



UNIVERSITÄT ZU LÜBECK

Aus dem Institut für Kardiogenetik
der Universität zu Lübeck
Direktorin: Prof. Dr. rer. nat. Jeanette Erdmann

In silico-Analysen zur Identifikation von ADAMTS7-Inhibitoren
für eine Therapie der koronaren Herzkrankheit

Inauguraldissertation
zur
Erlangung der Doktorwürde
der Universität zu Lübeck

Aus der Sektion Naturwissenschaften

vorgelegt von
Michaela Müller
aus Wismar

Lübeck 2017

1. Berichterstatterin: Prof. Dr. rer. nat. Jeanette Erdmann
 2. Berichterstatter: Prof. Dr. rer. nat. Hauke Busch
- Tag der mündlichen Prüfung: 4. Mai 2018

Zum Druck genehmigt. Lübeck, den 15. Mai 2018

«*Nobody said it was easy*»
Coldplay, The Scientist

Abstract

Genome-wide association studies identified *ADAMTS7* as a risk locus for coronary artery disease. In carotid arteries of rats, neointima formation after balloon-mediated injury goes along with enhanced *Adamts7* expression. vascular injury. Although a causal link between *ADAMTS7* and coronary artery disease remains to be proven, inhibition of *ADAMTS7* represents a potential new target for intervention in this disease. Therefore the aim of this thesis is the identification of potential effective, selective *ADAMTS7* inhibitors with *in silico* methods.

ADAMTS7, a member of the ‘a disintegrin and metalloproteinase with thrombospondin motifs’ (*ADAMTS*) family of proteins, contains a catalytic zinc ion in the binding site of its metalloproteinase domain. The structure of *ADAMTS7* and its inhibitors are unknown. Homology modeling was used to model a part of the three-dimensional structure of the protein (metalloproteinase domain and disintegrin-like domain). In a molecular dynamics simulation flexible regions were observed in the *ADAMTS7* binding site. They were modelled in a loop sampling. The conformational ensemble of *ADAMTS7* consists of five binding site conformations. These allow the modeling of the adjustments of the binding site to different (yet unknown) ligands.

Sequences comparisons of the *ADAMTS* family confirmed the high conservation among family members. A loop region (L1) in the binding site is less conserved. Despite strong sequential conservation there are differences in the three-dimensional structures of *ADAMTS* proteins, especially in the L1 loop region. This loop region represents a potential binding region for selective *ADAMTS7* inhibitors.

A structure-based pharmacophore was proposed for the *ADAMTS7* binding site. In a virtual screening the pharmacophore was employed to screen a compound library. Candidate ligands were docked into the *ADAMTS7* conformational ensemble and additional filter criteria were applied to the resulting docking poses. An enrichment of potential *ADAMTS7* inhibitors with high affinity was assumed to be in the first 5 % of compounds. This corresponds to 79 compounds with lowest docking score.

The binding affinity of these 79 compounds to the other *ADAMTS* proteins was tested in an inverse screening. 18 compounds were identified to bind worse to the other *ADAMTS* binding sites compared to *ADAMTS7* and were therefore classified as potential selective for *ADAMTS7*. Also seven potential non selective hits were identified that bind to at least nine other *ADAMTS* proteins equal or better compared to *ADAMTS7*.

Analysing the docking poses of the 79 compounds with *ADAMTS7* a unique binding which explains affinity respectively selectivity was not revealed. In nearly all of the docking poses of the 18 potential selective hits hydrogen bonds to the L1 loop were observed. In addition a correlation between size and volume of the potential *ADAMTS7* inhibitors and *ADAMTS7* selectivity was revealed.

The main contribution of this thesis is the proposal of 79 potential *ADAMTS7* inhibitors. One quarter was classified as potential selective *ADAMTS7* inhibitors. Selectivity might be the result of interactions with the less conserved L1 loop. These are the first steps toward developing a new therapeutic intervention for coronary artery disease.

Zusammenfassung

Mittels genomweiter Assoziationsstudien konnte *ADAMTS7* als Risikolokus für die koronare Herzkrankheit (KHK) identifiziert werden. Tierversuche an Ratten zeigten eine erhöhte Expression von *Adamts7* nach einer gezielten Verletzung der Halsschlagader. Die Inhibition der *ADAMTS7*-Proteinfunktion könnte demnach protektiv gegen Atherosklerose und KHK wirken und ein neuer Ansatz für eine KHK-Therapie sein. Daher war das Ziel dieser Arbeit die Suche nach und die Identifizierung von potenziell wirksamen, selektiven *ADAMTS7*-Inhibitoren mittels *in silico* Methoden.

