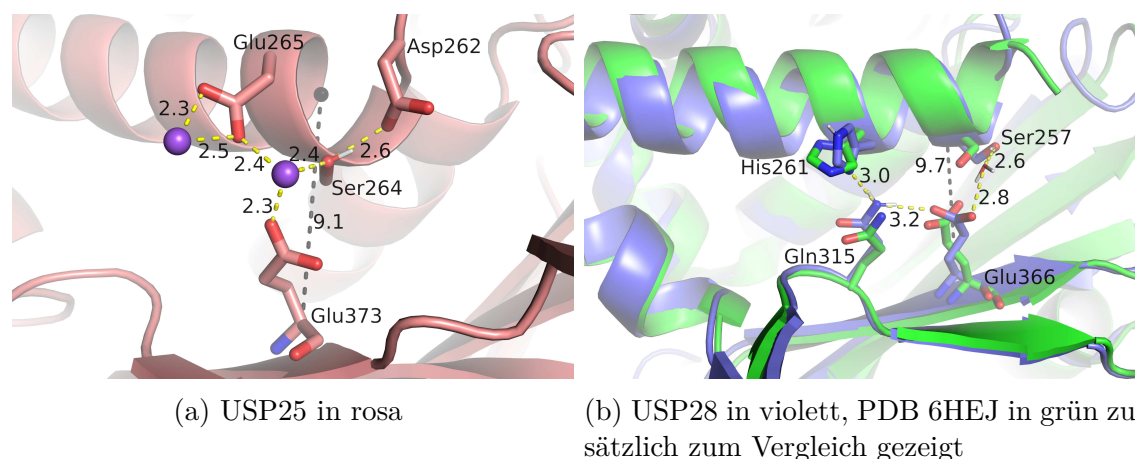


Abbildung 4.74.: Zentroid-Frames der USP- $\delta$ UCID Simulationen. In schwarz ist die Distanz vom  $\alpha$ -Kohlenstoff von Glu373/366 zum Mittelpunkt des Backbones der letzten  $\alpha$ 5-Helix zu sehen.

#### 4.2.6.3. MD-Simulationen der mutierten Varianten

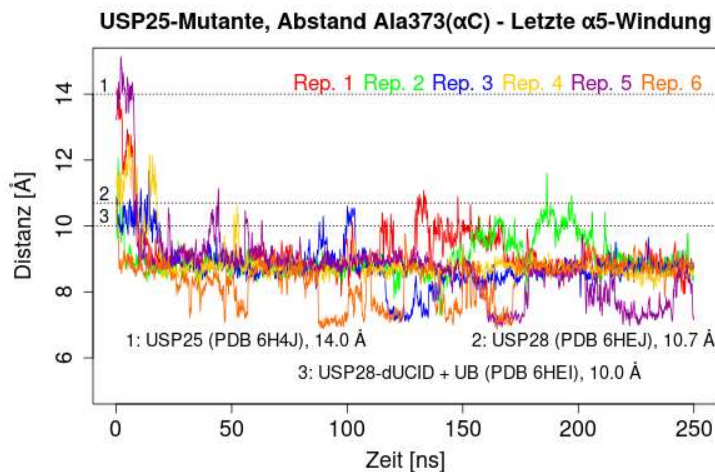
Zusätzlich zu den unmutierten Konstrukten wurden auch mutierte Varianten von USP25- $\Delta$ UCID und USP28- $\Delta$ UCID simuliert. Dabei wurde Glu373/366 zu Alanin umgewandelt, um eine Erklärung dafür zu finden, warum diese Mutante im Experiment eine höhere Aktivität aufweist als unmutiertes USP25/28.

Abbildung 4.75 zeigt den Abstand zwischen dem  $\alpha$ -Kohlenstoffatom von Glu373/366Ala zum Mittelpunkt des Aminosäuren-Rückgrats der letzten Windung der  $\alpha$ 5-Helix. Damit ist Abbildung 4.75 analog zu Abbildung 4.72, allerdings für die Simulationen für die mutierten Systeme. Abbildung 4.76a zeigt die Abstände der unterschiedlichen Systeme miteinander verglichen in Form von Histogrammen.

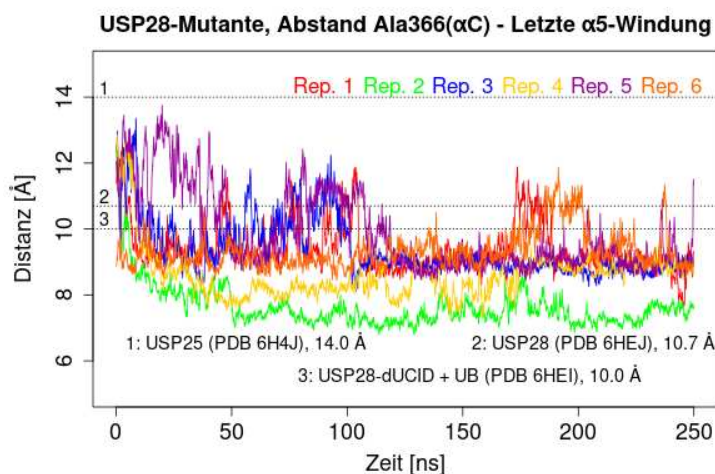
Es wird klar, dass es für die mutierten Systeme möglich ist, kürzere Abstände einzunehmen, als in den Wildtyp-Simulationen beobachtet werden konnte. Dies ist zum Beispiel sehr deutlich für Replik 5 und 6 in Abbildung 4.75a oder für Replik 2 in Abbildung 4.75b zu sehen. Anscheinend haben die Interaktionen zu Glu373/366 dazu beigetragen, die letzte Windung der  $\alpha$ -Helix auf Abstand zu halten. Ohne diese Wechselwirkungen als Abstandshalter kann die Helix nun deutlich näher an die benachbarten  $\beta$ -Faltblätter herankommen.

Problematisch bei der bisherigen Analyse der Simulationen ist, dass sich auf einen einzelnen Abstand konzentriert wurde. Dieser muss nicht zwangsläufig ein gutes Maß für die Öffnung der Bindetasche sein, da er sich auch aufgrund von anderen Bewegungen ändern kann, z. B. durch eine Biegung der  $\alpha$ -Helix parallel zu den  $\beta$ -Faltblättern.

Daher wurde getestet, ob mit Hilfe einer PCA ebenfalls Unterschiede zwischen den mutierten



(a) USP25-Mutante

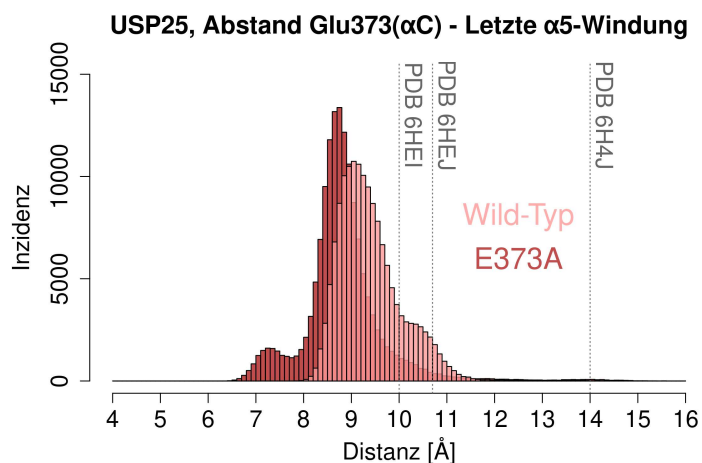


(b) USP28-Mutante

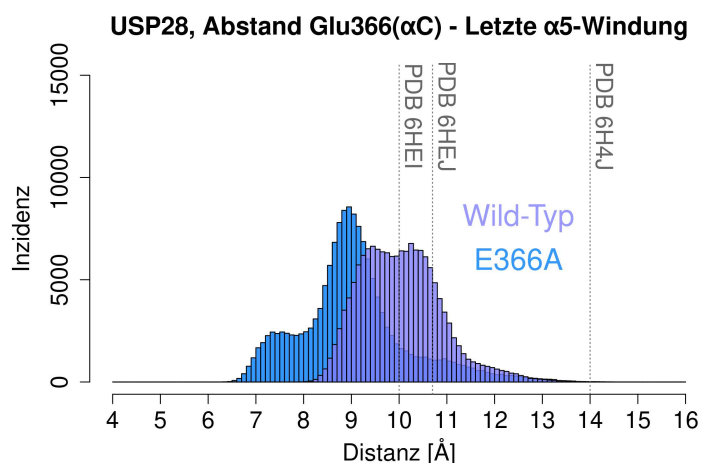
Abbildung 4.75.: Distanz zwischen Ala373/366 und dem Massenschwerpunkt des Backbones der letzten  $\alpha$ 5-Helixwindung innerhalb der Mutanten-Simulationen.

Simulationen feststellbar sind. Wie in Kapitel 4.2.6.1 beschrieben, wurde hierbei die PCA der USP28-Mutanten-Simulation genutzt, um die Hauptkomponenten zu definieren, über die dann anschließend auch die anderen Trajektorien ausgedrückt wurden. Die Ergebnisse sind in Abbildung 4.77 zu sehen.

Es wird deutlich, dass die beiden mutierten Systeme einen zusätzlichen Bereich im PC-Raum einnehmen, der nicht von den unmodifizierten Simulationen erreicht wird. Diese neuen Konformationen haben ähnliche PC1- bzw. PC2-Werte in den beiden mutierten Systemen, d. h. es liegt nahe, dass es sich hierbei um ähnliche Konformationen handelt, die in beiden



(a) USP25-Systeme

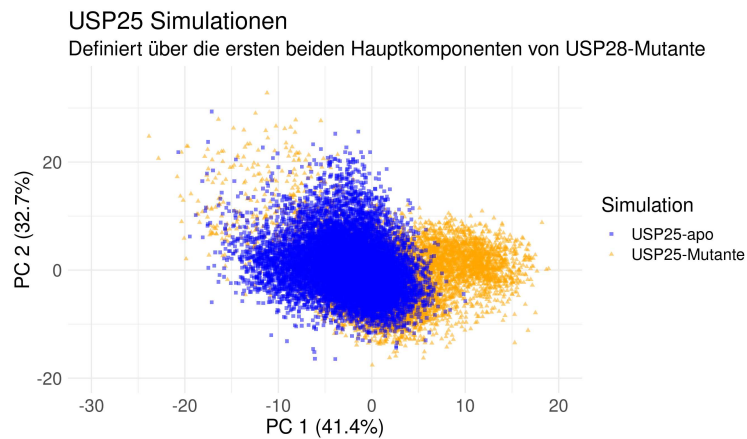


(b) USP28-Systeme

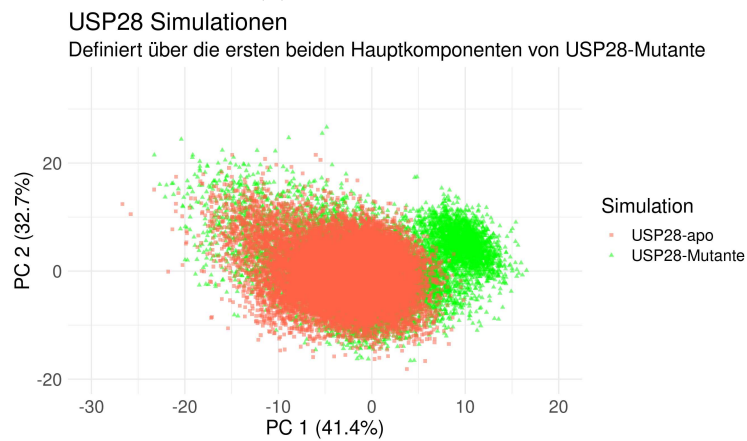
Abbildung 4.76.: Histogramm der Distanz zwischen Glu373/366(Ala) und dem Massenschwerpunkt des Backbones der letzten  $\alpha$ 5-Helixwindung innerhalb der unterschiedlichen USP-Simulationen.

Trajektorien beobachtet werden können.

Die PCA bestätigt somit, dass durch die Glu373/366Ala-Mutation sowohl für USP25 als auch für USP28 ein Shift der Helix möglich ist, der so vorher nicht beobachtbar war. Es liegt nahe zu vermuten, dass diese neue Konformation in irgendeiner Weise mit der erhöhten katalytischen Effizienz der mutierten Systeme assoziiert ist, möglicherweise weil es mit der neu positionierten Helix energetisch günstiger ist, Ubiquitin zu binden. Um diese Frage aber abschließend zu klären, wären weiterführende Simulationen der Systeme mit gebundenem Ubiquitin notwendig, bestenfalls basierend auf entsprechenden Kristallstrukturen, welche



(a) USP25-Systeme



(b) USP28-Systeme

Abbildung 4.77.: Hauptkomponentenanalyse der durchgeführten USP-Simulationen. Alle Trajektorien wurden über die beiden größten Hauptkomponenten der USP28-Mutanten-Simulationen ausgedrückt.

noch nicht existieren. Die hier durchgeführte Analyse muss daher zum jetzigen Zeitpunkt genügen.



### 4.3. ChlaDUB1: Abschätzung der Reaktivität verschiedener Nitril-Warheads

Die Reaktivität eines Warheads muss sorgfältig auf sein Target abgestimmt werden. Eine zu hohe Reaktivität erhöht das Risiko der unselektiven Bindung an off-Targets, während eine zu niedrige Reaktivität eine zu geringe Modifikation der zu adressierenden Aminosäure bedeuten kann. Daher ist es nützlich, theoretische Methoden zur Verfügung zu haben, um die Reaktivität eines kovalenten Liganden abschätzen zu können, um den Synthese-Aufwand auf die vielversprechendsten Verbindungen zu fokussieren.<sup>186, 194</sup>

Thomas Zimmermann aus der Arbeitsgruppe von Prof. Dr. Michael Decker entwarf und synthetisierte eine Reihe an ChlaDUB1-Inhibitoren, die unterschiedliche Nitril-Warheads beinhalteten (siehe Abbildung 4.78). Für diese Warheads soll ein Protokoll aus quantenmechanischen Rechnungen etabliert werden, um die Reaktivität abschätzen zu können.

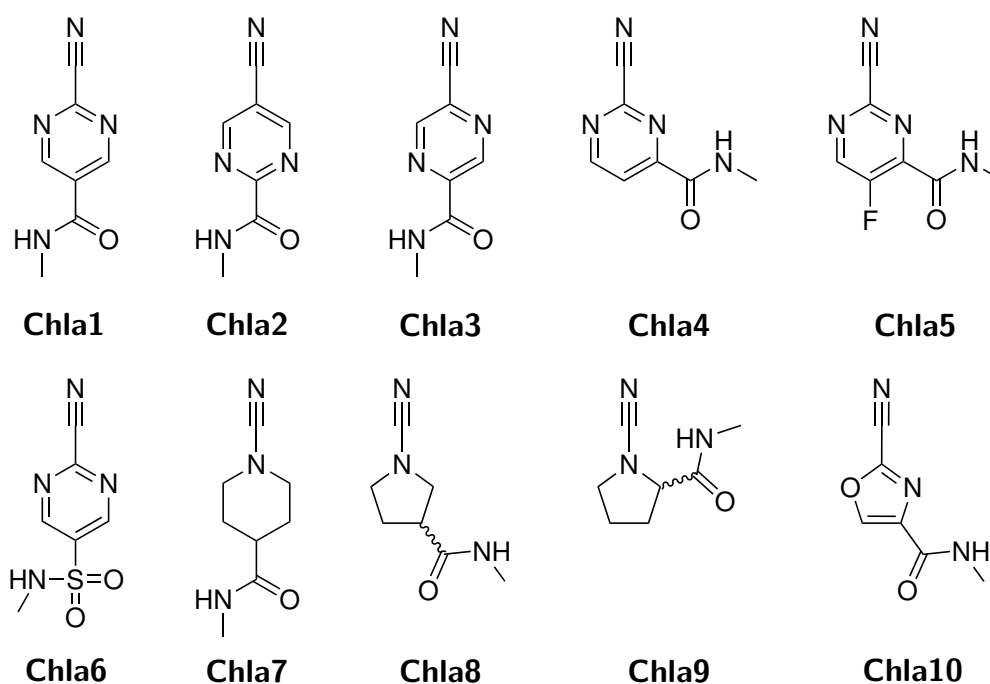


Abbildung 4.78.: Unterschiedliche Warheads zur Inhibition von ChlaDUB1. Stereochemische Information wurde für die Berechnung der Aktivierungsenergie nicht berücksichtigt.

Es lassen sich viele Parallelen innerhalb dieses Datensatzes ziehen: **Chla1**, **Chla2** und **Chla3** weisen ein *para*-Substitutionsmuster auf, allerdings unterschiedliche heteroaromatische Systeme. Auf **Chla3** wird aufgrund einiger besonderer Konformationsmöglichkeiten an späterer Stelle genauer eingegangen (siehe unten). **Chla4** weist das Amid in *meta*-Position auf, anstel-

le der *para*-Position von **Chla1**. **Chla5** ist analog zu **Chla4**, bis auf ein zusätzliches Fluoratom in *para*-Position. **Chla6** ersetzt das Amid von **Chla1** durch ein Sulfonamid. **Chla7** bis **Chla10** variieren den zentralen Heteroaromaten durch ein Piperidin-, ein Pyrrolidin- und ein Oxazolsystem. **Chla8** und **Chla9** sind hierbei Konstitutionsisomere zueinander.

Von Interesse war nun, wie die unterschiedlichen Heterozyklen und Substitutionen die Reaktivität der Verbindungen beeinflussen. Daher wurde die Aktivierungsenergie dieser Nitrile für die Modell-Reaktion aus Kapitel 4.2.4.5 berechnet (siehe Abbildung 4.61). Es wurde dasselbe Protokoll wie in Kapitel 4.2.4.5 angewendet, d. h. es wurden auf dem B3LYP/6-311++G\*\* - Level mit Gaussian09 die Energien für den Referenz- und Übergangszustand berechnet und die Differenz zwischen diesen Energien gebildet.

Die Ergebnisse für **Chla1**, **Chla2**, **Chla4**, **Chla5** und **Chla6** sind in Abbildung 4.79 gezeigt.

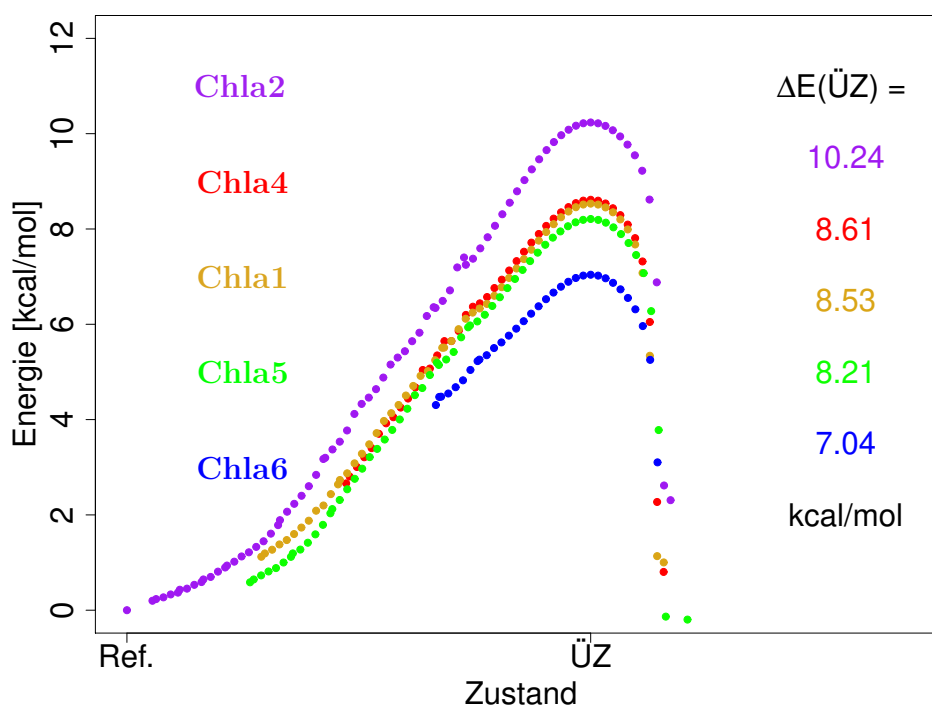


Abbildung 4.79.: Energieprofile der ChlaDUB1-Warheads **Chla1** - **Chla6** in der Region des Übergangszustandes. Auf **Chla3** wird gesondert eingegangen.

Die höchste Aktivierungsenergie dieser fünf Heteroaromaten zeigt **Chla2** mit 10.24 kcal/mol. Die Heteroaromaten **Chla1**, **Chla4** und **Chla5** liegen aufgrund ihrer hohen chemischen Ähnlichkeit sehr eng beieinander. Die Positionierung des Amids in Relation zum Nitril scheint nur einen geringen Effekt auf die notwendige Aktivierungsenergie zu haben (vgl. **Chla4**, 8.61 kcal/mol, mit **Chla1**, 8.53 kcal/mol). Die Einführung eines Fluoratoms senkt die Reakti-

onsbarriere; **Chla5** liegt mit 8.21 kcal/mol etwa 0.4 kcal/mol niedriger als **Chla4**. Wenig überraschend weist das Sulfonamid **Chla6** eine deutlich niedrigere Aktivierungsenergie von 7.04 kcal/mol auf, da ein Sulfon deutlich elektronenziehender ist als ein Amid. Allerdings verursacht das tetraedrisch konfigurierte Schwefelatom dieser Verbindung auch eine deutlich andere Positionierung des Warheads innerhalb der Bindetasche, im Vergleich zu planaren Amiden. Dies schränkt die Vergleichbarkeit zu den anderen hier betrachteten Warheads ein. Wie eingangs erwähnt stellt **Chla3** einen Spezialfall dar. Es können zwei unterschiedliche Konformere den Diederwinkel zwischen Amid und Aromat definiert werden (Szenario 1 und 2 in Abbildung 4.80). Für ein Konformer können sterische Wechselwirkungen zwischen dem Amid-Proton und dem aromatischen Proton entstehen, sodass hier keine vollständige Planarität zwischen den  $\pi$ -Orbitalen des Amids und des Aromaten vorliegt. Daraus sollte theoretisch eine schlechtere Stabilisierung des *lowest unoccupied molecular orbital* (LUMO) resultieren. Die senkrechte Stellung des Amids in Bezug auf das aromatische System stellt eine deutliche energetische Barriere dar, die im Rahmen der Geometrieoptimierung nicht überwunden werden kann. Folglich müssen getrennte Rechnungen für diese beiden Möglichkeiten durchgeführt werden, und anschließend die erhaltenen Ergebnisse miteinander verglichen werden.

Eine ähnliche sterische Problematik ergibt sich auf der Seite des Warheads bei der Reaktion mit dem Thiol (siehe Szenario 3 und 4 in Abbildung 4.80). Das sich ausbildende Thioimidat kann entweder das Schwefelatom auf die Seite des aromatischen Stickstoffatoms drehen (Szenario 3), oder auf die Seite der aromatischen CH-Einheit (Szenario 4). Szenario 4 zeigt erneut eine sterische Wechselwirkung, sodass keine vollständige Planarität zwischen dem Thioimidat und dem Aromaten, und somit keine vollständige Konjugation zwischen allen  $\pi$ -Bindungen erreicht wird, anders als bei Szenario 3. Vor dem Hintergrund, dass viele dieser heteroaromatischen Verbindungen sehr ähnliche Aktivierungsenergien aufweisen, scheint es relevant, diese Problematik genauer zu untersuchen (auch wenn zugegebenermaßen nur **Chla3** betroffen ist, da sich nur hier die beiden *ortho*-Positionen des Nitrils bzw. des Amids unterscheiden.)