ADAMTS7 gehört zur *ADAMTS*-Proteinfamilie, welche eine Bindetasche mit einem katalytischen Zink-Ion in der Metalloproteinase-Domäne besitzen. Es sind keine Inhibitoren oder strukturellen Daten von *ADAMTS7* bekannt. Ein Teil der 3D-Struktur von *ADAMTS7* (Metalloproteinase- und Disintegrin-ähnliche Domäne) konnte mit einem Homology Modeling erstellt werden. In einer Moleküldynamik-Simulationen wurden flexible Loop-Bereiche in der *ADAMTS7*-Bindetasche beobachtet, welche anhand eines Samplings modelliert wurden. Das *ADAMTS7*-Conformational Ensemble stellte insgesamt fünf Konformationen der *ADAMTS7*-Bindetasche dar. Dies erlaubte die Modellierung der möglichen Anpassung der Bindetasche an verschiedene (noch unbekannte) Liganden.

Sequenzvergleiche bestätigten eine hohe Konservierung innerhalb der *ADAMTS*-Familie. In der Bindetasche wurde ein Loop-Bereich (L1) aufgedeckt, welcher vergleichsweise weniger stark konserviert ist. Trotz hoher sequenzieller Konservierung konnten Unterschiede in der räumlichen Struktur der *ADAMTS*-Proteine, vor allem in der L1-Loop-Region, beobachtet werden. Der L1-Loop stellt eine potenzielle Binde-region selektiver *ADAMTS7*-Inhibitoren dar.

Die Charakterisierung der *ADAMTS7*-Bindetasche wurde mithilfe eines strukturbasierten Pharmacophores vorgenommen. Das *ADAMTS7*-Pharmacophore besteht aus sieben Features. Im Virtual Screening wurde eine Datenbank (rund 23 Mio. Verbindungen) mit dem *ADAMTS7*-Pharmacophore gescreent. Die verbleibenden Moleküle wurden in das *ADAMTS7*-Conformational Ensemble gedockt und die Docking-Posen anhand weiterer Filterkriterien bewertet. Im letzten Filterschritt wurde eine Anreicherung von potenziellen *ADAMTS7*-Inhibitoren mit einer hoher Bindeaffinität in den besten 5 % der Verbindungen prognostiziert. Dies entspricht 79 Verbindungen mit den niedrigsten Docking-Scores.

Mithilfe eines inversen Screenings wurde die Bindeaffinität der 79 Verbindungen zu den 18 weiteren *ADAMTS*-Proteinen und somit deren *ADAMTS7*-Selektivität abgeschätzt. Es gelang 18 Verbindungen zu identifizieren, die an die anderen *ADAMTS*-Bindetaschen schlechter als an *ADAMTS7* binden und als potenziell selektiv klassifiziert wurden. Außerdem konnten sieben potenziell nicht selektive Hits identifiziert werden, welche an mindestens neun weitere *ADAMTS*-Proteine gleich gut oder besser als an *ADAMTS7* binden.

Anhand der Analyse der Wechselwirkungen der 79 Verbindungen zur *ADAMTS7*-Bindetasche konnte kein eindeutiges Bindemuster auf dem eine Affinität bzw. Selektivität zu *ADAMTS7* basieren könnte aufgedeckt werden. Bei den 18 potenziell selektiven Hits zeigte sich, dass in fast allen Docking-Posen Wasserstoffbrücken zum L1-Loop vorliegen. Zudem konnte ein Zusammenhang zwischen Größe und Volumen der potenziellen Inhibitoren und der Selektivität zu *ADAMTS7* erkannt werden.

Als Resultat wurden somit 79 mögliche *ADAMTS7*-Inhibitoren postuliert. Etwa ein Viertel dieser wurde als potenziell selektiv klassifiziert, wobei die Selektivität wahrscheinlich durch das Wechselwirken mit dem sequenziell weniger stark konservierten und strukturell differenziellen L1-Loop erreicht wird. Dies sind die ersten Schritte zur Entwicklung einer neuen KHK-Therapie.

Danksagung

An dieser Stelle möchte ich mich bei den Menschen bedanken, die mir durch ihre Mithilfe bei der Erstellung dieser Arbeit eine große Unterstützung waren. Zunächst möchte ich meiner Doktormutter Jeanette Erdmann für die vertrauensvolle Aufnahme im Institut für Kardiogenetik und die Gelegenheit an diesem interessanten Thema zu arbeiten danken. Außerdem danke ich Hauke Busch für die Übernahme des Zweitgutachtens.

Ein besonderer Dank an dieser Stelle gilt Stephanie Tennstedt. Bereits bei meiner Bachelorarbeit hat sie mich begleitet. Acht Jahre später schließt sich nun meine akademische Laufbahn und unser gemeinsames wissenschaftliches Kapitel mit dieser Dissertation. Ohne Stephe wäre die Arbeit nicht das, was sie heute ist; vielen Dank für die vielen Gespräche und Anregungen.