Abbildung 4.81 zeigt die unterschiedlichen Aktivierungsenergien, die erhalten werden, abhängig von der verwendeten Kombination an Szenarien bzw. vom Konformer. Die unterschiedlichen Amid-Positionen scheinen nur einen minimalen Effekt auf die Aktivierungsenergie auszuüben (0.01 bzw. 0.02 kcal/mol, in beiden Fällen weniger als 0.2 % bezogen auf die Aktivierungsenergien). Hier lässt sich keine klare Tendenz zwischen den beiden Szenarien 1 und 2 anhand der QM-Rechnungen herleiten. Vor dem Hintergrund, dass die Positionierung des Amids am Aromaten nur einen geringen Einfluss auf die Aktivierungsenergie hat (vergleiche

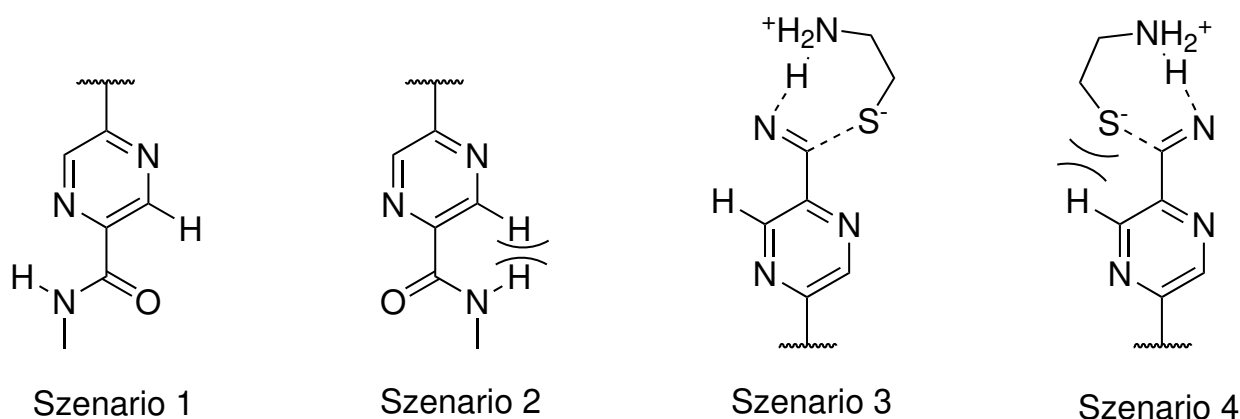


Abbildung 4.80.: Mögliche Konformere für das Pyrazin **Chla3**. Die Szenarien 1 und 2 können kombiniert mit Szenario 3 oder 4 auftreten.

erneut die Aktivierungsenergie von **Chla4** und **Chla1**), ist es einleuchtend, dass ein etwas schlechterer Überlapp der Orbitale nur eine minimale Änderung der Aktivierungsenergie erzeugt.

Deutlich relevanter scheint die Orientierung des sich ausbildenden Thioimidats zu sein. Hier kann ein Unterschied von 0.51 kcal/mol bzw. 0.52 kcal/mol zwischen den beiden möglichen Szenarien beobachtet werden. Diese Differenz entspricht etwa dem Unterschied zwischen **Chla4** und **Chla5**. Das unterstreicht die Relevanz, auf eine sinnvolle Konfiguration des Übergangszustandes zu achten, da sich sonst verfälschte Rangfolgen an Reaktivitäten ergeben könnten. In diesem Datensatz scheint aber **Chla3** selbst mit der bestmöglichen Aktivierungsenergie von 8.97 kcal/mol unreaktiver als **Chla4** zu sein (siehe Abbildung 4.79).

Abbildung 4.82 zeigt die Energie im Verlauf der Reaktion für die übrigen Verbindungen **Chla7** bis **Chla10**. Die nichtaromatischen Cyanamide **Chla7** und **Chla9** weisen hierbei Aktivierungsenergien von ca. 13 kcal/mol auf. Das Cyanamid **Chla8** ergibt eine etwa um 2 kcal/mol geringere Aktivierungsenergie, vermutlich aufgrund der höheren Ringspannung des Fünfrings im Vergleich zum Sechsring in **Chla7** und der geringeren sterischen Hinderung im Vergleich zu **Chla9**. Das Oxazol **Chla10** liegt mit einer berechneten Aktivierungsenergie von 7.57 kcal/mol deutlich niedriger. Aufgrund des heteroaromatischen und dadurch elektronenarmen Fünfrings ist dies wenig verwunderlich.

Abbildung 4.83 zeigt die von Thomas Zimmermann experimentell untersuchten ChlaDUB1-Inhibitoren. **Chla11** verhält sich hierbei zu **Chla12** wie **Chla1** zu **Chla4**; statt einer *para*-Substitution liegt eine *meta*-Substitution vor. **Chla13** weist lediglich eine zusätzliche Methoxygruppe im Vergleich zu **Chla11** auf, welche nicht in Konjugation zum Warhead

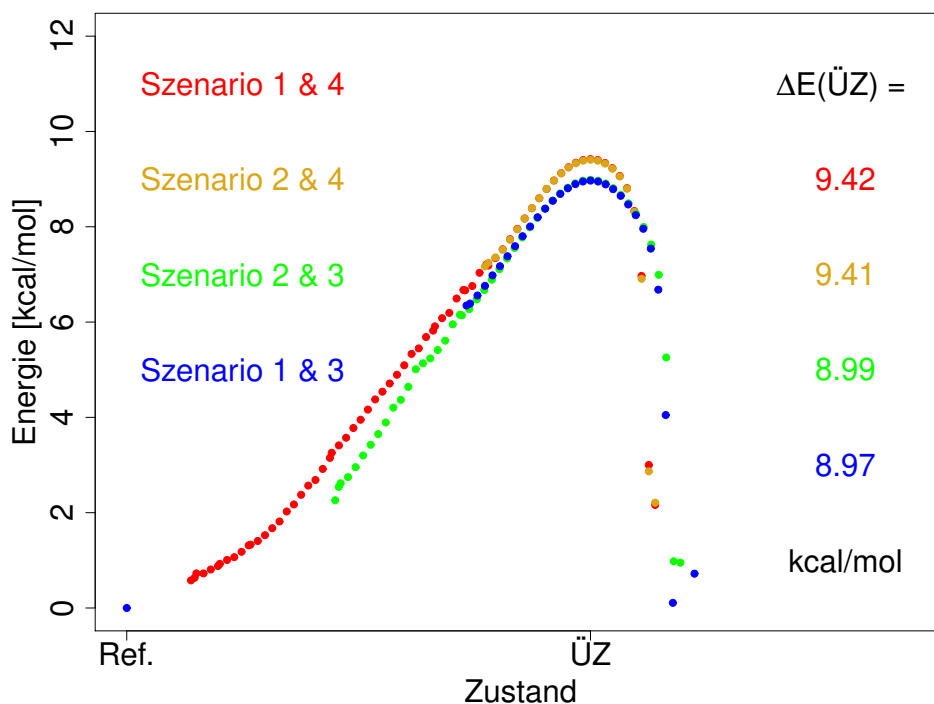


Abbildung 4.81.: Aktivierungsenergien für unterschiedliche Konformere von **Chla3**. Die unterschiedlichen Szenarien sind in Abbildung 4.80 erklärt.

steht und damit die chemischen Eigenschaften des Warheads nicht beeinflusst. Die Reaktivität dieser beiden Verbindungen sollte daher identisch sein, sofern die Methoxygruppe den nichtkovalenten Bindemodus nicht signifikant verändert. **Chla14** besitzt **Chla5** als reaktive Einheit und somit ein zusätzliches Fluoratom in *para*-Position zum Warhead im Vergleich zu **Chla12**.

Tabelle 4.23 fasst die experimentell bestimmten  $k_{inact}$ -Werte für die Verbindungen zusammen und stellt sie den Reaktivitäten der Warheads gegenüber. Zusätzlich zur Aktivierungsenergie  $\Delta E_A$  ist in dieser Tabelle auch die berechnete, freie Aktivierungsenthalpie  $\Delta G_A$  aufgeführt, die aus  $\Delta E_A$  unter Zuhilfenahme thermodynamischer Korrekturen abgeleitet werden kann.<sup>25</sup> (Für die übrigen Verbindungen sind die entsprechenden Werte im Anhang A.3.1 zu finden.) Typischerweise ist es aufgrund der komplexen kinetischen Zusammenhänge, die bei der kovalenten Inhibition eines Enzyms vorliegen, schwierig, Berechnungen zur Reaktivität auf experimentell bestimmte Kinetik-Daten zu übertragen.<sup>19</sup> Aufgrund der hohen Ähnlichkeit der Warheads kann aber bei einigen Einflussfaktoren davon ausgegangen werden, dass diese effektiv identisch für die unterschiedlichen Verbindungen sein müssten und somit nicht verantwortlich für experimentell beobachtete Unterschiede sind. Da zum Beispiel der Raumbedarf

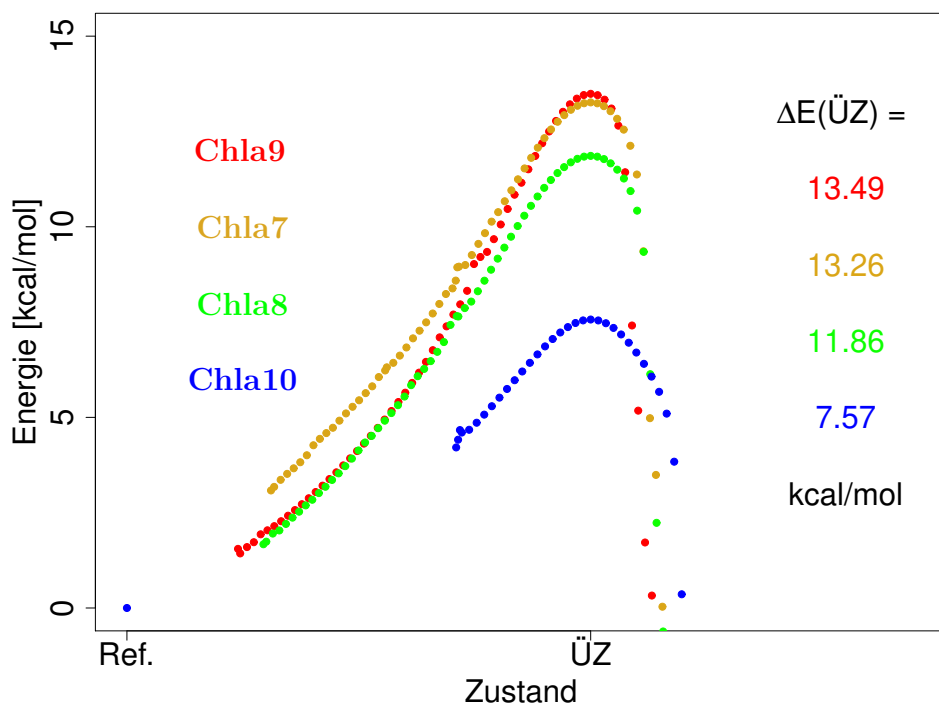


Abbildung 4.82.: Energieprofile für die Verbindungen **Chla7** bis **Chla10**. Die Energieskala ist in Vergleich zu Abbildung 4.79 und Abbildung 4.81 leicht erweitert, um die erhöhten Aktivierungsenergien ebenfalls zu erfassen.

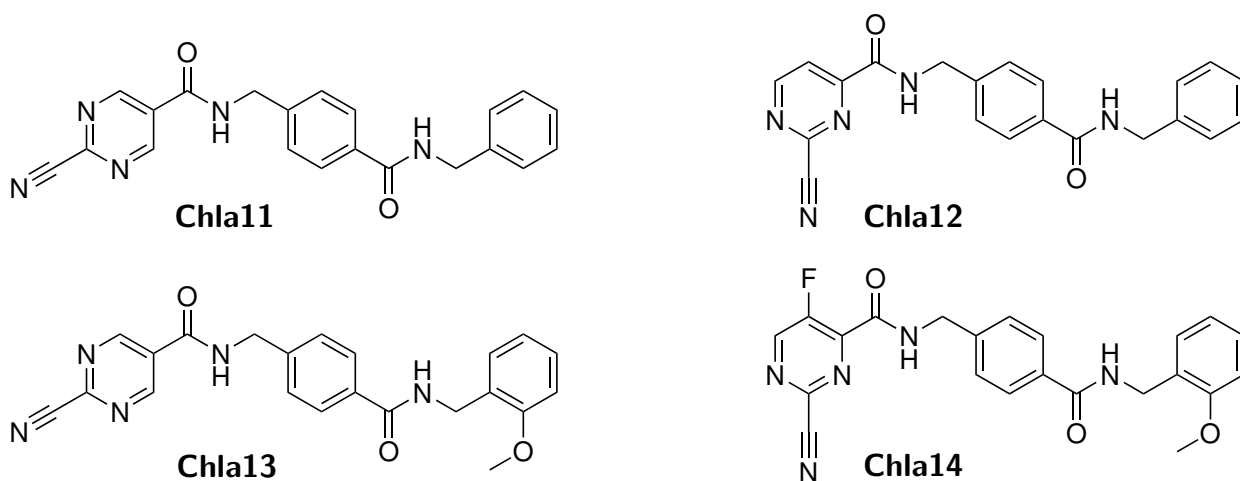


Abbildung 4.83.: Die von Thomas Zimmermann experimentell untersuchten Verbindungen. **Chla11** und **Chla13** weisen dabei den Warhead **Chla1** auf, **Chla12** den Warhead **Chla4** und **Chla14** besitzt **Chla5** als reaktive Einheit.

bei **Chla4** und **Chla5** nahezu identisch ist, nahezu dieselben nichtkovalenten Wechselwirkungen zum Protein ausgebildet werden und in beiden Fällen der Warhead des Liganden gleich orientiert in der Bindetasche liegen müsste, sollten beide Liganden etwa gleich schnell nichtkovalente Komplexe mit dem Enzym ausbilden, die eine ähnliche Stabilität aufweisen, und etwa gleich oft für eine Reaktion günstig orientiert sein. Folglich sollten Unterschiede in der Inhibition hauptsächlich auf Unterschiede in der Geschwindigkeitskonstante der kovalenten Reaktion,  $k_3$  (siehe Kapitel 1.1.2), zurückzuführen sein.

	<b>Chla11</b>	<b>Chla12</b>	<b>Chla13</b>	<b>Chla14</b>
Warhead	<b>Chla1</b>	<b>Chla4</b>	<b>Chla1</b>	<b>Chla5</b>
$\Delta E_A$ [kcal/mol]	8.53	8.61	8.53	8.21
$\Delta G_A$ [kcal/mol]	18.4	18.7	18.4	18.3
$k_{inact}$ [min <sup>-1</sup> ]	0.13	0.16	0.16	0.37
	$\pm 0.02585$	$\pm 0.01$	$\pm 0.01$	$\pm 0.018$

Tabelle 4.23.: Berechnete Aktivierungsenergien und von Thomas Zimmermann experimentell bestimmte  $k_{inact}$ -Werte für die Verbindungen **Chla11**, **Chla12**, **Chla13** und **Chla14**.

Diese Unterschiede können über die Arrhenius-Gleichung vorausgesagt werden (siehe Gleichung 4.8). Die Arrhenius-Gleichung stellt einen Zusammenhang zwischen der Aktivierungsenergie  $\Delta G_A$  und der Geschwindigkeitskonstante  $k$  einer Reaktion her (hier  $k_3$ ). Der Parameter  $A$  ist ein Maß für die Häufigkeit einer produktiven Kollision zwischen zwei Reaktanden.<sup>97</sup> Solange die Warheads zweier Verbindungen in der Bindetasche gleich orientiert vorliegen, müsste  $A$  daher nahezu identisch sein.  $R$  und  $T$  sind die universelle Gaskonstante und die Temperatur.

$$k_3 = A \cdot e^{\frac{-\Delta G_A}{RT}} \quad (4.8)$$

Dabei sollte allerdings beachtet werden, dass der hier diskutierte Mechanismus lediglich eine Modellreaktion darstellt und daher nicht dazu geeignet ist, absolute Geschwindigkeitskonstanten aus den berechneten Aktivierungsenergien abzuleiten. Zusätzlich ist  $A$  nicht bekannt. Diese Problematik kann umgangen werden, indem die Reaktivitäten relativ zueinander betrachtet werden. So können sich methodische Fehler innerhalb der quantenmechanischen Rechnungen gegenseitig ausgleichen und gleichzeitig der Einfluss von  $A$  eliminiert werden, sofern es vertretbar ist, bei den beiden Verbindungen von einem identischen  $A$  auszugehen. Unter der Annahme, dass der  $k_{inact}$ -Wert einer Verbindung proportional zu der Geschwindigkeitskonstante  $k_3$  ist, da alle anderen kinetischen Parameter ungefähr gleich sind, sollte somit für zwei Verbindungen  $a$  und  $b$  gelten:

$$\frac{k_{inact,a}}{k_{inact,b}} \approx \frac{k_{3,a}}{k_{3,b}} \approx \frac{e^{-\frac{\Delta G_{A,a}}{RT}}}{e^{-\frac{\Delta G_{A,b}}{RT}}} \approx e^{\frac{-\Delta G_{A,a} + \Delta G_{A,b}}{RT}} \quad (4.9)$$

Für die Verbindungen **Chla12** und **Chla14**, die die nötige Ähnlichkeit zueinander aufweisen, bzw. deren Warheads **Chla4** und **Chla5**, ergäbe sich somit:

$$e^{\frac{-18.7\text{kcal/mol} + 18.3\text{kcal/mol}}{300\text{K} \cdot 1.987 \cdot 10^{-3}\text{kcal}/(\text{mol} \cdot \text{K})}} = e^{\frac{-0.4\text{kcal/mol}}{300\text{K} \cdot 1.987 \cdot 10^{-3}\text{kcal}/(\text{mol} \cdot \text{K})}} \approx 0.51 \quad (4.10)$$

Verglichen mit dem experimentell bestimmten Verhältnis der  $k_{inact}$ -Werten von **Chla12** und **Chla14** von  $\frac{0.1604}{0.3737} = 0.43$  und unter Berücksichtigung der experimentellen Unsicherheit ist dies in relativ guter Übereinstimmung. Für analoge Vergleiche zwischen den Verbindungen **Chla11**, **Chla12** und **Chla13** überlappen die dazugehörigen Fehlerbereiche zu stark. Tendenziell kann aber auch hier erkannt werden, dass die drei Verbindungen ähnliche Aktivierungsenergien und auch im Experiment sehr ähnliche  $k_{inact}$ -Werte aufweisen.

Rein auf der Aktivierungsenergie basierend würde für eine Verbindung mit dem Sulfonamid **Chla6** als Warhead ein deutlich niedrigerer  $k_{inact}$ -Wert im Vergleich zu den anderen Verbindungen erwartet werden. Allerdings muss hier berücksichtigt werden, dass aller Wahrscheinlichkeit nach von **Chla6** nicht derselbe nichtkovalente Bindemodus eingenommen werden kann wie z. B. bei dem Warhead **Chla4**, da das Sulfonamid deutlich größer ist und eine andere Geometrie besitzt als ein Amid. Folglich wird vermutlich der präexponentielle Faktor  $A$  (siehe Gleichung 4.8) nicht mehr vergleichbar sein. Zum Zeitpunkt der Entstehung dieser Arbeit werden noch Experimente mit Verbindungen mit dem **Chla6**-Warhead durchgeführt. Erste Messungen deuten daraufhin, dass eine solche Verbindung tatsächlich schlechter an das Protein bindet als vergleichbare Inhibitoren, exakte Werte stehen allerdings noch aus.

Die Verbindungen mit den Warheads **Chla7**, **Chla8** und **Chla9** wiesen im Experiment selbst bei der höchsten gemessenen Konzentration an Inhibitor keine Inhibition der Chla-DUB1 auf. In Analogie dazu wurde für diese Verbindungen hier eine um ca. 3-5 kcal/mol höhere Aktivierungsenergie berechnet, was einer um mehrere Größenordnungen langsameren Reaktionsgeschwindigkeit entspricht.

Zusammenfassend lässt sich sagen, dass es mit Hilfe des hier genutzten QM-Protokolls durchaus möglich ist, Aktivierungsenergien der Warheads zu berechnen, was eine Abschätzung der experimentellen Geschwindigkeitskonstante erlaubt, solange diese Betrachtung zwischen zwei Inhibitoren erfolgt, deren Erkennungseinheit sehr ähnlich zueinander sind. Für eine vollwertige statistische Auswertung, wie gut die Korrelation zwischen Theorie und Praxis ist, wären allerdings mehr Datenpunkte nötig, die einen größeren Reaktivitätsbereich über-



spannen. Dies ist zwangsläufig mit experimentellen Problemen verknüpft, da nur bis zu einer gewissen Konzentrationsgrenze die Inhibition effektiv gemessen werden kann. Zusätzlich können bereits geringfügige geometrische Änderungen innerhalb des Liganden ausreichen, um die zugrunde liegenden theoretischen Annahmen (Gleichheit des Arrhenius-Parameters  $A$ , ausreichende Übereinstimmung im nichtkovalenten Bindemodus) hinfällig werden zu lassen. Zum Zeitpunkt der Entstehung dieser Arbeit werden allerdings noch weitere Messungen von Thomas Zimmermann vorgenommen, sodass in der Zukunft eine statistisch umfassendere Auswertung möglich sein könnte.

Sowohl auf Seite des theoretischen Protokolls als auch auf Seite des Experiments bestehen Möglichkeiten, um potenziell eine bessere Korrelation zu erreichen. Die thermodynamischen Korrekturen, die zur Berechnung von  $\Delta G_A$  der Modellreaktion genutzt werden, beinhalten Terme, die abhängig von der Translation und Rotation des Moleküls sind. Ein Inhibitor, der sich gebunden in der Bindetasche eines Proteins befindet, kann allerdings weder vor noch nach der Reaktion frei rotieren und sich nicht beliebig im Raum bewegen; lediglich die möglichen Vibrationen ändern sich durch die Reaktion mit dem Protein. Daher ist es vorstellbar, dass eine gezielter Einbezug von lediglich den Vibrationskorrekturen eine verbesserte Vorhersage der Reaktivität erlaubt.

Das hier genutzte Protokoll, bestehend aus einer DFT-Methode in Kombination mit einem Pople-Basissatz, B3LYP/6-311++G\*\*,<sup>25</sup> orientierte sich stark an einem bereits publizierten Vorgehen,<sup>186</sup> da diese Vorgehensweise ähnliche Ergebnisse wie die *ab initio*-Methode MP2 (Møller-Plesset-Störungstheorie zweiter Ordnung) produzierte, gut mit dem Experiment korrelierte und relativ gering im Rechenaufwand ist. Dort wurde allerdings nicht getestet, ob eine empirische Dispersionskorrektur die Genauigkeit der Ergebnisse weiter steigern könnte. Aber auch wenn diese Faktoren theoretisch in der Lage sind, die Präzision der Rechnung zu erhöhen, so bleibt dadurch dennoch der hohe relative Fehler der Messung bestehen (siehe Tabelle 4.23). Vor diesem Hintergrund stellt das hier angewendete Protokoll eine ausreichende Genauigkeit für die Abschätzung der Reaktivität zur Verfügung.

Um Einflüsse der nichtkovalenten Bindung auf die gemessenen Geschwindigkeitskonstanten komplett zu eliminieren, könnten analoge Experimente mit Cystein oder Glutathion (GSH) statt dem Protein durchgeführt werden. Insbesondere wäre diese Messung für solche Liganden interessant, die laut der berechneten Aktivierungsenergie reaktiver sein sollten als vergleichbare Moleküle, in der Messung mit Protein allerdings schlechter bzw. langsamer binden (z. B. **Chla6**). In der zusätzlichen Messreihe mit Cystein/GSH sollten solche Inhibitoren wieder dem berechneten Trend folgen, was zeigen würde, dass die vergleichsweise langsame Reaktion mit dem Protein auf einen ungünstigen nichtkovalenten Bindungsmodus

Verbindung	$\Delta E_A$ Berteotti <i>et al.</i> <sup>186</sup> [kcal/mol]	$\Delta E_A$ diese Arbeit [kcal/mol]
<b>U19</b>	12.0	11.8
<b>U26</b>	7.7	9.54
<b>U27</b>	15.7	15.5

Tabelle 4.24.: Vergleich der berechneten Aktivierungsenergien zwischen Berteotti *et al.*<sup>186</sup> und dieser Arbeit.

zurückzuführen ist. Im Rahmen des Wirkstoffdesigns ist eine schnelle Reaktion mit Cystein oder GSH aber prinzipiell unerwünscht, da dies ein Anzeichen für eine unselektive Reaktion mit Thiolen ist.

Ein direkter Vergleich mit anderen bereits publizierten quantenmechanischen Protokollen ist schwierig, da meist für die individuellen Projekte spezifische Moleküle berechnet wurden und methodische Details für deutliche Abweichungen sorgen können. Oballa *et al.*<sup>194</sup> nutzten einen systematisch anderen Ansatz, indem sie Reaktionsenergien statt Aktivierungsenergien für das Ranking der Warheads nutzten, was den Vergleich zu den hier erhaltenen Werten erschwert. Das zugrundeliegende Protokoll von Berteotti *et al.*<sup>186</sup> besitzt z. B. keine Amid-substituierten oder fluorierten Verbindungen in seiner Testreihe, sodass nur ein geringer Überlapp mit den in dieser Arbeit berechneten Verbindungen besteht. Tabelle 4.24 zeigt daher ein paar ausgewählte Aktivierungsenergien als Beispiel, die sowohl von Berteotti *et al.* als auch in dieser Arbeit berechnet wurden. In einigen Fällen können die bereits publizierten Werte gut reproduziert werden, während für z. B. **U26** deutliche Abweichungen feststellbar sind. Hauptgrund für solche Abweichungen ist vermutlich der unterschiedlich definierte Referenz-Zustand (siehe Kapitel 4.2.4.5). Für Acrylamide existieren Protokolle, die lediglich die Energien des höchsten besetzten und des niedrigsten unbesetzten Molekülorbitals (HOMO und LUMO) nutzen, um die Reaktivität vorherzusagen.<sup>26</sup> Diese Werte sind mit einem deutlich geringeren Rechenaufwand verbunden, allerdings sind solche Protokolle für Nitrile noch nicht etabliert.

Bis größere Datensätze für Inhibitoren mit verschiedenen Nitril-Warheads vorliegen, sollte das hier beschriebene Protokoll eine ausreichende Abschätzung bei akzeptablem Rechenaufwand zur Verfügung stellen.

## 4.4. BChE: Nichtkovalentes Docking eines kovalent agierenden Photoswitches

Molekulares Docking erlaubt, Modelle für den Bindemodus eines Liganden zu erstellen. An den Bindemodus eines kovalenten Inhibitors müssen dabei zusätzliche Anforderungen gestellt werden: Der Warhead der Verbindung muss bei der zu adressierenden Aminosäure platziert sein, da ansonsten eine Reaktion im Rahmen des Modells nicht möglich wäre. Besonders für Liganden mit sterisch anspruchsvollen Erkennungseinheiten stellt diese zusätzliche Bedingung ein wichtiges Kriterium dar, um plausible Bindungsmodelle zu generieren und so Struktur-Wirkung-Beziehungen besser erklären zu können. Als Fallbeispiel hierfür können die Photoswitches mit Carbamat-Warhead, die im Arbeitskreis von Prof. Decker zur Inhibition von BChE entwickelt wurden, herangezogen werden.<sup>68</sup>

Als Erkennungseinheit für diese Inhibitoren dient ein viergliedriges Ringsystem, das bereits in anderen BChE-Projekten Anwendung gefunden hat.<sup>195–197</sup> Der Fokus in dieser Arbeit liegt auf dem Photoswitch **B1**, der mit Hilfe von Licht mit einer Wellenlänge von 365 nm in die *cis*-Form umgesetzt, und mit Hilfe von 455 nm in die *trans*-Form zurückgebracht werden kann (siehe Abbildung 4.84). Das *cis*-Isomer ist hierbei etwa zehnmal aktiver als das *trans*-Isomer ( $IC_{50}$  von 44.6 nM vs. 424 nM).<sup>68</sup> Es liegen allerdings zum Zeitpunkt dieser Arbeit keine kristallographischen Daten vor, die diesen Umstand erklären könnten. Prinzipiell sollte dieser Umstand über den nichtkovalenten Bindungsmodus erklärbar sein sollen, da hier der Photoswitch noch Teil des Moleküls ist und noch nicht abgespalten wurde.

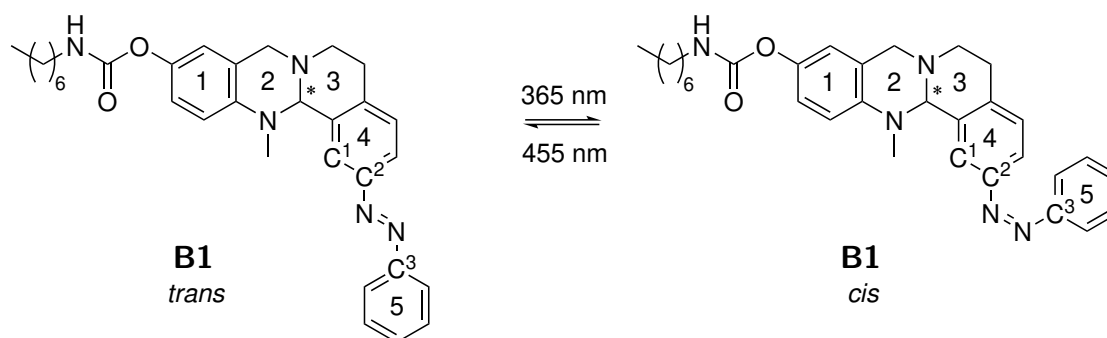


Abbildung 4.84.: Der Photoswitch **B1** in seiner *trans*- und *cis*-Form. Einzelne Ringe und Atome wurden zur leichteren Referenzierung hervorgehoben.

Mithilfe von molekularem Docking soll dieser Unterschied in der Inhibition qualitativ erklärt werden. Dabei soll darauf geachtet werden, dass in dem formulierten Bindungsmodell das Carbamat für eine Reaktion plausibel in der Bindetasche von BChE platziert wird, d. h. in

der Nähe des katalytisch aktiven Serins Ser198. Teile der hier gezeigten Ergebnisse wurden bereits publiziert.<sup>68</sup>

#### 4.4.1. Stereochemische Analyse des Photoswitches

Verbindung **B1** weist einige stereochemische Problematiken auf, weil neben einem chiralen Kohlenstoffatom noch weitere Stereo-Elemente in der Verbindung vorliegen. Das viergliedrige Ringsystem (siehe Abbildung 4.85), weist ein pyramidales Stickstoffatom zwischen Ring 2 und 3 auf. Dieses Stickstoffatom stellt im Rahmen des Modelings ein Stereozentrum dar, falls dessen Inversion nicht durch die entsprechende Modeling-Methode abgebildet wird (siehe Abbildung 4.85, a)-d)).

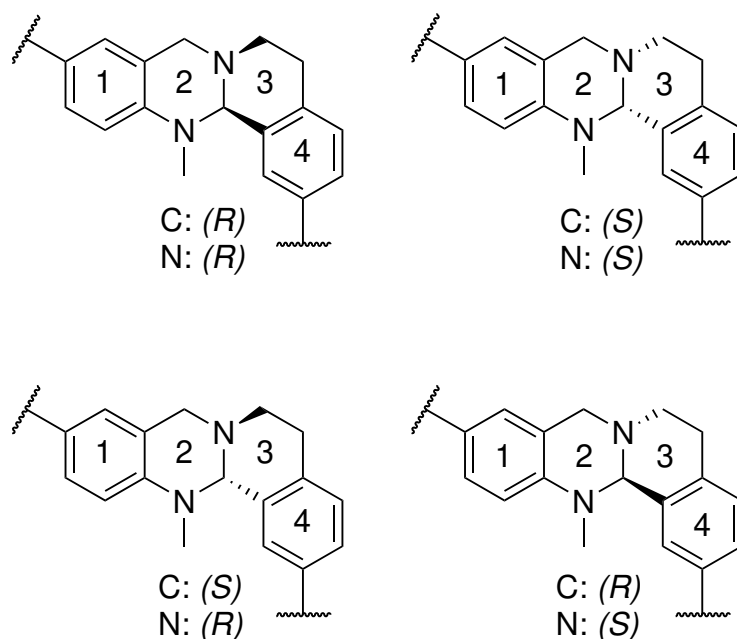
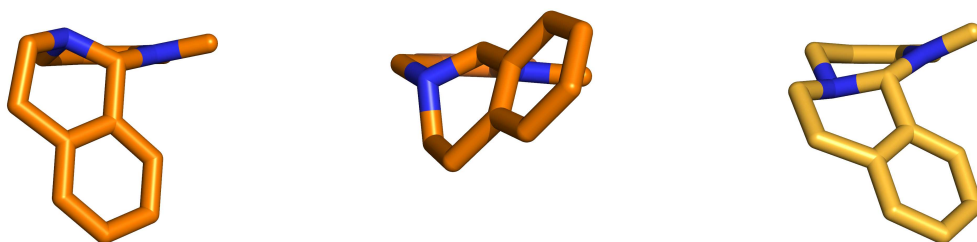


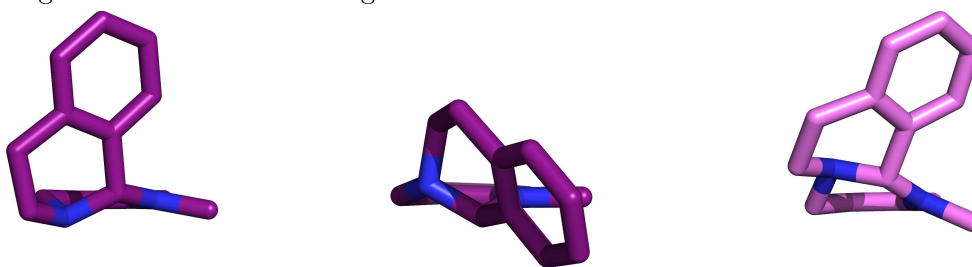
Abbildung 4.85.: Unterschiedliche Konfigurationen des viergliedrigen Ringsystems von **B1**.

Damit ergibt sich für die Ringe 2 und 3 eine stereochemische Situation, die der des Decalins ähnelt. Die Substituenten an den beiden anellierten Atomen können zueinander *cis*-konfiguriert vorliegen, wofür wiederum zwei unterschiedliche, ineinander invertierbare Ringkonformere existieren (vergleiche Abbildung 4.86a mit Abbildung 4.86b bzw. Abbildung 4.86d mit Abbildung 4.86e). Sie können aber auch in einer *trans*-Konfiguration vorliegen, für die ein Ringflip unmöglich ist, weil die beiden Substituenten nicht gleichzeitig axial vorliegen können. Somit ist bei einer *trans*-Konfiguration nur ein Konformer denkbar (siehe Abbildung 4.86c und Abbildung 4.86f).

Für das nachfolgende Docking wird für die *cis*-konfigurierten Varianten nur dasjenige Konformer betrachtet, bei dem die Bindung des chiralen Kohlenstoffatoms axial ausgerichtet ist, da so eine sterisch ungünstige Positionierung der anilischen *N*-Methylgruppe zu einer aromatischen CH-Gruppe vermieden wird.



(a) C (*R*), N (*R*), pyramidales N mit äquatorialer Bindungsorientierung (b) C (*R*), N (*R*), pyramidales N mit axialer Bindungsorientierung (c) C (*R*), N (*S*), *trans*-Ausrichtung



(d) C (*S*), N (*S*), pyramidales N mit äquatorialer Bindungsorientierung (e) C (*S*), N (*S*), pyramidales N mit axialer Bindungsorientierung (f) C (*S*), N (*R*), *trans*-Ausrichtung

Abbildung 4.86.: Räumliche Darstellung der unterschiedlichen Stereoisomere des Ringsystems. Ring 1 ist dabei senkrecht zur Bildebene ausgerichtet und im Hintergrund, während Ring 4 im Vordergrund ist.

In allen möglichen Konfigurationen ergibt sich eine „Krümmung“ innerhalb des Ringsystems, die aber abhängig von der relativen Konfiguration unterschiedlich stark ausgeprägt ist. Für den Fall, dass das Kohlenstoffatom und das pyramidale Stickstoffatom gleich konfiguriert sind (d. h. beide (*R*)- oder beide (*S*)-konfiguriert), sind die Ringe 3 und 4 nahezu senkrecht zu den Ringen 1 und 2 (siehe Abbildung 4.86a und Abbildung 4.86d), während bei entgegengesetzt konfigurierten Chiralitätszentren sich eine wesentlich flachere Biegung ergibt (siehe Abbildung 4.86c und Abbildung 4.86f). Die absolute Konfiguration des Kohlenstoffatom ist ausschlaggebend für die Richtung der Krümmung: Bei einem (*R*)-konfigurierten Kohlenstoffatom bilden die Orientierung der *N*-Methylgruppe und die beiden Phenylringe

1 und 4 ein rechtshändiges System (in Abbildung 4.86 in Orange-Tönen eingefärbt), bei einem (*S*)-konfigurierten Kohlenstoffatom ein linkshändiges System (in 4.86 in Lila-Tönen eingefärbt).

Azobenzole stellen in ihrer *cis*-Form außerdem noch ein zusätzliches stereochemisches Element dar. Die beiden Phenylringe 4 und 5 können aufgrund von sterischen Wechselwirkungen nicht komplett planar in einer Ebene liegen. Stattdessen wird eine verdrehte Konfiguration angenommen, bei der ein C<sup>1</sup>-C<sup>2</sup>-N-N-Diederwinkel (bzw. ein N-N-C<sup>3</sup>-C<sup>4</sup>-Diederwinkel) von etwa 51° vorliegt (siehe Abbildung 4.87).



Abbildung 4.87.: Helikale Chiralität eines *cis*-Photoswitches. Die beiden gezeigten Konfigurationen können nur durch Spiegelung ineinander überführt werden.

Diese Drehung aus der Phenylring-*N*-Bindungsebene kann im oder gegen den Uhrzeigersinn stattfinden. Folglich weist bereits das simpelste Azobenzol eine helikale Chiralität auf (siehe Abbildung 4.87).<sup>198</sup> Sollte einer der Phenylringe zusätzliche Substitutionen aufweisen, ergeben sich komplexere stereochemische Szenarien.<sup>198</sup> In dem hier diskutierten Fall stellt das viergliedrige Ringsystem mit seinen Stereozentren einen sterisch großen Referenzpunkt dar, bei der der Phenylring 5 des Photoswitches entweder über oder unter der Ringebene platziert sein kann.

Diese Helizität erhält im Kontext der Enzym-Inhibition eine größere Bedeutung: Ganz analog zu klassischer, zentraler Stereoisomerie ist es denkbar, dass ein bestimmtes helikales Enantiomer besser in die Bindetasche passt als sein spiegelbildliches Gegenstück, und somit für die beiden Enantiomere individuelle Inhibitionskonstanten und IC<sub>50</sub>-Werte existieren können.

Den entsprechenden Effekt allerdings experimentell nachzuweisen ist schwierig, da die unterschiedlichen Stereoisomere durch Bestrahlung derselben *trans*-Verbindung generiert werden, d. h. nach dem „Switchen“ liegt vermutlich ein Gemisch der beiden helikalen Enantiomere vor. Typischerweise kann nur dann ein bestimmtes helikales Enantiomer in seiner Reinform gebildet werden, wenn die beiden Azobenzol-Ringe über einen (chiralen) Linker miteinander

verknüpft werden.<sup>199–201</sup> Es ist zwar denkbar, dass durch die stereochemische Information innerhalb des Ringsystems ein bestimmtes Diastereomer bevorzugt beim „Switchen“ generiert wird, jedoch sind derartige Präferenzen kaum literaturbekannt.<sup>198</sup>

Diese geometrischen Besonderheiten der Verbindung **B1** haben konkrete Konsequenzen für die Gestaltung des Docking-Protokolls. Falls die gewählte Docking-Methode nicht selbstständig das optimale Isomer bestimmen kann, müssen separate Dockings für die entsprechenden Diastereomere durchgeführt werden, und die erhaltenen Docking-Ergebnisse miteinander verglichen werden. Für Verbindung **B1**, vor allem für deren *cis*-Konfiguration, gilt dies insbesondere, da die Diederwinkel der Azobindung selten in Docking-Programmen parametrisiert wird.

Jede Kombinationsmöglichkeit an Heterozyklus-Isomer und Photoswitch-Positionierung separat zu docken stellt allerdings einen enormen Aufwand dar. Daher wurde mit Hilfe einiger Überlegungen versucht, vermutlich ungünstige Stereoisomere auszuschließen.

Durch die zuvor beschriebene Krümmung des Ringsystems existiert eine konvexe und eine konkave Seite des Liganden. Den *cis*-Photoswitch auf der konvexen Seite zu platzieren erhöht effektiv den Platzbedarf des bereits ohnehin großen Inhibitors. Um somit den sterischen Ansprüchen der Bindetasche besser gerecht zu werden, wurden nur die *cis*-Isomere betrachtet, bei denen der Phenylring auf der konkaven Seite und über dem Ringsystem befindlich ist (siehe Abbildung 4.88). Diese Annahme reduziert die Anzahl der relevanten Diastereomere enorm.

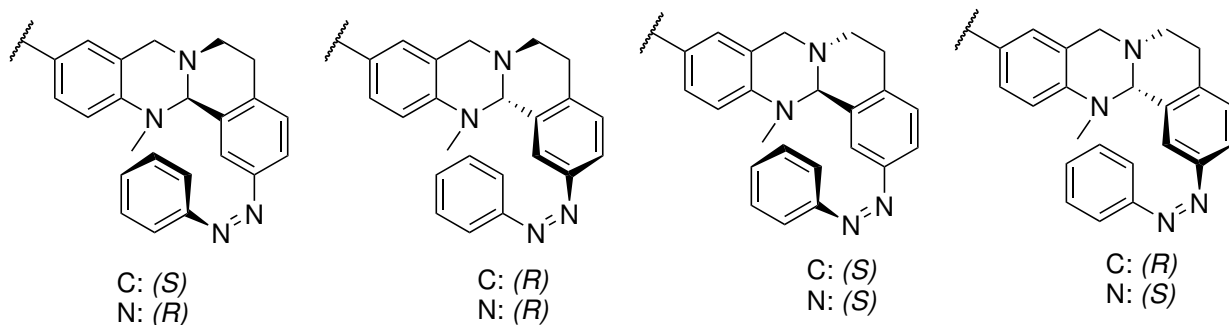


Abbildung 4.88.: Alle untersuchten Konfigurationen des *cis*-Photoswitches.

#### 4.4.2. Methodische Details zum Docken eines Photoswitches

AutoDock3<sup>93</sup> wurde für diese Docking-Studie genutzt, da es erlaubt, viele der Eigenheiten der Photoswitches durch den Benutzer zu definieren. Als Startpunkt wurde eine Kristallstruktur von BChE genutzt (PDB-Code 6QAA). Die Vorbereitung der Proteinstruktur wurde in MOE

durchgeführt.<sup>148</sup> Vor dem Docking wurden alle Wasser- und Puffermoleküle für den Kristallisationsprozess entfernt. Trp82 wurde mit Hilfe von Amber14:EHT minimiert, da vorherige Docking-Tests von Praktikantin Constanze Weber nahelegten, dass es mit dieser Aminosäure zu einem sterischen Clash kommen könnte, wenn der Heterozyklus des Moleküls in der Nähe der katalytisch aktiven Seite platziert werden soll. Protonierungszustände wurden mit Hilfe des *protonate3D*-Programms<sup>149</sup> von MOE bei pH=7 festgelegt, was zu Standardprotonierungszuständen für das Protein führte. Das katalytische His438 war  $\delta$ -protoniert, während das tertiäre Amin von **B1** nicht protoniert war.

Ligandstrukturen wurden in MOE konstruiert. Der Ligand wurde, bis auf das Azobenzol, mit Hilfe des Kraftfeldes MMFF94x minimiert,<sup>150</sup> bis die Wurzel des Quadrats des mittleren Gradienten weniger als 0.001 kcal/(mol Å) betrug. Anschließend wurden dem Liganden Gasteiger-Ladungen zugewiesen.<sup>181</sup> Um besser die Geometrie des Photoswitches zu beschreiben wurde eine QM-Optimierung mit Gaussian09, Revision C.01,<sup>181</sup> eines trunkierten Modells des *cis*-Photoswitches auf dem B3LYP/6-311+G\*\* - Level durchgeführt. Die so erhaltenen Diederwinkel für die C-C-N-N-Torsion (51°) und die C<sup>2</sup>-N-N-C<sup>3</sup>-Torsion (2°) wurden für den Liganden übernommen. Die Carbamat- und die Diederwinkel des *cis*-Photoswitches wurden im Docking nicht variiert. Die N-N-C<sup>3</sup>-C<sup>4</sup>-Torsion des terminalen Phenylrings des *trans*-Photoswitches wurde mit einem gaussförmigen Potenzial von 6 kcal/mol modifiziert, um eine leichte Drehung des Rings aus der Ebene des restlichen Photoswitches zu erlauben. Die Stärke dieses Potenzials wurde mit Hilfe eines QM-Scans der Torsion auf dem B3LYP/6-311+G\*\* - Level in 15° Schritten abgeschätzt.

Basierend auf der zuvor durchgeführten Diskussion der Stereochemie der Verbindung **B1** wurde anschließend eine Auswahl an Stereoisomeren gedockt. 100 Durchläufe des Lamarck-schen genetischen Algorithmus wurden durchgeführt, mit 300 Individuen in der initialen Population, maximal 27 000 Generationen oder 12 Millionen Energiebewertungen. Die Raten für Mutation und Crossover betragen jeweils 0.02 und 0.80. Der Elitismus wurde auf 1 gesetzt. Die lokale Suche mit dem pseudo-Solis-Wets-Suchalgorithmus wurde mit einer Rate von 0.06 durchgeführt. 300 Iterationen dieser lokalen Suche wurden durchgeführt, wobei nach vier konsekutiven Erfolgen oder Misserfolgen die Schrittweite der Suche angepasst wurde. Falls ein Parameter hier nicht explizit gelistet wurde, wurden dessen Default-Wert übernommen. Die erhaltenen Docking-Lösungen wurden mit einem Cutoff von 2 Å geclustert, wobei die bestbewertete, noch nicht einem Cluster zugewiesene Pose immer den Referenzpunkt für den nächsten Cluster bildete.

Alle Posen wurden anhand der *Final Docked Energy* bewertet, welche die Summe aus einem intermolekularen Term, der *Final Intermolecular Energy* und einem Term für Wechselwir-



kungen innerhalb des Liganden, der *Final Internal Energy*, ist. Zusätzlich wurde der Abstand des Carbamats zur katalytischen Einheit als Bewertungskriterium betrachtet, da die Verbindung **B1** als kovalent agierende Verbindung nahe an diesen Aminosäuren platziert sein muss.

### 4.4.3. Ergebnisse des nichtkovalenten Dockings

Die Docking-Ergebnisse der Top-Posen für alle untersuchten Stereoisomere sind in Tabelle 4.25 zusammengefasst. Lediglich für zwei Konfigurationen kann eine kurze Distanz zwischen dem reaktiven Carbamat-Kohlenstoffatom des Liganden und dem Hydroxyl-Sauerstoffatom von Ser198 beobachtet werden, nämlich bei den beiden *cis*-Formen der am chiralen Kohlenstoffatom *R*-konfigurierten Diastereomere. Die am Stickstoffatom (*S*)-konfigurierte Variante des Liganden wird dabei leicht besser bewertet als die (*R*)-konfigurierte (-15.29 kcal/mol vs. -14.95 kcal/mol) und weist ebenfalls eine verbesserte Clustergröße auf (17 anstelle von 3 Posen).

Konfiguration		Final Inter-molecular Energy [kcal/mol]	Final Internal Energy [kcal/mol]	Final Docked Energy [kcal/mol]	Distanz [Å]	Clustergröße	
Azo-Switch	C						N
<i>cis</i>	( <i>R</i> )	( <i>R</i> )	-14.83	-0.12	-14.95	3.64	3
<i>cis</i>	( <i>R</i> )	( <i>S</i> )	-14.76	-0.53	-15.29	3.02	17
<i>cis</i>	( <i>S</i> )	( <i>R</i> )	-13.97	-0.48	-14.45	10.7	3
<i>cis</i>	( <i>S</i> )	( <i>S</i> )	-14.41	-0.36	-14.78	6.67	10
<i>trans</i>	( <i>R</i> )	( <i>R</i> )	-13.21	-0.05	-13.26	11.39	11
<i>trans</i>	( <i>R</i> )	( <i>S</i> )	-13.17	-0.36	-13.53	13.95	3
<i>trans</i>	( <i>S</i> )	( <i>R</i> )	-11.78	-0.04	-11.82	4.71	12
<i>trans</i>	( <i>S</i> )	( <i>S</i> )	-13.36	-0.08	-13.44	8.70	5

Tabelle 4.25.: Docking-Ergebnisse der unterschiedlichen Diastereomere der Verbindung **B1**.

Abbildung 4.89 zeigt die beiden entsprechenden Posen. In beiden Fällen ist das Ringsystem in der Cholin-Bindetasche platziert, in direkter Nachbarschaft zu Trp82, Thr120 und Tyr128. Die Heptylketten folgen der Oberfläche der Acyl-Bindestelle von BChE; und bei beiden Verbindungen ist das Carbamat bei der katalytischen Einheit platziert, mit dem Carbonyl-Sauerstoffatom nahe dem Oxyanionenloch. Damit erscheint eine kovalente Reaktion für diese Top-Posen plausibel. Zusätzlich ist der Bindemodus zu bereits publizierten Docking-Ergebnissen von ähnlichen Verbindungen.<sup>197</sup>

Diese modellierten Bindeposen erlauben qualitativ eine Hypothese aufzustellen, warum die entsprechenden *trans*-Verbindungen schlechtere Inhibitoren der BChE sind. Bei einem vergleichbaren Bindemodus der *trans*-Formen würde ein Phenylring des Photoswitches sehr na-

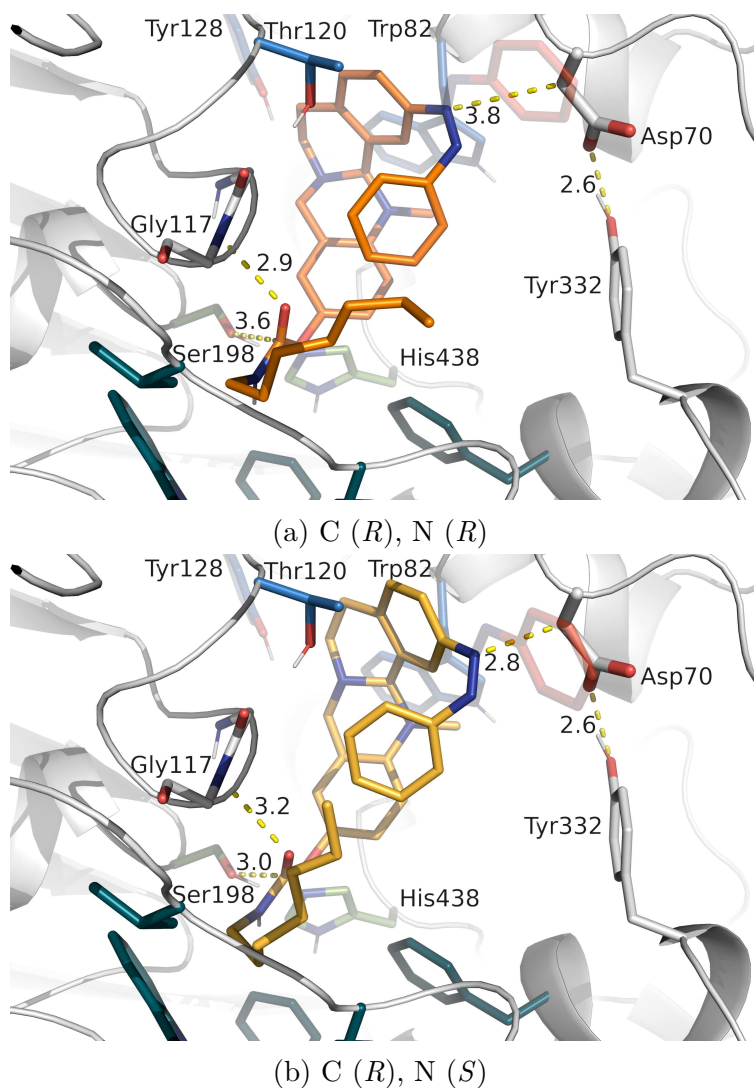


Abbildung 4.89.: Top-Docking-Posen für die Stereoisomere von *cis*-**B1**, die das Carbamat günstig zur katalytischen Einheit platziert haben. Diese Abbildung wurde mit Genehmigung adaptiert,<sup>68</sup> Copyright 2022 American Chemical Society.

he an Asp70 platziert werden (in Abbildung 4.89 transparent angedeutet). Um einen Clash oder ungünstige Wechselwirkungen zwischen dem lipophilen Phenylring und den hydrophilen Aminosäuren zu vermeiden, müsste entweder die Wasserstoffbrückenbindung zwischen Asp70 und Tyr332 unterbrochen oder ein komplett anderer Bindemodus eingenommen werden. Letzteres erscheint schwierig, da das Ringsystem sehr rigide ist und sich konformativ kaum anpassen kann. Lediglich ein *trans*-Isomer (mit (*S*)-konfiguriertem Kohlenstoffatom und (*R*)-konfiguriertem Stickstoffatom) platziert annähernd das Carbamat so nahe wie die *cis*-Stereoisomere, wird aber energetisch mit ca. 3 kcal/mol deutlich schlechter bewertet.

Qualitativ lässt sich auch nachvollziehen, warum nur für die am chiralen Kohlenstoffatom (*R*)-konfigurierten Verbindungen plausible Bindungsposen identifiziert werden können. Unter der Annahme, dass das sterisch anspruchsvolle Ringsystem lipophile Wechselwirkungen zu z. B. Trp82 innerhalb der Cholin-Subtasche ausbilden soll, wird nur bei einem rechtsgekrümmten Ringsystem gleichzeitig das Carbamat nahe beim katalytischen Ser198 platziert. Bei einem linksgekrümmten Ringsystem würde nicht nur der Photoswitch eine analoge Positionierung verhindern, sondern auch das Carbamat von der katalytischen Einheit weg deuten. Dieses Fallbeispiel des Dockings von **B1** in die Bindetasche von BChE zeigt, dass über den nichtkovalenten Bindungsmodus bereits essenzielle Informationen für die Struktur-Wirkungs-Beziehung eines kovalenten Inhibitors gewonnen werden können. Basierend auf diesem Bindungsmodell können nun Entscheidungen getroffen werden, wie **B1** optimiert werden kann (z. B. ob und wie der *cis/trans*-Unterschied verstärkt werden kann). Auch wenn dieses konkrete Docking-Problem durch die stereochemische Komplexität und fehlende Parameter für die Geometrie des Photoswitches anspruchsvoll ist, kann dennoch ein plausibler Bindungsmodus modelliert werden, in dem der Warhead günstig für eine Reaktion positioniert ist. Es ist erwähnenswert, dass die Generierung einer kovalenten Docking-Pose potenziell noch aufwendiger wäre, da zusätzlich zu den bereits diskutierten Problematiken noch Parameter für die kovalente Verknüpfung gefunden werden müssen. Das nichtkovalente Docking eines kovalenten Inhibitors erscheint somit prinzipiell vorteilhaft.

## 5. Zusammenfassung

Kovalente Inhibition stellt einen effektiven Weg dar, die Verweildauer des Liganden innerhalb einer Bindetasche zu erhöhen. In dieser Arbeit wurden theoretische Methoden angewendet, um die Reaktivität und den nichtkovalenten Zustand vor der Reaktion zu modellieren.

Im Rahmen einer Fallstudie zu Cathepsin K wurden nichtkovalente Modelle von kovalenten Inhibitoren generiert. Für verschiedene Komplexe aus Cathepsin K und einem kovalent gebundenem Liganden wurde der Zustand vor der Reaktion modelliert und dessen Stabilität im Rahmen einer klassischen MD-Simulation überprüft. In vielen Fällen konnte in der Simulation eine stabile Positionierung des Warheads in direkter Nachbarschaft zur katalytischen Einheit beobachtet werden, die über mehrere hundert Nanosekunden aufrecht erhalten blieb. Die Stabilität dieser Positionierung hing hauptsächlich vom gewählten Protonierungszustand der katalytischen Aminosäuren ab: der ungeladene Zustand erlaubte ein Netzwerk aus Wasserstoffbrückenbindungen zwischen einem Wassermolekül, dem Warhead und der katalytischen Einheit, während im zwitterionischen Zustand deutlich mehr Fluktuationen beobachtbar waren. Dieselben Beobachtungen konnten für das verwandte Cathepsin S gemacht werden.

Im Rahmen eines Wirkstoffdesign-Projektes zur Deubiquitinase USP28 wurden von unpublizierten Kristallstrukturen ausgehend erste Docking-Experimente durchgeführt. Es konnte gezeigt werden, dass ein literaturbekannter Inhibitor von USP28 mit einem Warhead so modifiziert werden kann, dass die reaktive Einheit in direkter Nachbarschaft zu einem Cystein positioniert wird. Für diese Warheads wurden ebenfalls quantenmechanische Rechnungen zur Bestimmung der Aktivierungsenergie durchgeführt. So kann abgeschätzt werden, wie stark sich experimentelle Geschwindigkeitskonstanten aufgrund unterschiedlicher Substitutionen verändern, und wie das Design eines kovalenten Inhibitors rational verbessert werden kann. Für eine Reihe von Inhibitoren der ChlaDUB1 wurde ein Protokoll aus quantenmechanischen Rechnungen genutzt, um die Reaktivität verschiedener Warheads abzuschätzen. Die erhaltenen Aktivierungsenergien korrelierten mit experimentell bestimmten Raten zur Inaktivierung des Enzyms. Es konnten dann deutliche Abweichungen zum berechneten Trend beobachtet werden, wenn innerhalb der Warhead-Struktur geometrische Änderungen auftraten, wie z. B. sterisch anspruchsvolle Sulfonamide anstelle von Amid.

Um besser nachvollziehen zu können, warum bei einem Photoswitch-Inhibitor der Butyrylcholin-

---

Esterase der *cis*-Zustand des Moleküls besser inhibiert als der *trans*-Zustand, wurde eine Docking-Studie des Zustandes vor der Reaktion durchgeführt. Für diese Studie war es von großer Bedeutung, sowohl die unterschiedlichen Ringkonformere des Liganden als auch die stereochemischen Besonderheiten des Photoswitches zu beachten und angemessen im Docking widerzuspiegeln. Es konnte ein qualitatives Modell aufgestellt werden, das zeigt, dass der *trans*-Zustand aufgrund seiner längeren Form mit wichtigen Aminosäuren am Eingang der Bindungstasche kollidiert.

Zusammenfassend konnte gezeigt werden, dass durch die Modellierung des nichtkovalenten Zustandes eines kovalenten Inhibitors bereits signifikante Erkenntnisse über die Struktur-Wirkung-Beziehung und für die Optimierung der Verbindung gewonnen werden können. Docking-Studien, Molekulardynamik-Simulationen und quantenmechanische Berechnungen sind, solange ihre jeweiligen Limitationen beachtet werden, nützliche Werkzeuge, um jeden Punkt entlang der Reaktionskoordinate darzustellen.

## 6. Summary

Covalent inhibition is an effective way to increase the residence time of a ligand within the active site. In this work theoretical methods were used to model the reactivity and the noncovalent pre-reaction state.

Noncovalent models of covalent inhibitors were generated as part of a case study of Cathepsin K. Several complexes of Cathepsin K and a covalently bound ligand were modeled in their state before the reaction, and their stability was assessed by classical molecular dynamics simulations. In most cases the warhead was positioned in close proximity to the catalytic unit, remaining there for up to several hundred nanoseconds. This stable positioning was largely dependent on the protonation state of the catalytic amino acids: In the uncharged state allowed a network of hydrogen bonds between the warhead, a water molecule and the catalytic site, while the charged state caused significantly more fluctuations. The same observations could be made for Cathepsin S.

Using unpublished crystal structures, first design steps for the inhibition of the deubiquitinase USP28 were performed. Docking studies showed that modification of a literature-known inhibitor of USP28 with a warhead allowed to place this reactive unit close to a cysteine. Activation energies were also obtained for these structures via quantum mechanical calculations. This allows to estimate how drastically rate constants change upon different warhead substitutions, and how the rational design of inhibitors can be improved.

To estimate the reactivity of a series of ChlaDUB1 inhibitors, a protocol of quantum mechanical calculations was adapted. The obtained activation energies correlated with experimentally obtained rate constants of enzyme inactivation. Deviation from the observed trend occurred if the geometrical structure of the warhead differed significantly, e. g. if a sulfone amide is integrated into the structure instead of an amide.

To better rationalize the differences in inhibition between the *cis*- and *trans*-state of a photoswitch inhibitor of butyrylcholine esterase, a docking study of the noncovalent state was performed. The different ring conformers and stereochemical properties of the photoswitch were critical for a sensible model of the ligand. A qualitative model could be obtained which explains that the *cis*-isomer is more active than the *trans*-isomer due to a steric clash of the latter with amino acids at the entrance of the pocket.

Therefore it could be demonstrated that the modeling of the noncovalent state of a cova-

---

lent inhibitor already allows significant insights into the structure-activity-relationship and into possible paths for optimization. Docking studies, molecular dynamics simulations and quantum mechanical calculations are useful tools to model any point along the reaction coordinate, as long as their respective limitations are kept in mind.

## Literaturverzeichnis

- <sup>1</sup> Warren DeLano, Schrödinger, LLC. Pymol, 2023.
- <sup>2</sup> Lydia Boike, Nathaniel J. Henning, and Daniel K. Nomura. Advances in covalent drug discovery. *Nature reviews. Drug discovery*, 2022.
- <sup>3</sup> J. R. Vane. Inhibition of prostaglandin synthesis as a mechanism of action for aspirin-like drugs. *Nature: New biology*, 231(25):232–235, 1971.
- <sup>4</sup> Anna L. Blobaum and Lawrence J. Marnett. Structural and functional basis of cyclooxygenase inhibition. *Journal of medicinal chemistry*, 50(7):1425–1441, 2007.
- <sup>5</sup> G. J. Roth and C. J. Siok. Acetylation of the nh2-terminal serine of prostaglandin synthetase by aspirin. *Journal of Biological Chemistry*, 253(11):3782–3784, 1978.
- <sup>6</sup> Karen Bush and Patricia A. Bradford.  $\beta$ -lactams and  $\beta$ -lactamase inhibitors: An overview. *Cold Spring Harbor perspectives in medicine*, 6(8), 2016.
- <sup>7</sup> Christoph Sotriffer. Docking of covalent ligands: Challenges and approaches. *Molecular informatics*, 37(9-10):e1800062, 2018.
- <sup>8</sup> M. Raymond V. Finlay, Mark Anderton, Susan Ashton, Peter Ballard, Paul A. Bethel, Matthew R. Box, Robert H. Bradbury, Simon J. Brown, Sam Butterworth, Andrew Campbell, Christopher Chorley, Nicola Colclough, Darren A. E. Cross, Gordon S. Currie, Matthew Grist, Lorraine Hassall, George B. Hill, Daniel James, Michael James, Paul Kemmitt, Teresa Klinowska, Gillian Lamont, Scott G. Lamont, Nathaniel Martin, Heather L. McFarland, Martine J. Mellor, Jonathon P. Orme, David Perkins, Paula Perkins, Graham Richmond, Peter Smith, Richard A. Ward, Michael J. Waring, David Whittaker, Stuart Wells, and Gail L. Wrigley. Discovery of a potent and selective egfr inhibitor (azd9291) of both sensitizing and t790m resistance mutations that spares the wild type form of the receptor. *Journal of medicinal chemistry*, 57(20):8249–8267, 2014.
- <sup>9</sup> Darren A. E. Cross, Susan E. Ashton, Serban Ghiorghiu, Cath Eberlein, Caroline A. Nebhan, Paula J. Spitzler, Jonathon P. Orme, M. Raymond V. Finlay, Richard A. Ward,



- Martine J. Mellor, Gareth Hughes, Amar Rahi, Vivien N. Jacobs, Monica Red Brewer, Eiki Ichihara, Jing Sun, Hailing Jin, Peter Ballard, Katherine Al-Kadhimi, Rachel Rowlinson, Teresa Klinowska, Graham H. P. Richmond, Mireille Cantarini, Dong-Wan Kim, Malcolm R. Ranson, and William Pao. Azd9291, an irreversible egfr tki, overcomes t790m-mediated resistance to egfr inhibitors in lung cancer. *Cancer discovery*, 4(9):1046–1061, 2014.
- <sup>10</sup> Tony S. Mok, Yi-Long Wu, Myung-Ju Ahn, Marina C. Garassino, Hye R. Kim, Suresh S. Ramalingam, Frances A. Shepherd, Yong He, Hiroaki Akamatsu, Willemijn S. M. E. Theelen, Chee K. Lee, Martin Sebastian, Alison Templeton, Helen Mann, Marcelo Marotti, Serban Ghiorghiu, and Vassiliki A. Papadimitrakopoulou. Osimertinib or platinum-pemetrexed in egfr t790m-positive lung cancer. *The New England journal of medicine*, 376(7):629–640, 2017.
- <sup>11</sup> Zhengying Pan, Heleen Scheerens, Shyr-Jiann Li, Brian E. Schultz, Paul A. Sprengeler, L. Chuck Burrill, Rohan V. Mendonca, Michael D. Sweeney, Keana C. K. Scott, Paul G. Grothaus, Douglas A. Jeffery, Jill M. Spoerke, Lee A. Honigberg, Peter R. Young, Stacie A. Dalrymple, and James T. Palmer. Discovery of selective irreversible inhibitors for bruton’s tyrosine kinase. *ChemMedChem*, 2(1):58–61, 2007.
- <sup>12</sup> Lee A. Honigberg, Ashley M. Smith, Mint Sirisawad, Erik Verner, David Loury, Betty Chang, Shyr Li, Zhengying Pan, Douglas H. Thamm, Richard A. Miller, and Joseph J. Buggy. The bruton tyrosine kinase inhibitor pci-32765 blocks b-cell activation and is efficacious in models of autoimmune disease and b-cell malignancy. *Proceedings of the National Academy of Sciences of the United States of America*, 107(29):13075–13080, 2010.
- <sup>13</sup> Shan Wang, Yu Tian, Min Wang, Gui-Bo Sun, and Xiao-Bo Sun. Advanced activity-based protein profiling application strategies for drug development. *Frontiers in pharmacology*, 9:353, 2018.
- <sup>14</sup> Matthias Gehringer and Stefan A. Laufer. Emerging and re-emerging warheads for targeted covalent inhibitors: Applications in medicinal chemistry and chemical biology. *Journal of medicinal chemistry*, 62(12):5673–5724, 2019.
- <sup>15</sup> H. R. Tsou, N. Mamuya, B. D. Johnson, M. F. Reich, B. C. Gruber, F. Ye, R. Nilakantan, R. Shen, C. Discafani, R. DeBlanc, R. Davis, F. E. Koehn, L. M. Greenberger, Y. F. Wang, and A. Wissner. 6-substituted-4-(3-bromophenylamino)quinazolines as putative

- irreversible inhibitors of the epidermal growth factor receptor (egfr) and human epidermal growth factor receptor (her-2) tyrosine kinases with enhanced antitumor activity. *Journal of medicinal chemistry*, 44(17):2719–2734, 2001.
- <sup>16</sup> Allan Wissner, Elsebe Overbeek, Marvin F. Reich, M. Brawner Floyd, Bernard D. Johnson, Nellie Mamuya, Edward C. Rosfjord, Carolyn Discafani, Rachel Davis, Xiaoqing Shi, Sridhar K. Rabindran, Brian C. Gruber, Fei Ye, William A. Hallett, Ramaswamy Nilakantan, Ru Shen, Yu-Fen Wang, Lee M. Greenberger, and Hwei-Ru Tsou. Synthesis and structure-activity relationships of 6,7-disubstituted 4-anilinoquinoline-3-carbonitriles. the design of an orally active, irreversible inhibitor of the tyrosine kinase activity of the epidermal growth factor receptor (egfr) and the human epidermal growth factor receptor-2 (her-2). *Journal of medicinal chemistry*, 46(1):49–63, 2003.
- <sup>17</sup> Edgar R. Wood, Lisa M. Shewchuk, Byron Ellis, Perry Brignola, Ronald L. Brashear, Thomas R. Caferro, Scott H. Dickerson, Hamilton D. Dickson, Kelly H. Donaldson, Michael Gaul, Robert J. Griffin, Anne M. Hassell, Barry Keith, Robert Mullin, Kimberly G. Petrov, Michael J. Reno, David W. Rusnak, Sarva M. Tadepalli, John C. Ulrich, Craig D. Wagner, Dana E. Vanderwall, Alex G. Waterson, Jon D. Williams, Wendy L. White, and David E. Uehling. 6-ethynylthieno3,2-d- and 6-ethynylthieno2,3-dpyrimidin-4-anilines as tunable covalent modifiers of erbb kinases. *Proceedings of the National Academy of Sciences of the United States of America*, 105(8):2773–2778, 2008.
- <sup>18</sup> Jin Han, Santosh L. Saraf, and Victor R. Gordeuk. Systematic review of voxelotor: A first-in-class sickle hemoglobin polymerization inhibitor for management of sickle cell disease. *Pharmacotherapy*, 40(6):525–534, 2020.
- <sup>19</sup> Robert A. Copeland. *Evaluation of enzyme inhibitors in drug discovery: A guide for medicinal chemists and pharmacologists*. Wiley, Hoboken, N.J, second edition edition, 2013.
- <sup>20</sup> John M. Strelow. A perspective on the kinetics of covalent and irreversible inhibition. *SLAS discovery : advancing life sciences R & D*, 22(1):3–20, 2017.
- <sup>21</sup> Matthias Gehring. Covalent inhibitors: back on track? *Future medicinal chemistry*, 12(15):1363–1368, 2020.
- <sup>22</sup> G. Jones, P. Willett, R. C. Glen, A. R. Leach, and R. Taylor. Development and validation of a genetic algorithm for flexible docking. *Journal of molecular biology*, 267(3):727–748, 1997.

- <sup>23</sup> Andrea Scarpino, György G. Ferenczy, and György M. Keserű. Comparative evaluation of covalent docking tools. *Journal of Chemical Information and Modeling*, 58(7):1441–1458, 2018.
- <sup>24</sup> Kai Zhu, Kenneth W. Borrelli, Jeremy R. Greenwood, Tyler Day, Robert Abel, Ramy S. Farid, and Edward Harder. Docking covalent inhibitors: a parameter free approach to pose prediction and scoring. *Journal of Chemical Information and Modeling*, 54(7):1932–1940, 2014.
- <sup>25</sup> James B. Foresman and Aeleen Frisch. *Exploring chemistry with electronic structure methods*. Gaussian Inc, Wallingford, CT USA, third edition edition, October 2015.
- <sup>26</sup> Ferruccio Palazzesi, Marc A. Grundl, Alexander Pautsch, Alexander Weber, and Christopher S. Tautermann. A fast ab initio predictor tool for covalent reactivity estimation of acrylamides. *Journal of Chemical Information and Modeling*, 59(8):3565–3571, 2019.
- <sup>27</sup> Christopher S. Gondi and Jasti S. Rao. Cathepsin b as a cancer target. *Expert opinion on therapeutic targets*, 17(3):281–291, 2013.
- <sup>28</sup> Dieter Brömme and Fabien Lecaille. Cathepsin k inhibitors for osteoporosis and potential off-target effects. *Expert opinion on investigational drugs*, 18(5):585–600, 2009.
- <sup>29</sup> Asher Mullard. Merck &co. drops osteoporosis drug odanacatib. *Nature reviews. Drug discovery*, 15(10):669, 2016.
- <sup>30</sup> Jakub Benýšek, Michal Buša, Petra Rubešová, Jindřich Fanfrlík, Martin Lepšík, Jiří Brynda, Zuzana Matoušková, Ulrike Bartz, Martin Horn, Michael Gütschow, and Michael Mareš. Highly potent inhibitors of cathepsin k with a differently positioned cyanohydrazide warhead: structural analysis of binding mode to mature and zymogen-like enzymes. *Journal of enzyme inhibition and medicinal chemistry*, 37(1):515–526, 2022.
- <sup>31</sup> Igor José Dos Santos Nascimento, Thiago Mendonça de Aquino, and Edeildo Ferreira Da Silva-Júnior. Cruzain and rhodesain inhibitors: Last decade of advances in seeking for new compounds against american and african trypanosomiasis. *Current topics in medicinal chemistry*, 21(21):1871–1899, 2021.
- <sup>32</sup> Donghyuk Shin, Rukmini Mukherjee, Diana Grewe, Denisa Bojkova, Kheewoong Baek, Anshu Bhattacharya, Laura Schulz, Marek Widera, Ahmad Reza Mehdipour, Georg Tasscher, Paul P. Geurink, Alexander Wilhelm, Gerbrand J. van Heden Noort, Huib Ovaa,

- Stefan Müller, Klaus-Peter Knobeloch, Krishnaraj Rajalingam, Brenda A. Schulman, Jindrich Cinatl, Gerhard Hummer, Sandra Ciesek, and Ivan Dikic. Papain-like protease regulates sars-cov-2 viral spread and innate immunity. *Nature*, 587(7835):657–662, 2020.
- <sup>33</sup> Shuhua Ma, Lakshmi S. Devi-Kesavan, and Jiali Gao. Molecular dynamics simulations of the catalytic pathway of a cysteine protease: a combined qm/mm study of human cathepsin k. *Journal of the American Chemical Society*, 129(44):13633–13645, 2007.
- <sup>34</sup> Syeed Hussain, Akavish Khan, Sheraz Gul, Marina Resmini, Chandra S. Verma, Emrys W. Thomas, and Keith Brocklehurst. Identification of interactions involved in the generation of nucleophilic reactivity and of catalytic competence in the catalytic site cys/his ion pair of papain. *Biochemistry*, 50(49):10732–10742, 2011.
- <sup>35</sup> Alexander G. Dossetter, Howard Beeley, Jonathan Bowyer, Calum R. Cook, James J. Crawford, Jonathan E. Finlayson, Nicola M. Heron, Christine Heyes, Adrian J. Highton, Julian A. Hudson, Anja Jestel, Peter W. Kenny, Stephan Krapp, Scott Martin, Philip A. MacFaul, Thomas M. McGuire, Pablo Morentin Gutierrez, Andrew D. Morley, Jeffrey J. Morris, Ken M. Page, Lyn Rosenbrier Ribeiro, Helen Sawney, Stefan Steinbacher, Caroline Smith, and Madeleine Vickers. (1r,2r)-n-(1-cyanocyclopropyl)-2-(6-methoxy-1,3,4,5-tetrahydropyrido4,3-bindole-2-carbonyl)cyclohexanecarboxamide (azd4996): a potent and highly selective cathepsin k inhibitor for the treatment of osteoarthritis. *Journal of medicinal chemistry*, 55(14):6363–6374, 2012.
- <sup>36</sup> M. E. McGrath, J. L. Klaus, M. G. Barnes, and D. Brömme. Crystal structure of human cathepsin k complexed with a potent inhibitor. *Nature structural biology*, 4(2):105–109, 1997.
- <sup>37</sup> Gerald R. Grimsley, J. Martin Scholtz, and C. Nick Pace. A summary of the measured pk values of the ionizable groups in folded proteins. *Protein science : a publication of the Protein Society*, 18(1):247–251, 2009.
- <sup>38</sup> M. Olivia Kim and J. Andrew McCammon. Computation of ph-dependent binding free energies. *Biopolymers*, 105(1):43–49, 2016.
- <sup>39</sup> L. Polgár. On the mode of activation of the catalytically essential sulfhydryl group of papain. *European journal of biochemistry*, 33(1):104–109, 1973.
- <sup>40</sup> S. Pinitglang, A. B. Watts, M. Patel, J. D. Reid, M. A. Noble, S. Gul, A. Bokth, A. Naeem, H. Patel, E. W. Thomas, S. K. Sreedharan, C. Verma, and K. Brocklehurst. A classical

- enzyme active center motif lacks catalytic competence until modulated electrostatically. *Biochemistry*, 36(33):9968–9982, 1997.
- <sup>41</sup> Keith Brocklehurst and Mike P. Philpott. Cysteine proteases: mode of action and role in epidermal differentiation. *Cell and tissue research*, 351(2):237–244, 2013.
- <sup>42</sup> Brigitta Elsässer, Florian B. Zauner, Johann Messner, Wai Tuck Soh, Elfriede Dall, and Hans Brandstetter. Distinct roles of catalytic cysteine and histidine in the protease and ligase mechanisms of human legumain as revealed by dft-based qm/mm simulations. *ACS catalysis*, 7(9):5585–5593, 2017.
- <sup>43</sup> Cécilia Hognon, Marco Marazzi, and Cristina García-Iriepa. Atomistic-level description of the covalent inhibition of sars-cov-2 papain-like protease. *International journal of molecular sciences*, 23(10), 2022.
- <sup>44</sup> Michael Shokhen, Netaly Khazanov, and Amnon Albeck. Challenging a paradigm: theoretical calculations of the protonation state of the cys25-his159 catalytic diad in free papain. *Proteins*, 77(4):916–926, 2009.
- <sup>45</sup> Michael Shokhen, Netaly Khazanov, and Amnon Albeck. The mechanism of papain inhibition by peptidyl aldehydes. *Proteins*, 79(3):975–985, 2011.
- <sup>46</sup> Milena Mladenovic, Reinhold F. Fink, Walter Thiel, Tanja Schirmeister, and Bernd Engels. On the origin of the stabilization of the zwitterionic resting state of cysteine proteases: a theoretical study. *Journal of the American Chemical Society*, 130(27):8696–8705, 2008.
- <sup>47</sup> Florian Hofer, Johannes Kraml, Ursula Kahler, Anna S. Kamenik, and Klaus R. Liedl. Catalytic site pka values of aspartic, cysteine, and serine proteases: Constant ph md simulations. *Journal of Chemical Information and Modeling*, 60(6):3030–3042, 2020.
- <sup>48</sup> Robert C. Harris, Ruibin Liu, and Jana Shen. Predicting reactive cysteines with implicit-solvent-based continuous constant ph molecular dynamics in amber. *Journal of chemical theory and computation*, 16(6):3689–3698, 2020.
- <sup>49</sup> Sören Rindfleisch, Matthias Krull, Jon Uranga, Tobias Schmidt, Fabian Rabe von Papenheim, Laura Liliana Kirck, Angeliki Balouri, Thomas Schneider, Ashwin Chari, Ronald Kluger, Gleb Bourenkov, Ulf Diederichsen, Ricardo A. Mata, and Kai Tittmann. Ground-state destabilization by electrostatic repulsion is not a driving force in orotidine-5′-monophosphate decarboxylase catalysis. *Nature Catalysis*, 5(4):332–341, 2022.

- <sup>50</sup> Matthias Hoffmann, Carina Stiller, Erik Endres, Matthias Scheiner, Sandra Gunesch, Christoph Sotriffer, Tangui Maurice, and Michael Decker. Highly selective butyrylcholinesterase inhibitors with tunable duration of action by chemical modification of transferable carbamate units exhibit pronounced neuroprotective effect in an alzheimer's disease mouse model. *Journal of medicinal chemistry*, 62(20):9116–9140, 2019.
- <sup>51</sup> Zeinab Breijyeh and Rafik Karaman. Comprehensive review on alzheimer's disease: Causes and treatment. *Molecules (Basel, Switzerland)*, 25(24), 2020.
- <sup>52</sup> John Hardy. The amyloid hypothesis for alzheimer's disease: a critical reappraisal. *Journal of neurochemistry*, 110(4):1129–1134, 2009.
- <sup>53</sup> Miroslav Pohanka. Alzheimer's disease and oxidative stress: a review. *Current medicinal chemistry*, 21(3):356–364, 2014.
- <sup>54</sup> Heiko Braak and Kelly Del Tredici. Are cases with tau pathology occurring in the absence of  $\text{a}\beta$  deposits part of the ad-related pathological process? *Acta neuropathologica*, 128(6):767–772, 2014.
- <sup>55</sup> Harald Hampel, M-Marsel Mesulam, A. Claudio Cuello, Martin R. Farlow, Ezio Giacobini, George T. Grossberg, Ara S. Khachaturian, Andrea Vergallo, Enrica Cavedo, Peter J. Snyder, and Zaven S. Khachaturian. The cholinergic system in the pathophysiology and treatment of alzheimer's disease. *Brain : a journal of neurology*, 141(7):1917–1933, 2018.
- <sup>56</sup> H. Hampel, M-M Mesulam, A. C. Cuello, A. S. Khachaturian, A. Vergallo, M. R. Farlow, P. J. Snyder, E. Giacobini, and Z. S. Khachaturian. Revisiting the cholinergic hypothesis in alzheimer's disease: Emerging evidence from translational and clinical research. *The journal of prevention of Alzheimer's disease*, 6(1):2–15, 2019.
- <sup>57</sup> Terrone L. Rosenberry, Xavier Brazzolotto, Ian R. Macdonald, Marielle Wandhammer, Marie Trovaslet-Leroy, Sultan Darvesh, and Florian Nachon. Comparison of the binding of reversible inhibitors to human butyrylcholinesterase and acetylcholinesterase: A crystallographic, kinetic and calorimetric study. *Molecules (Basel, Switzerland)*, 22(12), 2017.
- <sup>58</sup> B. Li, J. A. Stribley, A. Ticu, W. Xie, L. M. Schopfer, P. Hammond, S. Brimijoin, S. H. Hinrichs, and O. Lockridge. Abundant tissue butyrylcholinesterase and its possible function in the acetylcholinesterase knockout mouse. *Journal of neurochemistry*, 75(3):1320–1331, 2000.

- <sup>59</sup> M-M Mesulam, A. Guillozet, P. Shaw, A. Levey, E. G. Duysen, and O. Lockridge. Acetylcholinesterase knockouts establish central cholinergic pathways and can use butyrylcholinesterase to hydrolyze acetylcholine. *Neuroscience*, 110(4):627–639, 2002.
- <sup>60</sup> AnŽe Meden, Damijan Knez, Marko Jukič, Xavier Brazzolotto, Marija Gršič, Anja Pišlar, Abida Zahirović, Janko Kos, Florian Nachon, Jurij Svete, Stanislav Gobec, and Uroš Grošelj. Tryptophan-derived butyrylcholinesterase inhibitors as promising leads against alzheimer’s disease. *Chemical communications (Cambridge, England)*, 55(26):3765–3768, 2019.
- <sup>61</sup> Jonah Cheung, Michael J. Rudolph, Fiana Burshteyn, Michael S. Cassidy, Ebony N. Gary, James Love, Matthew C. Franklin, and Jude J. Height. Structures of human acetylcholinesterase in complex with pharmacologically important ligands. *Journal of medicinal chemistry*, 55(22):10282–10286, 2012.
- <sup>62</sup> Tarek Mohamed, Wesseem Osman, Gary Tin, and Praveen P. N. Rao. Selective inhibition of human acetylcholinesterase by xanthine derivatives: in vitro inhibition and molecular modeling investigations. *Bioorganic & medicinal chemistry letters*, 23(15):4336–4341, 2013.
- <sup>63</sup> Sulunay Parlar, Gulsah Bayraktar, Ayse Hande Tarikogullari, Vildan Alptüzün, and Er-cin Erciyas. Synthesis, biological evaluation and molecular docking study of hydrazone-containing pyridinium salts as cholinesterase inhibitors. *Chemical & pharmaceutical bulletin*, 64(9):1281–1287, 2016.
- <sup>64</sup> Maria Laura Bolognesi, Manuela Bartolini, Andrea Cavalli, Vincenza Andrisano, Michela Rosini, Anna Minarini, and Carlo Melchiorre. Design, synthesis, and biological evaluation of conformationally restricted rivastigmine analogues. *Journal of medicinal chemistry*, 47(24):5945–5952, 2004.
- <sup>65</sup> Diego Alejandro Rodríguez-Soacha and Michael Decker. Photopharmacology in alzheimer’s disease. *Advanced Therapeutics*, 1(3), 2018.
- <sup>66</sup> Katharina Hüll, Johannes Morstein, and Dirk Trauner. In vivo photopharmacology. *Chemical reviews*, 118(21):10710–10747, 2018.
- <sup>67</sup> Anna-Lena Leistner, Susanne Kirchner, Johannes Karcher, Tobias Bantle, Mariam L. Schulte, Peter Gödtel, Christian Fengler, and Zbigniew L. Pianowski. Fluorinated azo-

- benzenes switchable with red light. *Chemistry (Weinheim an der Bergstrasse, Germany)*, 27(31):8094–8099, 2021.
- <sup>68</sup> Matthias Scheiner, Alexandra Sink, Matthias Hoffmann, Cassandre Vrigneau, Erik Endres, Allison Carles, Christoph Sotriffer, Tanguy Maurice, and Michael Decker. Photoswitchable pseudoirreversible butyrylcholinesterase inhibitors allow optical control of inhibition in vitro and enable restoration of cognition in an alzheimer’s disease mouse model upon irradiation. *Journal of the American Chemical Society*, 144(7):3279–3284, 2022.
- <sup>69</sup> Fumiaki Ohtake and Hikaru Tsuchiya. The emerging complexity of ubiquitin architecture. *Journal of biochemistry*, 161(2):125–133, 2017.
- <sup>70</sup> Jonathan D. Wrigley, Gerald Gavory, Iain Simpson, Marian Preston, Helen Plant, Jenna Bradley, Anne U. Goeppert, Ewelina Rozycka, Gareth Davies, Jarrod Walsh, Andrea Valentine, Keeva McClelland, Krzysztofa Ewa Odrzywol, Jonathan Renshaw, Joanna Boros, Jonathan Tart, Lindsey Leach, Thorsten Nowak, Richard A. Ward, Timothy Harrison, and David M. Andrews. Identification and characterization of dual inhibitors of the usp25/28 deubiquitinating enzyme subfamily. *ACS chemical biology*, 12(12):3113–3125, 2017.
- <sup>71</sup> Cristian Prieto-Garcia, Oliver Hartmann, Michaela Reissland, Fabian Braun, Thomas Fischer, Susanne Walz, Christina Schülein-Völk, Ursula Eilers, Carsten P. Ade, Marco A. Calzado, Amir Orian, Hans M. Maric, Christian Münch, Mathias Rosenfeldt, Martin Eilers, and Markus E. Diefenbacher. Maintaining protein stability of  $\Delta$ np63 via usp28 is required by squamous cancer cells. *EMBO molecular medicine*, 12(4):e11101, 2020.
- <sup>72</sup> Markus E. Diefenbacher, Atanu Chakraborty, Sophia M. Blake, Richard Mitter, Nikita Popov, Martin Eilers, and Axel Behrens. Usp28 counteracts flw7 in intestinal homeostasis and cancer. *Cancer research*, 75(7):1181–1186, 2015.
- <sup>73</sup> Markus E. Diefenbacher, Nikita Popov, Sophia M. Blake, Christina Schülein-Völk, Emma Nye, Bradley Spencer-Dene, Laura A. Jaenicke, Martin Eilers, and Axel Behrens. The deubiquitinase usp28 controls intestinal homeostasis and promotes colorectal cancer. *The Journal of clinical investigation*, 124(8):3407–3418, 2014.
- <sup>74</sup> Nikita Popov, Steffi Herold, Maria Llamazares, Christina Schülein, and Martin Eilers. Fbw7 and usp28 regulate myc protein stability in response to dna damage. *Cell cycle (Georgetown, Tex.)*, 6(19):2327–2331, 2007.



- <sup>75</sup> Christina Schüle-Völk, Elmar Wolf, Jing Zhu, Wenshan Xu, Lyudmyla Taranets, Andreas Hellmann, Laura A. Jänicke, Markus E. Diefenbacher, Axel Behrens, Martin Eilers, and Nikita Popov. Dual regulation of flw7 function and oncogenic transformation by usp28. *Cell reports*, 9(3):1099–1109, 2014.
- <sup>76</sup> Yadi Wu, Yifan Wang, Xiuwei H. Yang, Tiebang Kang, Yongxiang Zhao, Chi Wang, B. Mark Evers, and Binhua P. Zhou. The deubiquitinase usp28 stabilizes lsd1 and confers stem-cell-like traits to breast cancer cells. *Cell reports*, 5(1):224–236, 2013.
- <sup>77</sup> Malte Gersch, Jane L. Wagstaff, Angela V. Toms, Bradford Graves, Stefan M.V. Freund, and David Komander. Distinct usp25 and usp28 oligomerization states regulate deubiquitinating activity. *Molecular Cell*, 74(3):436–451.e7, 2019.
- <sup>78</sup> Florian Sauer, Theresa Klemm, Ravi B. Kollampally, Ingrid Tessmer, Radhika K. Nair, Nikita Popov, and Caroline Kisker. Differential oligomerization of the deubiquitinases usp25 and usp28 regulates their activities. *Molecular Cell*, 74(3):421–435.e10, 2019.
- <sup>79</sup> Hui Wang, Qian Meng, Yiluan Ding, Muya Xiong, Mengying Zhu, Yuanyuan Yang, Haixia Su, Lei Gu, Yechun Xu, Li Shi, Hu Zhou, and Naixia Zhang. Usp28 and usp25 are downregulated by vismodegib in vitro and in colorectal cancer cell lines. *The FEBS journal*, 288(4):1325–1342, 2021.
- <sup>80</sup> Di Zhou, Zhuo Xu, Yaodong Huang, Hui Wang, Xiaoli Zhu, Wentao Zhang, Weiwei Song, Tong Gao, Tongchao Liu, Meng Wang, Li Shi, Naixia Zhang, and Bing Xiong. Structure-based discovery of potent usp28 inhibitors derived from vismodegib. *European journal of medicinal chemistry*, 254:115369, 2023.
- <sup>81</sup> Zhenzhen Liu, Taoqian Zhao, Zhonghua Li, Kai Sun, Yundong Fu, Ting Cheng, Jimin Guo, Bin Yu, Xiaojing Shi, and Hongmin Liu. Discovery of 1,2,3-triazolo[4,5-d]pyrimidine derivatives as highly potent, selective, and cellularly active usp28 inhibitors. *Acta pharmaceutica Sinica. B*, 10(8):1476–1491, 2020.
- <sup>82</sup> E. Josue Ruiz, Adan Pinto-Fernandez, Andrew P. Turnbull, Linxiang Lan, Thomas M. Charlton, Hannah C. Scott, Andreas Damianou, George Vere, Eva M. Riising, Clive Da Costa, Wojciech W. Krajewski, David Guerin, Jeffrey D. Kearns, Stephanos Ioannidis, Marie Katz, Crystal McKinnon, Jonathan O’Connell, Natalia Moncaut, Ian Rosewell, Emma Nye, Neil Jones, Claire Heride, Malte Gersch, Min Wu, Christopher J. Dinsmore, Tim R. Hammonds, Sunkyu Kim, David Komander, Sylvie Urbe, Michael J. Clague,

- Benedikt M. Kessler, and Axel Behrens. Usp28 deletion and small-molecule inhibition destabilizes c-myc and elicits regression of squamous cell lung carcinoma. *eLife*, 10, 2021.
- <sup>83</sup> Lori Newman, Jane Rowley, Stephen Vander Hoorn, Nalinka Saman Wijesooriya, Magnus Unemo, Nicola Low, Gretchen Stevens, Sami Gottlieb, James Kiarie, and Marleen Temmerman. Global estimates of the prevalence and incidence of four curable sexually transmitted infections in 2012 based on systematic review and global reporting. *PloS one*, 10(12):e0143304, 2015.
- <sup>84</sup> Jane Rowley, Stephen Vander Hoorn, Eline Korenromp, Nicola Low, Magnus Unemo, Laith J. Abu-Raddad, R. Matthew Chico, Alex Smolak, Lori Newman, Sami Gottlieb, Soe Soe Thwin, Nathalie Broutet, and Melanie M. Taylor. Chlamydia, gonorrhoea, trichomoniasis and syphilis: global prevalence and incidence estimates, 2016. *Bulletin of the World Health Organization*, 97(8):548–562P, 2019.
- <sup>85</sup> Krishnaraj Rajalingam, Manu Sharma, Christine Lohmann, Monique Oswald, Oliver Thieck, Christopher J. Froelich, and Thomas Rudel. Mcl-1 is a key regulator of apoptosis resistance in chlamydia trachomatis-infected cells. *PloS one*, 3(9):e3102, 2008.
- <sup>86</sup> Jonathan N. Pruneda, Charlotte H. Durkin, Paul P. Geurink, Huib Ovaa, Balaji Sathanam, David W. Holden, and David Komander. The molecular basis for ubiquitin and ubiquitin-like specificities in bacterial effector proteases. *Molecular Cell*, 63(2):261–276, 2016.
- <sup>87</sup> Lisa E. Manhart, Catherine W. Gillespie, M. Sylvan Lowens, Christine M. Khosropour, Danny V. Colombara, Matthew R. Golden, Navneet R. Hakhu, Katherine K. Thomas, James P. Hughes, Nicole L. Jensen, and Patricia A. Totten. Standard treatment regimens for nongonococcal urethritis have similar but declining cure rates: a randomized controlled trial. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 56(7):934–942, 2013.
- <sup>88</sup> Yesid A. Ramirez, Thomas B. Adler, Eva Altmann, Theresa Klemm, Christian Tiesmeyer, Florian Sauer, Stefan G. Kathman, Alexander V. Statsyuk, Christoph Sotriffer, and Caroline Kisker. Structural basis of substrate recognition and covalent inhibition of *cdul* from chlamydia trachomatis. *ChemMedChem*, 13(19):2014–2023, 2018.
- <sup>89</sup> Natasja Brooijmans and Irwin D. Kuntz. Molecular recognition and docking algorithms. *Annual review of biophysics and biomolecular structure*, 32:335–373, 2003.

- <sup>90</sup> Marcel L. Verdonk, Jason C. Cole, Michael J. Hartshorn, Christopher W. Murray, and Richard D. Taylor. Improved protein-ligand docking using gold. *Proteins*, 52(4):609–623, 2003.
- <sup>91</sup> Richard A. Friesner, Jay L. Banks, Robert B. Murphy, Thomas A. Halgren, Jasna J. Klicic, Daniel T. Mainz, Matthew P. Repasky, Eric H. Knoll, Mee Shelley, Jason K. Perry, David E. Shaw, Perry Francis, and Peter S. Shenkin. Glide: a new approach for rapid, accurate docking and scoring. 1. method and assessment of docking accuracy. *Journal of medicinal chemistry*, 47(7):1739–1749, 2004.
- <sup>92</sup> Thomas A. Halgren, Robert B. Murphy, Richard A. Friesner, Hege S. Beard, Leah L. Frye, W. Thomas Pollard, and Jay L. Banks. Glide: a new approach for rapid, accurate docking and scoring. 2. enrichment factors in database screening. *Journal of medicinal chemistry*, 47(7):1750–1759, 2004.
- <sup>93</sup> Garrett M. Morris, David S. Goodsell, Robert S. Halliday, Ruth Huey, William E. Hart, Richard K. Belew, and Arthur J. Olson. Automated docking using a lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry*, 19(14):1639–1662, 1998.
- <sup>94</sup> Garrett M. Morris, Ruth Huey, William Lindstrom, Michel F. Sanner, Richard K. Belew, David S. Goodsell, and Arthur J. Olson. Autodock4 and autodocktools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry*, 30(16):2785–2791, 2009.
- <sup>95</sup> A. Lavecchia and C. Di Giovanni. Virtual screening strategies in drug discovery: a critical review. *Current medicinal chemistry*, 20(23):2839–2860, 2013.
- <sup>96</sup> R. Rosenfeld, S. Vajda, and C. DeLisi. Flexible docking and design. *Annual review of biophysics and biomolecular structure*, 24:677–700, 1995.
- <sup>97</sup> Peter W. Atkins. *Physikalische Chemie*. VCH, Weinheim, 3., korr. Aufl., 1. nachdr edition, 2004.
- <sup>98</sup> Tiejun Cheng, Xun Li, Yan Li, Zhihai Liu, and Renxiao Wang. Comparative assessment of scoring functions on a diverse test set. *Journal of Chemical Information and Modeling*, 49(4):1079–1093, 2009.

- <sup>99</sup> Gerd Neudert and Gerhard Klebe. Dsx: a knowledge-based scoring function for the assessment of protein-ligand complexes. *Journal of Chemical Information and Modeling*, 51(10):2731–2745, 2011.
- <sup>100</sup> H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne. The protein data bank. *Nucleic acids research*, 28(1):235–242, 2000.
- <sup>101</sup> A. M. TURING. I.—computing machinery and intelligence. *Mind*, LIX(236):433–460, 1950.
- <sup>102</sup> Andrea Scarpino, László Petri, Damijan Knez, Tímea Imre, Péter Ábrányi-Balogh, György G. Ferenczy, Stanislav Gobec, and György M. Keserű. Widock: a reactive docking protocol for virtual screening of covalent inhibitors. *Journal of computer-aided molecular design*, 35(2):223–244, 2021.
- <sup>103</sup> Nir London, Rand M. Miller, Shyam Krishnan, Kenji Uchida, John J. Irwin, Oliv Eidam, Lucie Gibold, Peter Cimermančič, Richard Bonnet, Brian K. Shoichet, and Jack Taunton. Covalent docking of large libraries for the discovery of chemical probes. *Nature chemical biology*, 10(12):1066–1072, 2014.
- <sup>104</sup> Keriann M. Backus, Bruno E. Correia, Kenneth M. Lum, Stefano Forli, Benjamin D. Horning, Gonzalo E. González-Páez, Sandip Chatterjee, Bryan R. Lanning, John R. Teijaro, Arthur J. Olson, Dennis W. Wolan, and Benjamin F. Cravatt. Proteome-wide covalent ligand discovery in native biological systems. *Nature*, 534(7608):570–574, 2016.
- <sup>105</sup> Levi C. T. Pierce, Romelia Salomon-Ferrer, Cesar Augusto F de Oliveira, J. Andrew McCammon, and Ross C. Walker. Routine access to millisecond time scale events with accelerated molecular dynamics. *Journal of chemical theory and computation*, 8(9):2997–3002, 2012.
- <sup>106</sup> Zhifeng Jing, Chengwen Liu, Sara Y. Cheng, Rui Qi, Brandon D. Walker, Jean-Philip Piquemal, and Pengyu Ren. Polarizable force fields for biomolecular simulations: Recent advances and applications. *Annual review of biophysics*, 48:371–394, 2019.
- <sup>107</sup> Prajakta U. Kulkarni, Harshil Shah, and Vivek K. Vyas. Hybrid quantum mechanics/molecular mechanics (qm/mm) simulation: A tool for structure-based drug design and discovery. *Mini reviews in medicinal chemistry*, 22(8):1096–1107, 2022.

- <sup>108</sup> James A. Maier, Carmenza Martinez, Koushik Kasavajhala, Lauren Wickstrom, Kevin E. Hauser, and Carlos Simmerling. ff14sb: Improving the accuracy of protein side chain and backbone parameters from ff99sb. *Journal of chemical theory and computation*, 11(8):3696–3713, 2015.
- <sup>109</sup> Andrew R. Leach. *Molecular modelling: Principles and applications*. Pearson/Prentice Hall, Harlow, 2. ed., 12. [dr.] edition, 2009.
- <sup>110</sup> William L. Jorgensen. Quantum and statistical mechanical studies of liquids. 10. transferable intermolecular potential functions for water, alcohols, and ethers. application to liquid water. *Journal of the American Chemical Society*, 103(2):335–340, 1981.
- <sup>111</sup> William L. Jorgensen, Jayaraman Chandrasekhar, Jeffry D. Madura, Roger W. Impey, and Michael L. Klein. Comparison of simple potential functions for simulating liquid water. *The Journal of chemical physics*, 79(2):926–935, 1983.
- <sup>112</sup> Alexey V. Onufriev and Saeed Izadi. Water models for biomolecular simulations. *WIREs Computational Molecular Science*, 8(2), 2018.
- <sup>113</sup> Saeed Izadi, Ramu Anandakrishnan, and Alexey V. Onufriev. Building water models: A different approach. *The journal of physical chemistry letters*, 5(21):3863–3871, 2014.
- <sup>114</sup> Vincent Krutler, Wilfred F. van Gunsteren, and Philippe H. Hünenberger. A fast shake algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *Journal of Computational Chemistry*, 22(5):501–508, 2001.
- <sup>115</sup> Hans C. Andersen. Rattle: A “velocity” version of the shake algorithm for molecular dynamics calculations. *Journal of Computational Physics*, 52(1):24–34, 1983.
- <sup>116</sup> Christopher I. Bayly, Piotr Cieplak, Wendy Cornell, and Peter A. Kollman. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the resp model. *The Journal of Physical Chemistry*, 97(40):10269–10280, 1993.
- <sup>117</sup> Junmei Wang, Romain M. Wolf, James W. Caldwell, Peter A. Kollman, and David A. Case. Development and testing of a general amber force field. *Journal of Computational Chemistry*, 25(9):1157–1174, 2004.
- <sup>118</sup> Tom Darden, Darrin York, and Lee Pedersen. Particle mesh ewald: An  $n \cdot \log(n)$  method for ewald sums in large systems. *The Journal of chemical physics*, 98(12):10089–10092, 1993.

- <sup>119</sup> H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. Molecular dynamics with coupling to an external bath. *The Journal of chemical physics*, 81(8):3684–3690, 1984.
- <sup>120</sup> D. J. Evans and B. L. Holian. The nose–hoover thermostat. *The Journal of chemical physics*, 83(8):4069–4074, 1985.
- <sup>121</sup> Hans C. Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of chemical physics*, 72(4):2384–2393, 1980.
- <sup>122</sup> M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox. Gaussian09, revision c.01, 2010.
- <sup>123</sup> Attila Szabo and Neil S. Ostlund. *Modern quantum chemistry: Introduction to advanced electronic structure theory*. Dover Publications, Mineola, New York, 1. publ., unabridged, unaltered republ. of the 1. ed., new york 1989 edition, 1996.
- <sup>124</sup> P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Physical Review*, 136(3B):B864–B871, 1964.
- <sup>125</sup> W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Physical Review*, 140(4A):A1133–A1138, 1965.
- <sup>126</sup> C. Lee, W. Yang, and R. G. Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Physical review. B, Condensed matter*, 37(2):785–789, 1988.
- <sup>127</sup> Axel D. Becke. Density-functional thermochemistry. iii. the role of exact exchange. *The Journal of chemical physics*, 98(7):5648–5652, 1993.

- <sup>128</sup> P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *The Journal of Physical Chemistry*, 98(45):11623–11627, 1994.
- <sup>129</sup> Stefan Grimme, Andreas Hansen, Jan Gerit Brandenburg, and Christoph Bannwarth. Dispersion-corrected mean-field electronic structure methods. *Chemical reviews*, 116(9):5105–5154, 2016.
- <sup>130</sup> Takeshi Yanai, David P. Tew, and Nicholas C. Handy. A new hybrid exchange–correlation functional using the coulomb-attenuating method (cam-b3lyp). *Chemical Physics Letters*, 393(1-3):51–57, 2004.
- <sup>131</sup> Jeng-Da Chai and Martin Head-Gordon. Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Physical chemistry chemical physics : PCCP*, 10(44):6615–6620, 2008.
- <sup>132</sup> Kenichi Fukui. The path of chemical reactions - the irc approach. *Accounts of Chemical Research*, 14(12):363–368, 1981.
- <sup>133</sup> Erik Endres. *MD-Profilierung nichtkovalenter Anlagerungskomplexe von kovalenten Inhibitoren des Cathepsin K*. 2019.
- <sup>134</sup> Erik Endres, Natalia Yuan Chen, and Christoph Sotriffer. Md-based assessment of covalent inhibitors in noncovalent association complexes: Learning from cathepsin k as a test case. *Journal of Chemical Information and Modeling*, 63(10):3186–3197, 2023.
- <sup>135</sup> Gerhard Klebe. *Wirkstoffdesign*. Spektrum Akademischer Verlag, Heidelberg, 2009.
- <sup>136</sup> B. Zhao, C. A. Janson, B. Y. Amegadzie, K. D’Alessio, C. Griffin, C. R. Hanning, C. Jones, J. Kurdyla, M. McQueney, X. Qiu, W. W. Smith, and S. S. Abdel-Meguid. Crystal structure of human osteoclast cathepsin k complex with e-64. *Nature structural biology*, 4(2):109–111, 1997.
- <sup>137</sup> Eva Altmann, Sandra W. Cowan-Jacob, and Martin Missbach. Novel purine nitrile derived inhibitors of the cysteine protease cathepsin k. *Journal of medicinal chemistry*, 47(24):5833–5836, 2004.
- <sup>138</sup> Naoki Teno, Takahiro Miyake, Takeru Ehara, Osamu Irie, Junichi Sakaki, Osamu Ohmori, Hiroki Gunji, Naoko Matsuura, Keiichi Masuya, Yuko Hitomi, Kazuhiko Nonomura,

- Miyuki Horiuchi, Keigo Gohda, Atsuko Iwasaki, Ichiro Umemura, Sachiyo Tada, Motohiko Kometani, Genji Iwasaki, Sandra W. Cowan-Jacob, Martin Missbach, René Lattmann, and Claudia Betschart. Novel scaffold for cathepsin k inhibitors. *Bioorganic & medicinal chemistry letters*, 17(22):6096–6100, 2007.
- <sup>139</sup> Zoran Rankovic, Jiaqiang Cai, Jennifer Kerr, Xavier Fradera, John Robinson, Ashvin Mistry, Emma Hamilton, George McGarry, Fiona Andrews, Wilson Caulfield, Iain Cumming, Maureen Dempster, John Waller, Paul Scullion, Iain Martin, Ann Mitchell, Clive Long, Mark Baugh, Paul Westwood, Emma Kinghorn, John Bruin, William Hamilton, Joost Uitdehaag, Mario van Zeeland, Dominique Potin, Laurent Saniere, Andre Fouquet, François Chevallier, Hortense Deronzier, Cecile Dorleans, and Eric Nicolai. Design and optimization of a series of novel 2-cyano-pyrimidines as cathepsin k inhibitors. *Bioorganic & medicinal chemistry letters*, 20(5):1524–1527, 2010.
- <sup>140</sup> Zoran Rankovic, Jiaqiang Cai, Xavier Fradera, Maureen Dempster, Ashvin Mistry, Ann Mitchell, Clive Long, Emma Hamilton, Angela King, Sylviane Boucharens, Craig Jamieson, Jonathan Gillespie, Iain Cumming, Joost Uitdehaag, and Mario van Zeeland. Dioxotriazines as a novel series of cathepsin k inhibitors. *Bioorganic & medicinal chemistry letters*, 20(5):1488–1490, 2010.
- <sup>141</sup> Zoran Rankovic, Jiaqiang Cai, Jennifer Kerr, Xavier Fradera, John Robinson, Ashvin Mistry, William Finlay, George McGarry, Fiona Andrews, Wilson Caulfield, Iain Cumming, Maureen Dempster, John Waller, Wullie Arbuckle, Mark Anderson, Iain Martin, Ann Mitchell, Clive Long, Mark Baugh, Paul Westwood, Emma Kinghorn, Phil Jones, Joost C. M. Uitdehaag, Mario van Zeeland, Dominique Potin, Laurent Saniere, Andre Fouquet, François Chevallier, Hortense Deronzier, Cecile Dorleans, and Eric Nicolai. Optimisation of 2-cyano-pyrimidine inhibitors of cathepsin k: improving selectivity over herg. *Bioorganic & medicinal chemistry letters*, 20(21):6237–6241, 2010.
- <sup>142</sup> Jure Borišek, Matej Vizovišek, Piotr Sosnowski, Boris Turk, Dušan Turk, Barbara Mohar, and Marjana Novič. Development of n-(functionalized benzoyl)-homocycloleucylglycinonitriles as potent cathepsin k inhibitors. *Journal of medicinal chemistry*, 58(17):6928–6937, 2015.
- <sup>143</sup> Simon Law, Pierre-Marie Andrault, Adeleke H. Aguda, Nham T. Nguyen, Natasha Kruglyak, Gary D. Brayer, and Dieter Brömme. Identification of mouse cathepsin k structural elements that regulate the potency of odanacatib. *The Biochemical journal*, 474(5):851–864, 2017.



- <sup>144</sup> Renee L. DesJarlais, Dennis S. Yamashita, Hye-Ja Oh, Irene N. Uzinskas, Karl F. Erhard, Andrew C. Allen, R. Curtis Haltiwanger, Baoguang Zhao, Ward W. Smith, Sherin S. Abdel-Meguid, Karla D'Alessio, Cheryl A. Janson, Michael S. McQueney, Thaddeus A. Tomaszek, Mark A. Levy, and Daniel F. Veber. Use of x-ray co-crystal structures and molecular modeling to design potent and selective non-peptide inhibitors of cathepsin k. *Journal of the American Chemical Society*, 120(35):9114–9115, 1998.
- <sup>145</sup> R. W. Marquis, Y. Ru, S. M. LoCastro, J. Zeng, D. S. Yamashita, H. J. Oh, K. F. Erhard, L. D. Davis, T. A. Tomaszek, D. Tew, K. Salyers, J. Proksch, K. Ward, B. Smith, M. Levy, M. D. Cummings, R. C. Haltiwanger, G. Trescher, B. Wang, M. E. Hemling, C. J. Quinn, H. Y. Cheng, F. Lin, W. W. Smith, C. A. Janson, B. Zhao, M. S. McQueney, K. D'Alessio, C. P. Lee, A. Marzulli, R. A. Dodds, S. Blake, S. M. Hwang, I. E. James, C. J. Gress, B. R. Bradley, M. W. Lark, M. Gowen, and D. F. Veber. Azepanone-based inhibitors of human and rat cathepsin k. *Journal of medicinal chemistry*, 44(9):1380–1395, 2001.
- <sup>146</sup> Eric E. Boros, David N. Deaton, Anne M. Hassell, Robert B. McFadyen, Aaron B. Miller, Larry R. Miller, Margot G. Paulick, Lisa M. Shewchuk, James B. Thompson, Derril H. Willard, and Lois L. Wright. Exploration of the p2-p3 sar of aldehyde cathepsin k inhibitors. *Bioorganic & medicinal chemistry letters*, 14(13):3425–3429, 2004.
- <sup>147</sup> Elma Mons, Ineke D. C. Jansen, Jure Loboda, Bjorn R. van Doodewaerd, Jill Hermans, Martijn Verdoes, Constant A. A. van Boeckel, Peter A. van Veelen, Boris Turk, Dusan Turk, and Huib Ovaa. The alkyne moiety as a latent electrophile in irreversible covalent small molecule inhibitors of cathepsin k. *Journal of the American Chemical Society*, 141(8):3507–3514, 2019.
- <sup>148</sup> Chemical Computing Group ULC. Molecular operating environment (moe) 2019.01, 2021.
- <sup>149</sup> Paul Labute. Protonate3d: assignment of ionization states and hydrogen coordinates to macromolecular structures. *Proteins*, 75(1):187–205, 2009.
- <sup>150</sup> Thomas A. Halgren. Merck molecular force field. i. basis, form, scope, parameterization, and performance of mmff94. *Journal of Computational Chemistry*, 17(5-6):490–519, 1996.
- <sup>151</sup> David A. Case, Thomas E. Cheatham, Tom Darden, Holger Gohlke, Ray Luo, Kenneth M. Merz, Alexey Onufriev, Carlos Simmerling, Bing Wang, and Robert J. Woods. The amber biomolecular simulation programs. *Journal of Computational Chemistry*, 26(16):1668–1688, 2005.

- <sup>152</sup> Alexey Onufriev, Donald Bashford, and David A. Case. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins*, 55(2):383–394, 2004.
- <sup>153</sup> Romelia Salomon-Ferrer, Andreas W. Götz, Duncan Poole, Scott Le Grand, and Ross C. Walker. Routine microsecond molecular dynamics simulations with amber on gpus. 2. explicit solvent particle mesh ewald. *Journal of chemical theory and computation*, 9(9):3878–3888, 2013.
- <sup>154</sup> Daniel R. Roe and Thomas E. Cheatham. Ptraj and cpptraj: Software for processing and analysis of molecular dynamics trajectory data. *Journal of chemical theory and computation*, 9(7):3084–3095, 2013.
- <sup>155</sup> Gerd Neudert and Gerhard Klebe. fconv: Format conversion, manipulation and feature computation of molecular data. *Bioinformatics (Oxford, England)*, 27(7):1021–1022, 2011.
- <sup>156</sup> Alan J. Barrett and Neil D. Rawlings. Evolutionary lines of cysteine peptidases. *bchm*, 382(5):727–734, 2001.
- <sup>157</sup> Thomas A. Pauly, Traian Sulea, Mark Ammirati, J. Sivaraman, Dennis E. Danley, Matthew C. Griffor, Ajith V. Kamath, I-K Wang, Ellen R. Laird, Andrew P. Seddon, Robert Ménard, Miroslaw Cygler, and Virginia L. Rath. Specificity determinants of human cathepsin s revealed by crystal structures of complexes. *Biochemistry*, 42(11):3203–3213, 2003.
- <sup>158</sup> Jiaqiang Cai, John Robinson, Simone Belshaw, Kathryn Everett, Xavier Fradera, Mario van Zeeland, Leon van Berkomp, Peter van Rijnsbergen, Lucy Popplestone, Mark Baugh, Maureen Dempster, John Bruin, William Hamilton, Emma Kinghorn, Paul Westwood, Jennifer Kerr, Zoran Rankovic, Wullie Arbuckle, D. Jonathan Bennett, Philip S. Jones, Clive Long, Iain Martin, Joost C. M. Uitdehaag, and Tommi Meulemans. Trifluoromethylphenyl as p2 for ketoamide-based cathepsin s inhibitors. *Bioorganic & medicinal chemistry letters*, 20(23):6890–6894, 2010.
- <sup>159</sup> Yancey D. Ward, David S. Thomson, Leah L. Frye, Charles L. Cywin, Tina Morwick, Michel J. Emmanuel, Renée Zindell, Daniel McNeil, Younes Bekkali, Marc Girardot, Matt Hrapchak, Molly DeTuri, Kathy Crane, Della White, Susan Pav, Yong Wang, Ming-Hong Hao, Christine A. Grygon, Mark E. Labadia, Dorothy M. Freeman, Walter Davidson, Jerry L. Hopkins, Maryanne L. Brown, and Denice M. Spero. Design and synthesis of

- dipeptide nitriles as reversible and potent cathepsin s inhibitors. *Journal of medicinal chemistry*, 45(25):5471–5482, 2002.
- <sup>160</sup> J. R. Somoza. *Cathepsin S with nitrile inhibitor*. 2006.
- <sup>161</sup> J. R. Somoza. *Human Cathepsin S with CRA-27934, a Nitrile Inhibitor*. 2006.
- <sup>162</sup> J. R. Somoza. *Human Cathepsin S with Inhibitor CRA-26871*. 2006.
- <sup>163</sup> J. R. Somoza. *Human Cathepsin S with Inhibitor CRA-29728*. 2006.
- <sup>164</sup> Y. D. Ward, M. J. Emmanuel, D. S. Thomson, W. Liu, Y. Bekkali, L. L. Frye, M. Girardot, T. Morwick, E.R.R. Young, R. Zindell, M. Hrapchak, M. DeTuri, A. White, K. M. Crane, D. M. White, Y. Wang, M.-H. Hao, C. A. Grygon, M. E. Labadia, J. Wildeson, D. Freeman, R. Nelson, A. Capolino, J. D. Peterson, E. L. Raymond, M. L. Brown, and D. M. Spero. *Cathepsin S complexed with Compound 15*. 2007.
- <sup>165</sup> Y. D. Ward, M. J. Emmanuel, D. S. Thomson, W. Liu, Y. Bekkali, L. L. Frye, M. Girardot, T. Morwick, E.R.R. Young, R. Zindell, M. Hrapchak, M. DeTuri, A. White, K. M. Crane, D. M. White, Y. Wang, M.-H. Hao, C. A. Grygon, M. E. Labadia, J. Wildeson, D. Freeman, R. Nelson, A. Capolino, J. D. Peterson, E. L. Raymond, M. L. Brown, and D. M. Spero. *Cathepsin S complexed with Compound 26*. 2007.
- <sup>166</sup> Y. D. Ward, M. J. Emmanuel, D. S. Thomson, W. Liu, Y. Bekkali, L. L. Frye, M. Girardot, T. Morwick, E.R.R. Young, R. Zindell, M. Hrapchak, M. DeTuri, A. White, K. M. Crane, D. M. White, Y. Wang, M.-H. Hao, C. A. Grygon, M. E. Labadia, J. Wildeson, D. Freeman, R. Nelson, A. Capolino, J. D. Peterson, E. L. Raymond, M. L. Brown, and D. M. Spero. *Cathepsin S complexed with Compound 8*. 2007.
- <sup>167</sup> Jiaqiang Cai, Mark Baugh, Darcey Black, Clive Long, D. Jonathan Bennett, Maureen Dempster, Xavier Fradera, Jonathan Gillespie, Fiona Andrews, Sylviane Boucharens, John Bruin, Kenneth S. Cameron, Iain Cumming, William Hamilton, Philip S. Jones, Allard Kaptein, Emma Kinghorn, Maurice Maidment, Iain Martin, Ann Mitchell, Zoran Rankovic, John Robinson, Paul Scullion, Joost C. M. Uitdehaag, Paul Vink, Paul Westwood, Mario van Zeeland, Leon van Berkom, Martijn Bastiani, and Tommi Meulemans. 6-phenyl-1h-imidazo4,5-cpyridine-4-carbonitrile as cathepsin s inhibitors. *Bioorganic & medicinal chemistry letters*, 20(15):4350–4354, 2010.
- <sup>168</sup> J. R. Somoza. *Human Cathepsin S with inhibitor CRA-16981*. 2006.

- <sup>169</sup> M. D. Eldridge, C. W. Murray, T. R. Auton, G. V. Paolini, and R. P. Mee. Empirical scoring functions: I. the development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *Journal of computer-aided molecular design*, 11(5):425–445, 1997.
- <sup>170</sup> C. W. Murray, T. R. Auton, and M. D. Eldridge. Empirical scoring functions. ii. the testing of an empirical scoring function for the prediction of ligand-receptor binding affinities and the use of bayesian regression to improve the quality of the model. *Journal of computer-aided molecular design*, 12(5):503–519, 1998.
- <sup>171</sup> Oliver Korb, Thomas Stützle, and Thomas E. Exner. Empirical scoring functions for advanced protein-ligand docking with plants. *Journal of Chemical Information and Modeling*, 49(1):84–96, 2009.
- <sup>172</sup> Jörg Schröder, Anette Klinger, Frank Oellien, Richard J. Marhöfer, Michael Duszenko, and Paul M. Selzer. Docking-based virtual screening of covalently binding ligands: an orthogonal lead discovery approach. *Journal of medicinal chemistry*, 56(4):1478–1490, 2013.
- <sup>173</sup> Jason L. Sonnenberg, Kim F. Wong, Gregory A. Voth, and H. Bernhard Schlegel. Distributed gaussian valence bond surface derived from ab initio calculations. *Journal of chemical theory and computation*, 5(4):949–961, 2009.
- <sup>174</sup> Shaobo Dai, Lisa-Marie Funk, Fabian Rabe von Pappenheim, Viktor Sautner, Mirko Paulikat, Benjamin Schröder, Jon Uranga, Ricardo A. Mata, and Kai Tittmann. Low-barrier hydrogen bonds in enzyme cooperativity. *Nature*, 573(7775):609–613, 2019.
- <sup>175</sup> Alexander Paasche, Tanja Schirmeister, and Bernd Engels. Benchmark study for the cysteine-histidine proton transfer reaction in a protein environment: Gas phase, cosmo, qm/mm approaches. *Journal of chemical theory and computation*, 9(3):1765–1777, 2013.
- <sup>176</sup> Jeremie Mortier, Christin Rakers, Raphael Frederick, and Gerhard Wolber. Computational tools for in silico fragment-based drug design. *Current topics in medicinal chemistry*, 12(17):1935–1943, 2012.
- <sup>177</sup> Philine Kirsch, Alwin M. Hartman, Anna K. H. Hirsch, and Martin Empting. Concepts and core principles of fragment-based drug design. *Molecules (Basel, Switzerland)*, 24(23), 2019.

- <sup>178</sup> Peter G. M. Wuts and Theodora W. Greene. *Greene's Protective Groups in Organic Synthesis*. Wiley, 2006.
- <sup>179</sup> Scott J. Weiner, Peter A. Kollman, David A. Case, U. Chandra Singh, Caterina Ghio, Guliano Alagona, Salvatore Profeta, and Paul Weiner. A new force field for molecular mechanical simulation of nucleic acids and proteins. *Journal of the American Chemical Society*, 106(3):765–784, 1984.
- <sup>180</sup> Pieter F. W. Stouten, Cornelius Frömmel, Haruki Nakamura, and Chris Sander. An effective solvation term based on atomic occupancies for use in protein simulations. *Molecular Simulation*, 10(2-6):97–120, 1993.
- <sup>181</sup> Johann Gasteiger and Mario Marsili. Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges. *Tetrahedron*, 36(22):3219–3228, 1980.
- <sup>182</sup> Francisco J. Solis and Roger J.-B. Wets. Minimization by random search techniques. *Mathematics of Operations Research*, 6(1):19–30, 1981.
- <sup>183</sup> Clayden, Greeves, Warren, Wothers. *Organic chemistry*. Oxford University Press, Oxford, first published. edition, 2001.
- <sup>184</sup> Rory Korzan, Brian Upton, Kenneth Turnbull, and Paul G. Seybold. Quantum chemical study of the energetics and directionality of acid-catalyzed aromatic epoxide ring openings. *International Journal of Quantum Chemistry*, 110(15):2931–2937, 2010.
- <sup>185</sup> A. Keeley, P. Ábrányi-Balogh, and G. M. Keserű. Design and characterization of a heterocyclic electrophilic fragment library for the discovery of cysteine-targeted covalent inhibitors. *MedChemComm*, 10(2):263–267, 2019.
- <sup>186</sup> Anna Berteotti, Federica Vacondio, Alessio Lodola, Michele Bassi, Claudia Silva, Marco Mor, and Andrea Cavalli. Predicting the reactivity of nitrile-carrying compounds with cysteine: a combined computational and experimental study. *ACS medicinal chemistry letters*, 5(5):501–505, 2014.
- <sup>187</sup> Vincenzo Barone and Maurizio Cossi. Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. *The Journal of Physical Chemistry A*, 102(11):1995–2001, 1998.

- <sup>188</sup> Maurizio Cossi, Nadia Rega, Giovanni Scalmani, and Vincenzo Barone. Energies, structures, and electronic properties of molecules in solution with the c-pcm solvation model. *Journal of Computational Chemistry*, 24(6):669–681, 2003.
- <sup>189</sup> Elma Mons, Robbert Q. Kim, Bjorn R. van Doodewaerd, Peter A. van Veelen, Monique P. C. Mulder, and Huib Ovaa. Exploring the versatility of the covalent thiol-alkyne reaction with substituted propargyl warheads: A deciding role for the cysteine protease. *Journal of the American Chemical Society*, 143(17):6423–6433, 2021.
- <sup>190</sup> Paul D. Leeson and Brian Springthorpe. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nature reviews. Drug discovery*, 6(11):881–890, 2007.
- <sup>191</sup> Andrew R. Leach, Michael M. Hann, Jeremy N. Burrows, and Ed J. Griffen. Fragment screening: an introduction. *Molecular bioSystems*, 2(9):430–446, 2006.
- <sup>192</sup> Matías R. Machado and Sergio Pantano. Split the charge difference in two! a rule of thumb for adding proper amounts of ions in md simulations. *Journal of chemical theory and computation*, 16(3):1367–1372, 2020.
- <sup>193</sup> Barry J. Grant, Ana P. C. Rodrigues, Karim M. ElSawy, J. Andrew McCammon, and Leo S. D. Caves. Bio3d: an r package for the comparative analysis of protein structures. *Bioinformatics (Oxford, England)*, 22(21):2695–2696, 2006.
- <sup>194</sup> Renata M. Oballa, Jean-François Truchon, Christopher I. Bayly, Nathalie Chauret, Stephen Day, Sheldon Crane, and Carl Berthelette. A generally applicable method for assessing the electrophilicity and reactivity of diverse nitrile-containing compounds. *Bioorganic & medicinal chemistry letters*, 17(4):998–1002, 2007.
- <sup>195</sup> Michael Decker, Fabian Krauth, and Jochen Lehmann. Novel tricyclic quinazolinimines and related tetracyclic nitrogen bridgehead compounds as cholinesterase inhibitors with selectivity towards butyrylcholinesterase. *Bioorganic & medicinal chemistry*, 14(6):1966–1977, 2006.
- <sup>196</sup> Fouad H. Darras, Beata Kling, Jörg Heilmann, and Michael Decker. Neuroprotective tri- and tetracyclic bche inhibitors releasing reversible inhibitors upon carbamate transfer. *ACS medicinal chemistry letters*, 3(11):914–919, 2012.
- <sup>197</sup> Edgar Sawatzky, Sarah Wehle, Beata Kling, Jan Wendrich, Gerhard Bringmann, Christoph A. Sotriffer, Jörg Heilmann, and Michael Decker. Discovery of highly selective and

- nanomolar carbamate-based butyrylcholinesterase inhibitors by rational investigation into their inhibition mode. *Journal of medicinal chemistry*, 59(5):2067–2082, 2016.
- <sup>198</sup> Irene Núñez, Estíbaliz Merino, Mercedes Lecea, Silvia Pieraccini, Gian Piero Spada, Carlo Rosini, Giuseppe Mazzeo, María Ribagorda, and M. Carmen Carreño. Control of the helical chirality of enantiopure sulfinyl (z)-azobenzene-based photoswitches. *Chemistry (Weinheim an der Bergstrasse, Germany)*, 19(10):3397–3406, 2013.
- <sup>199</sup> Gebhard Haberhauer and Christine Kallweit. A bridged azobenzene derivative as a reversible, light-induced chirality switch. *Angewandte Chemie (International ed. in English)*, 49(13):2418–2421, 2010.
- <sup>200</sup> Kazuto Takaishi, Masuki Kawamoto, Kazunori Tsubaki, Taniyuki Furuyama, Atsuya Muranaka, and Masanobu Uchiyama. Helical chirality of azobenzenes induced by an intramolecular chiral axis and potential as chiroptical switches. *Chemistry (Weinheim an der Bergstrasse, Germany)*, 17(6):1778–1782, 2011.
- <sup>201</sup> Kazuto Takaishi, Atsuya Muranaka, Masuki Kawamoto, and Masanobu Uchiyama. Planar chirality of twisted trans-azobenzene structure induced by chiral transfer from binaphthyls. *The Journal of organic chemistry*, 76(18):7623–7628, 2011.
- <sup>202</sup> Aleksandr V. Marenich, Christopher J. Cramer, and Donald G. Truhlar. Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *The journal of physical chemistry. B*, 113(18):6378–6396, 2009.

# Abbildungsverzeichnis

1.1. Gleichgewichte bei kovalenter Inhibition . . . . .	8
1.2. CatK und dessen Bindetasche . . . . .	12
1.3. Acetylcholin und AChE-Inhibitoren . . . . .	15
1.4. AChE und BChE im Vergleich . . . . .	16
1.5. Reaktionsmechanismus von Serinhydrolasen . . . . .	17
1.6. Oligomerisierung von USP25/28 . . . . .	19
1.7. USP28-Inhibitoren . . . . .	20
1.8. Bindetasche von USP25/28 . . . . .	22
3.1. Piktographische Darstellung von GA-Operatoren . . . . .	29
3.2. Vorteil des LGA . . . . .	30
3.3. Paarpotenzial AutoDock vs WIDOCK . . . . .	32
3.4. Kraftfeldterme . . . . .	33
4.1. Liganden der CatK-Studie . . . . .	49
4.2. Ligand von 1ATK . . . . .	53
4.3. Reaktive Distanz in 1ATK-Simulationen . . . . .	53
4.4. Liganden von 1U9X, 1U9V und 1U9W . . . . .	54
4.5. Reaktive Distanz der 1U9X- und 1U9V-Simulationen . . . . .	55
4.6. Liganden von 4X6H, 4X6I und 4X6J . . . . .	55
4.7. 4X6-Simulationen . . . . .	58
4.8. Vergleich 4X6J- und 4X6I-Kristallstruktur . . . . .	59
4.9. Liganden von 6QBS und 5TDI . . . . .	60
4.10. Reaktiver Abstand der 6QBS- und 5TDI-Simulationen . . . . .	61
4.11. Histogramme zum reaktiven Abstand der 6QBS-Simulation . . . . .	62
4.12. Vergleich des Torsionspotenzials verschiedener Warheads . . . . .	62
4.13. Boxplots aller CatK-Simulationen . . . . .	64
4.14. Erklärung der Gitter-basierten Analyse . . . . .	66
4.15. Ligand von 4DMX . . . . .	66
4.16. Gitteranalyse der 4DMX-Simulation . . . . .	67
4.17. Ligand von 3KWB . . . . .	68



---

4.18. Gitteranalyse 1SNK . . . . .	70
4.19. Reaktive Distanzen der 1SNK-Simulationen . . . . .	71
4.20. Gitteranalyse der 1MEM-Simulationen . . . . .	71
4.21. Gitteranalyse der 1BGO-Simulationen . . . . .	72
4.22. Gitteranalyse der 1NLJ-Simulationen . . . . .	72
4.23. Zeitaufgelöste Analyse des Wasser-Clusters, CYM/HIP . . . . .	75
4.24. Zeitaufgelöste Distanzen des Wasser-Clusters, CYS/HIS . . . . .	76
4.25. Verlängerte 5TDI-Simulationen . . . . .	77
4.26. Verlängerte 1ATK-Simulationen . . . . .	78
4.27. Verlängerte 1MEM-Simulationen . . . . .	78
4.28. Verlängerte 4X6J-Simulationen . . . . .	79
4.29. Verlängerte 6QBS-Simulationen . . . . .	80
4.30. Überlagerung papainähnlicher Cysteinproteasen . . . . .	81
4.31. Alle Liganden der CatS-Studie . . . . .	82
4.32. Boxplots der CatS-Simulationen . . . . .	83
4.33. Gitteranalyse der 1MS6-Simulationen . . . . .	84
4.34. Reaktiver Abstand der 2G7Y-Simulationen. . . . .	85
4.35. Nirmatrelvir . . . . .	85
4.36. Reaktive Distanz in den 7VH8-Simulationen . . . . .	86
4.37. Die beiden dominanten Zustände der 7VH8-Simulationen . . . . .	87
4.38. Reaktiver Abstand der Simulationen basierend auf Docking-Lösungen . . . . .	91
4.39. Boxplots des Schröder-Nitril-Subsets . . . . .	94
4.40. Gitterbasierte Dichteanalyse des Subsets . . . . .	95
4.41. Überlagerung der Gitter-Analyse, CYS/HIS . . . . .	98
4.42. Mögliche Setups für eine QM/MM-Simulation . . . . .	100
4.43. AZ1 und Vismodegib . . . . .	101
4.44. Unveröffentlichte Kristallstrukturen des AK Kisker . . . . .	103
4.45. Vismodegib-Strukturen im Vergleich . . . . .	104
4.46. Überlapp der AZ1- und Vismodegib-Bindestellen . . . . .	105
4.47. AZ1-Variationen . . . . .	106
4.48. Postulierter Mechanismus eines Thiolats mit einem Alkin . . . . .	110
4.49. Mechanismen einer Epoxid-Öffnung . . . . .	110
4.50. Top-Pose eines Nitrils mit AD4 . . . . .	113
4.51. WIDOCK-Ergebnis mit linearem Nitril . . . . .	114
4.52. Aufbau des WIDOCK-Testsystems . . . . .	115

4.53. 2D-Lennard-Jones-Potenziale eines Nitrils . . . . .	117
4.54. Top-Pose eines gewinkelten Nitrils . . . . .	119
4.55. WIDOCK-Posen für Nitril mit CF <sub>3</sub> -Gruppe . . . . .	120
4.56. WIDOCK-Posen eines Alkins . . . . .	123
4.57. WIDOCK-Posen der Aziridine . . . . .	126
4.58. WIDOCK-Posen des Epoxids . . . . .	127
4.59. WIDOCK-Pose des Vinylsulfons . . . . .	129
4.60. Nitril-Warheads für Berechnungen . . . . .	130
4.61. Modell-Mechanismus zwischen Nitril und Cysteamin . . . . .	131
4.62. Ergebnisse der DFT-Rechnungen für die USP28-Warheads . . . . .	133
4.63. Nichtkovalente Anlagerungskomplexe . . . . .	134
4.64. Alkin-Warheads für DFT-Rechnungen im USP28-Projekt . . . . .	135
4.65. Modell-Mechanismus Alkin und Cysteamin . . . . .	136
4.66. Ergebnisse der DFT-Rechnungen mit den Alkin-Warheads . . . . .	137
4.67. Experimentelle Ergebnisse von Jonathan Patzke . . . . .	138
4.68. AZ3 im Vergleich zum experimentell getesteten Nitril . . . . .	139
4.69. Hypothetisches alkin mit intrinsischer Base . . . . .	140
4.70. Strukturvorschläge für den Merging Approach . . . . .	141
4.71. Docking-Posen für den Merging-Approach . . . . .	143
4.72. Distanz zwischen $\alpha$ -Helix und Glu373/366 . . . . .	147
4.73. Stacking-Distanz in USP25-Simulation . . . . .	148
4.74. Zentroid-Frames des Clusterings der USP25/USP28-Simulationen . . . . .	149
4.75. Distanz zwischen $\alpha$ -Helix des mutierten Glu373/366Ala . . . . .	150
4.76. Histogramm für die Helix-Öffnung . . . . .	151
4.77. PCA-Analyse der USP-Simulationen . . . . .	152
4.78. ChlaDUB1-Warheads . . . . .	153
4.79. Energieprofile der ChlaDUB1-Warheads . . . . .	154
4.80. Konformere für das Pyrazin . . . . .	156
4.81. Aktivierungsenergien für Pyrazin-Konformere . . . . .	157
4.82. Energieprofile für ChlaDUB1-Warheads mit höherer Aktivierungsenergie . . . . .	158
4.83. Experimentell getestete ChlaDUB1-Inhibitoren . . . . .	158
4.84. Der Photoswitch . . . . .	163
4.85. Konfigurationen des Ringsystems . . . . .	164
4.86. Räumliche Darstellung der Ringsystem-Stereoisomere . . . . .	165
4.87. Helikale Chiralität eines <i>cis</i> -Photoswitches . . . . .	166

---

4.88. Alle untersuchten Konfigurationen des <i>cis</i> -Photoswitches . . . . .	167
4.89. Top-Posen des Photoswitches . . . . .	170
A.1. Grids der Nitrile, aus den CYS/HIS-Simulationen. . . . .	208
A.2. Grids der Nitrile, aus den CYM/HIP-Simulationen. . . . .	209
A.3. Grids der Nitrile, aus den CYS/HIS-Simulationen von CatS. . . . .	210
A.4. Grids der Nitrile, aus den CYM/HIP-Simulationen von CatS. . . . .	211
A.5. Überlagerung der Grid-Analyse jedes CatS-Nitrils, basierend auf den CYS/HIS-Simulationen. Wasser-Cluster sind als rote, Nitril-Kohlenstoff-Cluster als grüne und Nitril-Stickstoff-Cluster als blaue Kugeln gezeigt. Die Cluster für die Kohlenstoffatome und die Stickstoffatome der Nitrile nehmen räumlich eine sehr ähnliche Position ein. . . . .	212

# Tabellenverzeichnis

4.1. PDBs für CatK-Projekt . . . . .	50
4.2. Fakten zu den 4X6-Komplexen . . . . .	56
4.3. Partialladungen der 4X6-Amidbindungen . . . . .	57
4.4. Distanz zwischen Cys25 und Nitril-Kohlenstoffatom-Cluster . . . . .	68
4.5. Distanzen des Wasser-Clusters zu umliegenden Aminosäuren . . . . .	69
4.6. Dauer der Wasserstoffbrückenbindungen . . . . .	73
4.7. Übersicht der simulierten CatS-Komplexe . . . . .	81
4.8. Auswertung des DSX-Rescorings. . . . .	89
4.9. Ausgewählte Posen für die MD-Simulationen. . . . .	90
4.10. Simulationsergebnisse für die Schröder-Verbindungen . . . . .	94
4.11. Docking der Nitrile mit AD4 . . . . .	112
4.12. Docking-Ergebnisse des WIDOCK-Tests mit Nitrilen . . . . .	114
4.13. Parameter für Modellsystem . . . . .	115
4.14. WIDOCK-Ergebnisse für gewinkelte Nitrile . . . . .	118
4.15. WIDOCK-Ergebnisse für Nitril mit CF <sub>3</sub> -Gruppe . . . . .	119
4.16. Docking-Ergebnisse für AD4 . . . . .	121
4.17. WIDOCK-Ergebnisse der Alkine . . . . .	122
4.18. Docking-Ergebnisse für Dreiring-Warheads mit AD4 . . . . .	124
4.19. WIDOCK-Ergebnisse der Dreiring-Warheads . . . . .	125
4.20. Docking-Ergebnisse für sterisch anspruchsvolle Warheads mit AD4 . . . . .	128
4.21. Docking-Ergebnisse des WIDOCK-Protokolls für die sterisch anspruchsvollen Warheads. Neben der <i>Estimated Free Enthalpy</i> wird der Substruktur-RMSD des AZ1-Gerüsts im Vergleich zur Kristallstruktur angegeben. Die angegebenen Distanzen wurden zwischen dem reaktiven Kohlenstoffatom und dem Schwefelatom von Cys644 gemessen. . . . .	128
4.22. Docking-Ergebnisse für den Merging Approach . . . . .	142
4.23. Berechnete Aktivierungsenergien und experimentell bestimmte Geschwindigkeitskonstanten . . . . .	159
4.24. Vergleich der berechneten Aktivierungsenergien zwischen Berteotti <i>et al.</i> <sup>186</sup> und dieser Arbeit. . . . .	162

4.25. Docking-Ergebnisse des Photoswitches . . . . .	169
A.1. Aktivierungsenergien und freie Aktivierungsenthalpien für die USP28-Warheads.	213
A.2. Aktivierungsenergien und freie Aktivierungsenthalpien für die USP28-Alkin-Warheads. . . . .	214
A.3. Aktivierungsenergien und freie Aktivierungsenthalpien für die ChlaDUB1-Nitril-Warheads. . . . .	214

# Abkürzungsverzeichnis

ABPP	<i>Activity-Based Protein Profiling</i>	MOE	<i>Molecular Operating Environment</i>
ACh	Acetylcholin	QM	Quantenmechanik
AChE	Acetylcholin-Esterase	MP2	Møller-Plesset-Störungstheorie zweiter Ordnung
AD	Alzheimer-Demenz	PAS	<i>per-atom-score</i>
AIM	autoinhibitorisches Motiv	PBC	<i>periodic boundary conditions</i>
AMBER	<i>Assisted Model Building with Energy Refinement</i>	PC	<i>Principal Component</i>
ATP	Adenosintriphosphat	PCA	<i>Principal Component Analysis</i>
BChE	Butyrylcholin-Esterase	PDB	<i>Protein Data Bank</i>
BTK	<i>Bruton tyrosine kinase</i>	RESP	<i>Restrained Electrostatic Potential</i>
Cat	Cathepsin	RMS(D)	<i>Root-Mean-Square(-Deviation)</i>
Chla-	<i>Chlamydia-Deubiquitinase</i>	SAR	<i>Structure-Activity-Relationship</i>
DUB1		SARS-	<i>severe acute respiratory</i>
DFT	Dichtefunktionaltheorie	CoV-2	<i>syndrome corona virus 2</i>
DSX	DrugScoreX	SCF	<i>Self-Consistent-Field</i>
DUB	Deubiquitinase	TCI	<i>Targeted Covalent Inhibitor</i>
EE	<i>Electronic Energy</i>	UB	Ubiquitin
EGFR	<i>Epidermal growth factor receptor</i>	UCID	USP25/28
GA	Genetischer Algorithmus		catalytic domain inserted domain
GOLD	<i>Genetic Optimisation for Ligand Docking</i>	USP	Ubiquitin-spezifische Protease
GSH	Glutathion		
HDX-MS	Wasserstoff-Deuterium- Austausch-Massenspektrometrie		
HF	Hartree-Fock		
HOMO	<i>Highest Occupied Molecular Orbital</i>		
IC <sub>50</sub>	mittlere inhibitorische Konzentration		
IRC	<i>intrinsic reaction coordinate</i>		
LGA	Lamarcksche Genetische Algorithmus		
LLE	<i>lipophilic ligand efficiency</i>		
LUMO	<i>Lowest unoccupied Molecular Orbital</i>		
MD	Molekulardynamik		
MM	Molekularmechanik		
MO	Molekülorbital		

---

## **A. Anhang**

## A.1. CatK

### A.1.1. Gitterbasierte Analysen zu CatK

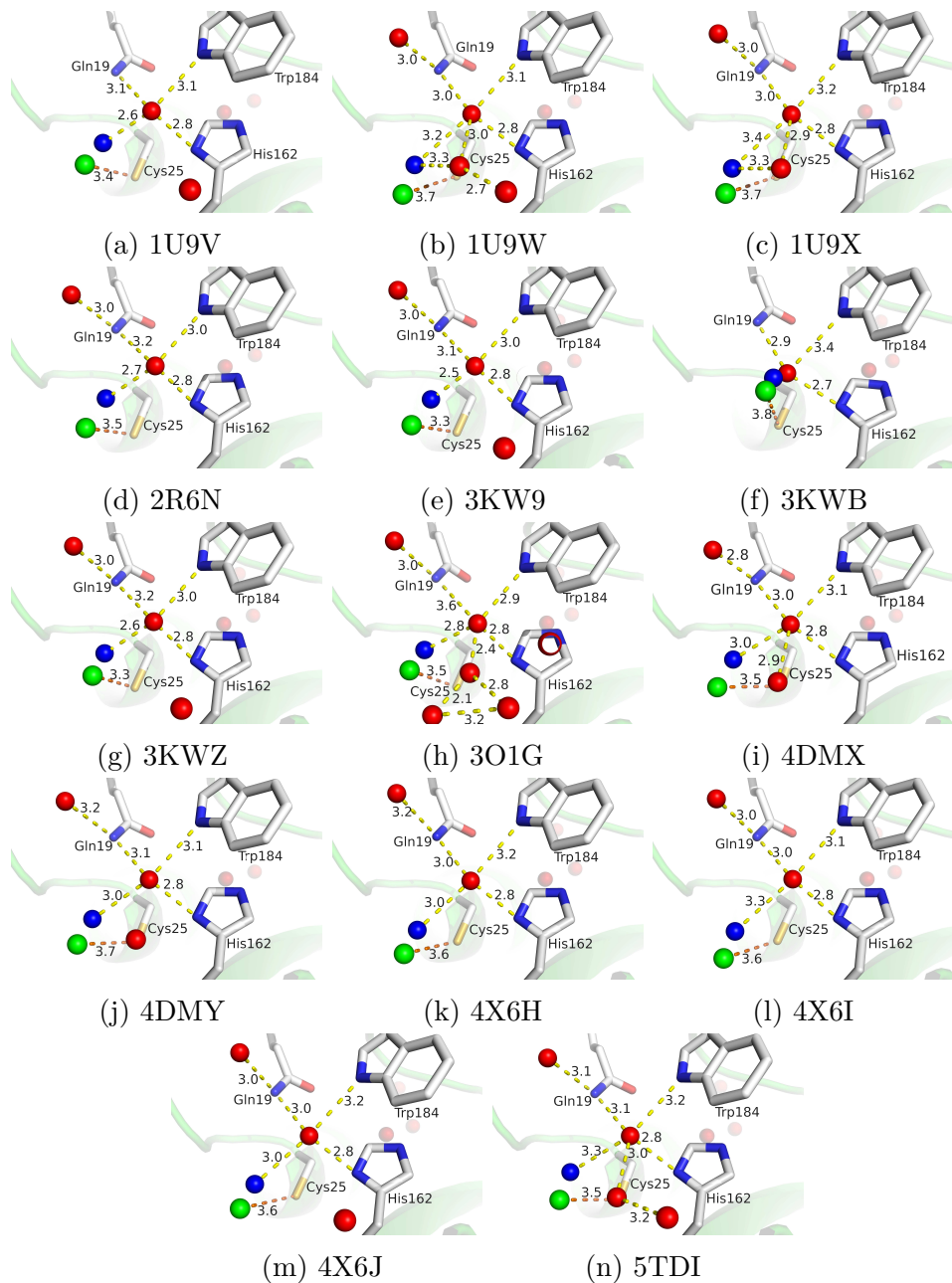


Abbildung A.1.: Grids der Nitrile, aus den CYS/HIS-Simulationen.



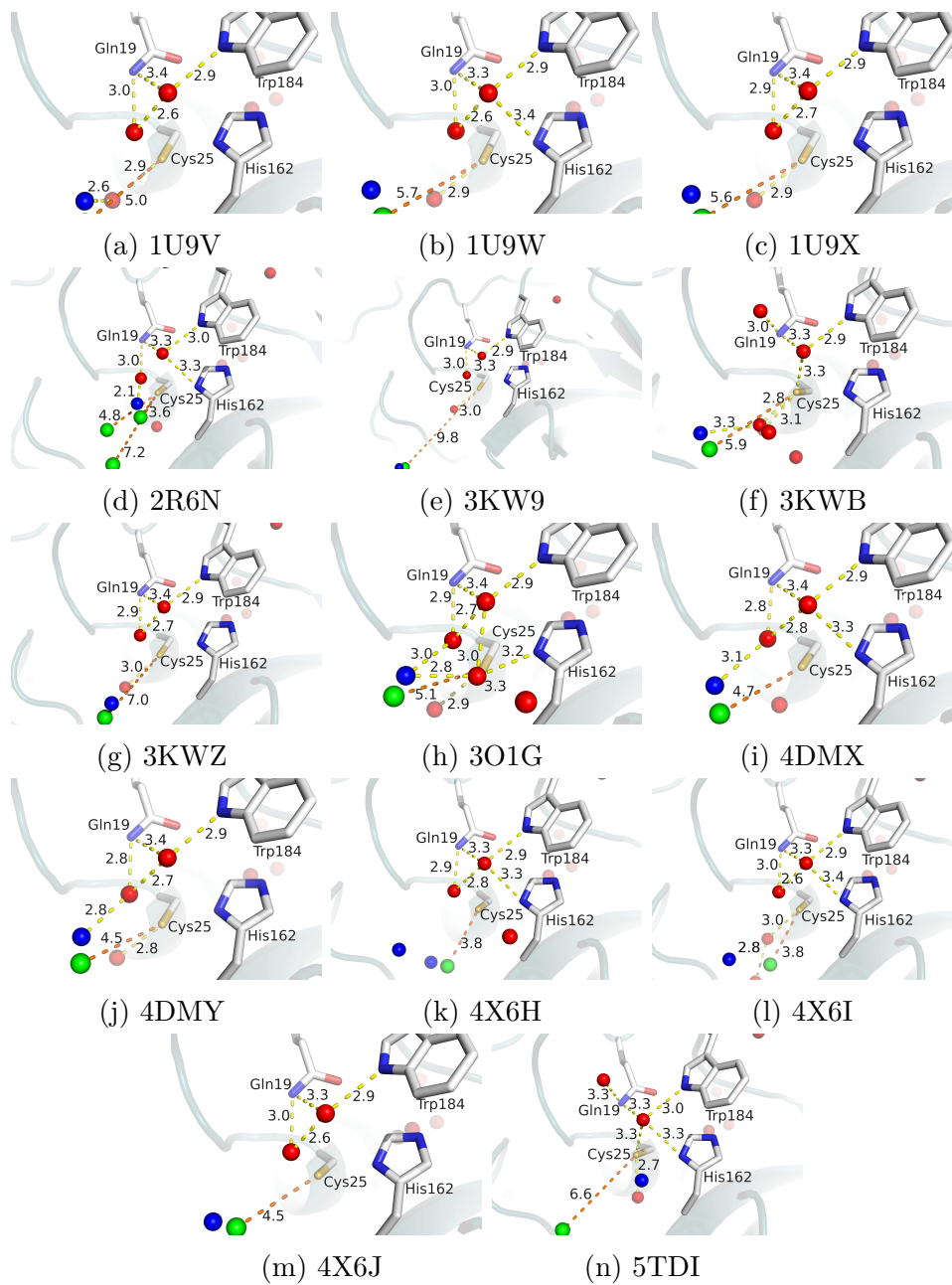


Abbildung A.2.: Grids der Nitrile, aus den CYM/HIP-Simulationen.

### A.1.2. Gitterbasierte Analysen zu *CatS*

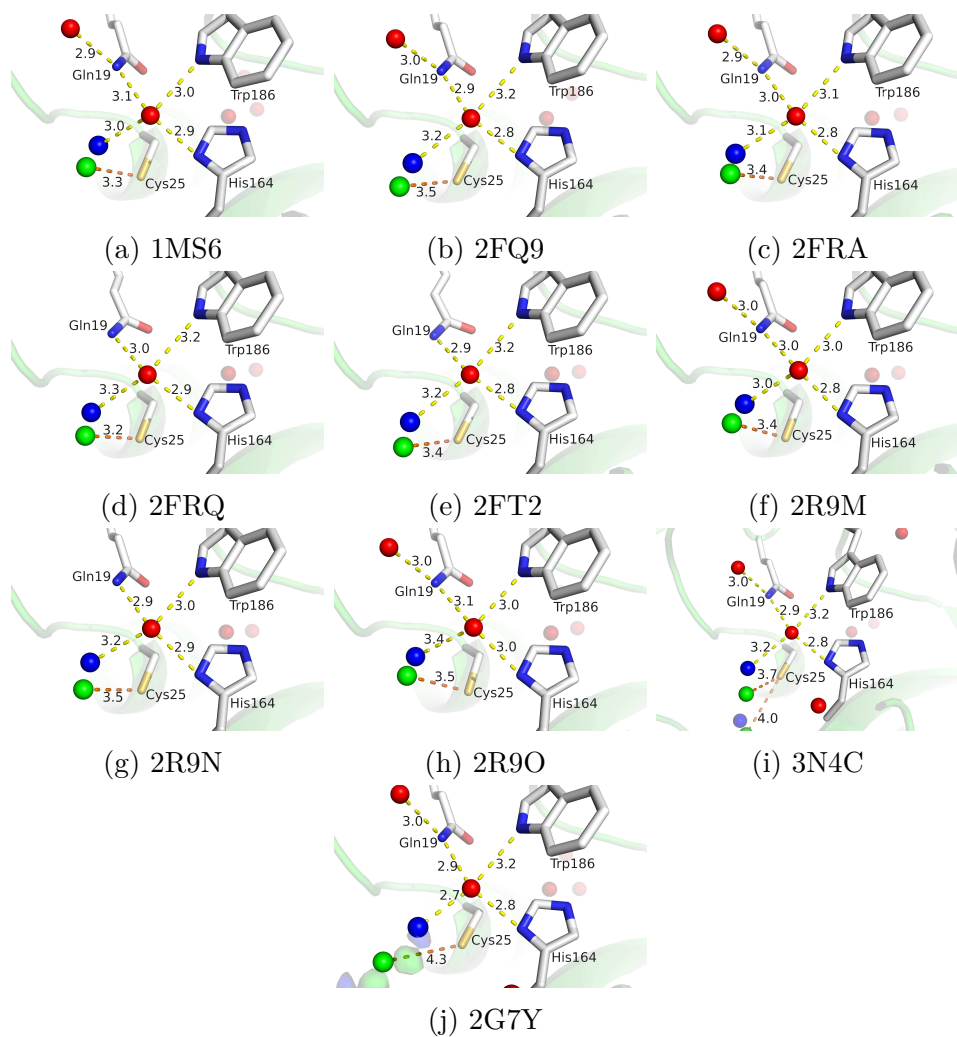


Abbildung A.3.: Grids der Nitrile, aus den CYS/HIS-Simulationen von CatS.

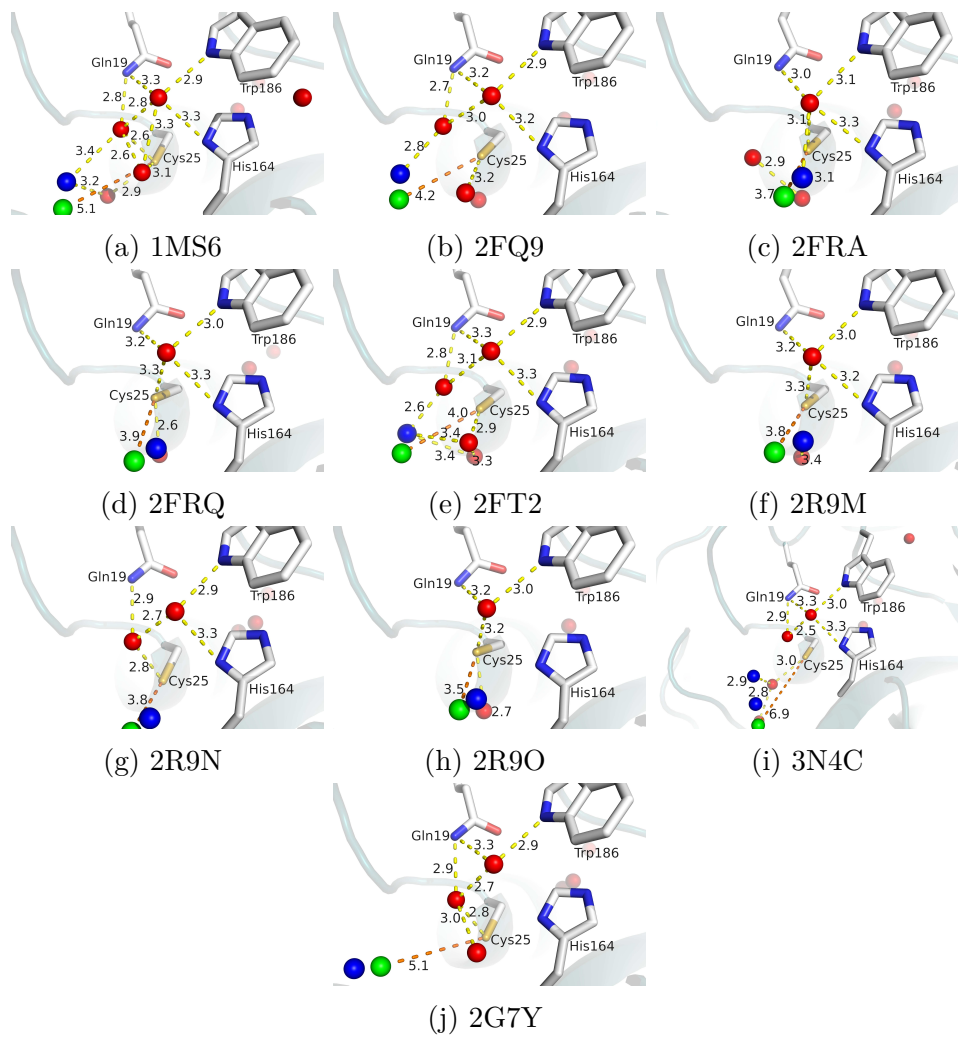


Abbildung A.4.: Grids der Nitrile, aus den CYM/HIP-Simulationen von CatS.

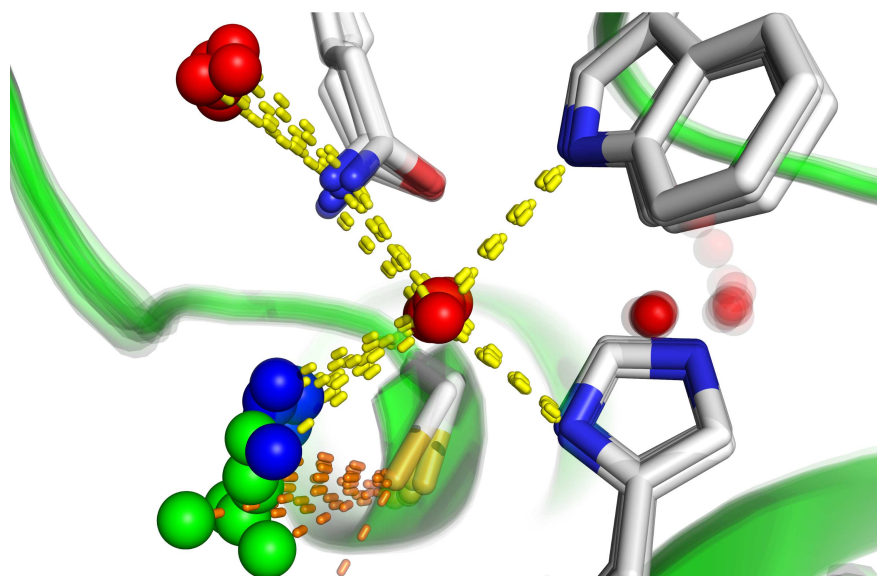


Abbildung A.5.: Überlagerung der Grid-Analyse jedes CatS-Nitrils, basierend auf den CYS/HIS-Simulationen. Wasser-Cluster sind als rote, Nitril-Kohlenstoff-Cluster als grüne und Nitril-Stickstoff-Cluster als blaue Kugeln gezeigt. Die Cluster für die Kohlenstoffatome und die Stickstoffatome der Nitrile nehmen räumlich eine sehr ähnliche Position ein.

## A.2. USP28

### A.2.1. Thermodynamische Korrekturen zu den SCF-Energien

Zur Berechnung von  $\Delta G_A$  wurden die in Gaussian berechnete Korrektur (*Thermal Correction to Free Energy*) für jede Verbindung zur SCF-Energie (*Electronic Energy (EE)*) dazu addiert, und anschließend die Differenz zwischen Übergangszustand und Referenz-Zustand (Cysteamin + Nitrilverbindung) gebildet. Da die von Gaussian berechnete Korrektur von einem gasförmigen Standard-Zustand ausgeht (d. h. mit einem Standard-Druck von 1 atm rechnet, nicht mit einer Standard-Konzentration von 1 mol/l), muss bei jeder Verbindung 1.89 kcal/mol addiert werden, um Werte entsprechend einer Verbindung in Lösung zu erhalten (in der Literatur oft als *Standard State Conversion* beschrieben).<sup>202</sup> Bei unimolekularen Reaktionen hebt sich diese Konversion auf; da aber bei der hier diskutierten bimolekularen Reaktion nur ein Molekül gebildet wird, „sinkt“ die Aktivierungsenergie effektiv um 1.89 kcal/mol im Vergleich zu den direkt aus Gaussian übernommenen Werten. Zusätzlich zu den Energien des Übergangszustandes relativ zum Referenz-Zustand sind analog die relativen Energien des Produkts angegeben ( $\Delta E_{Prod.}$  und  $\Delta G_{Prod.}$ ). Die hier berechnete freie Reaktionsenthalpien ergibt Werte niedrige positive Werte, was qualitativ in Übereinstimmung mit der Reversibilität der Reaktion ist.

Verbindung	$\Delta E_A$ (SCF) [kcal/mol]	$\Delta G_A$ [kcal/mol]	$\Delta E_{Prod.}$ (SCF) [kcal/mol]	$\Delta G_{Prod.}$ [kcal/mol]
<b>U19</b>	11.8	21.9	-8.36	2.06
<b>U20</b>	13.1	22.8	-5.35	4.75
<b>U21</b>	11.0	20.9	-7.28	2.94
<b>U22</b>	13.9	23.2	-4.44	5.68
<b>U23</b>	13.3	23.0	-4.93	5.26
<b>U24</b>	12.0	21.8	-6.35	3.74
<b>U25</b>	12.8	22.7	-5.45	4.68
<b>U26</b>	9.54	19.1	-9.33	1.11
<b>U27</b>	15.5	24.0	-3.32	6.80

Tabelle A.1.: Aktivierungsenergien und freie Aktivierungsenthalpien für die USP28-Warheads.

Verbindung	$\Delta E_A$ (SCF) [kcal/mol]	$\Delta G_A$ [kcal/mol]	$\Delta E_{Prod.}$ (SCF) [kcal/mol]	$\Delta G_{Prod.}$ [kcal/mol]
<b>U28</b>	19.6	28.6	16.2	26.6
<b>U29</b>	21.1	30.1	18.2	28.6
<b>U30</b>	18.4	29.6	14.7	26.9
<b>U31</b>	21.6	30.8	18.7	29.1
<b>U32</b>	21.4	30.2	18.6	28.9
<b>U33</b>	19.3	28.5	15.8	26.1
<b>U34</b>	20.7	29.2	17.8	27.9
<b>U35</b>	30.2	39.4	29.5	39.1
<b>U36</b>	23.5	32.6	21.4	31.6
<b>U37</b>	15.5	24.6	10.6	21.1

Tabelle A.2.: Aktivierungsenergien und freie Aktivierungsenthalpien für die USP28-Alkin-Warheads.

## A.3. ChlaDUB1

### A.3.1. Thermodynamische Korrekturen zu den SCF-Energien

Verbindung	$\Delta E_A$ (SCF) [kcal/mol]	$\Delta G_A$ [kcal/mol]	$\Delta E_{Prod.}$ (SCF) [kcal/mol]	$\Delta G_{Prod.}$ [kcal/mol]
<b>Chla1</b>	8.53	18.4	-10.3	0.31
<b>Chla2</b>	10.2	20.6	-7.99	2.34
<b>Chla4</b>	8.61	18.7	-10.4	0.47
<b>Chla5</b>	8.21	18.3	-10.9	-0.32
<b>Chla6</b>	7.04	17.6	-11.5	-0.57
<b>Chla7</b>	13.5	23.2	-10.0	-0.50
<b>Chla8</b>	11.9	21.5	-12.3	-1.50
<b>Chla9</b>	13.3	23.6	-11.0	-0.05
<b>Chla10</b>	7.57	17.9	-14.4	-3.61
<b>Chla3</b>	8.97	14.4	-11.1	-4.52
Szenario 1 u. 3				
<b>Chla3</b>	9.42	19.7	-8.01	2.60
Szenario 1 u. 4				
<b>Chla3</b>	8.99	19.1	-11.1	-0.31
Szenario 2 u. 3				
<b>Chla3</b>	9.41	19.6	-8.00	2.40
Szenario 2 u. 4				

Tabelle A.3.: Aktivierungsenergien und freie Aktivierungsenthalpien für die ChlaDUB1-Nitril-Warheads.