Ebenso gilt mein Dank allen Mitarbeitern und Alumni des Instituts für Kardiogenetik für die angenehme und hilfsbereite Arbeitsatmosphäre. Besonders erwähnt sei hier die Bioinformatik-Sektion des Instituts, vor allem die *Bioinformagicians* Mariana Barnes, Ingrid Brænne und Benedikt Reiz.

Nicht vergessen werden sollen außerdem all jene, die moralischen Beistand leisteten und für Ablenkung außerhalb des Wissenschaftskosmos sorgten: Meine Familie, vor allem meine Eltern Beate und Dietmar und meine Schwester Janine, sowie meine Freunde, die mich auf allen Etappen dieser Arbeit begleiteten. Besonders hervorgehoben werden muss an dieser Stelle mein Freund Thomas. Während wir zur gleichen Zeit promovierten, haben wir es geschafft uns gegenseitig zu (unter)stützen. Vielen Dank dafür!

Inhaltsverzeichnis

Abstract	VII
Zusammenfassung	IX
Danksagung	XI
Abkürzungsverzeichnis	XVII
1 Einleitung	1
1.1 Aktueller Stand der kardiovaskulären Genomforschung	2
1.2 Risikogen <i>ADAMTS7</i>	6
1.3 ADAMTS-Proteine	7
1.3.1 Domänenstruktur	8
1.3.2 Räumliche Gestalt der ADAMTS-Proteine anhand bekannter Kristallstrukturen	9
1.3.3 Substrate und Inhibitoren der ADAMTS-Proteine und assoziierte Erkrankungen	12
1.4 <i>ADAMTS7</i>	14
1.5 Zielsetzung	15
2 Methoden	17
2.1 Sequenzvergleich	17
2.2 Sekundärstrukturvorhersage	18
2.3 3D-Proteinstrukturen und deren Modellierung	19
2.3.1 Homology Modeling	20
2.3.2 Conformational Sampling	22
2.4 3D-Proteinstrukturen und deren Validierung	23
2.5 Molekülmechanik	24
2.5.1 Kraftfelder	24
2.5.2 Energieminimierung	25
2.5.3 Moleküldynamik-Simulation	26
2.5.4 Methoden zum Strukturvergleich: RMSD und RMSF	27
2.6 Pharmacophoremmodellierung	28
2.6.1 Definition und Charakterisierung der Proteinbindetasche	28
2.6.2 Pharmacophore	29
2.6.3 <i>e-Pharmacophore</i> -Methode	29
2.7 Docking und Scoring	30
2.7.1 Docking mit AutoDock4 und AutoDockVina	31

2.7.2	Docking mit PLANTS	32
2.7.3	Docking mit Schrödinger Glide	33
2.7.4	Inverses Screening	34
2.8	Virtual Screening	35
2.8.1	Rule of 5	36
2.8.2	Diversität von Liganden	37
3	Ergebnisse und Diskussion: Modellierung der ADAMTS7-Proteinstruktur	39
3.1	Wahl des/der Template(s)	39
3.2	Modellerstellung	42
3.3	Modellvalidierung der ADAMTS7-Modelle	44
3.3.1	Abdeckung der Sequenz und Strukturvergleich	44
3.3.2	Disulfidbrücken	45
3.3.3	Zink-Koordinierung	46
3.3.4	Globale und lokale Qualität	46
3.3.5	MD-Simulation des ADAMTS7-Modells	47
3.4	Conformational Ensembles von flexiblen Regionen in ADAMTS7-Bindetasche	50
3.5	Zusammenfassung	53
4	Ergebnisse und Diskussion: Untersuchungen der ADAMTS-Familie	55
4.1	ADAMTS-Modelle	55
4.2	Sequenz- und Strukturvergleich	59
4.2.1	Vergleich der Metalloproteinase-Domäne	59
4.2.2	Vergleich der Bindetasche	62
4.3	Bindemuster von bekannten Inhibitoren in ADAMTS-Proteinen	65
4.4	Zusammenfassung	68
5	Ergebnisse und Diskussion: ADAMTS7-Virtual Screening	71
5.1	ADAMTS7-Pharmacophore	71
5.1.1	Evaluierung der ADAMTS7-Pharmacophore-Features	72
5.1.2	Erweiterung des ADAMTS7-Pharmacophores	74
5.1.3	Beschreibung des ADAMTS7-Pharmacophore-Modells	75
5.2	Redocking zur Evaluierung von Dockingmethoden	75
5.3	Virtual Screening	78
5.3.1	Datenbank	79
5.3.2	Pharmacophore-Screening	80
5.3.3	Ensemble Docking	80
5.3.4	Screening und Filterung der Docking-Posen	81
5.4	Zusammenfassung	83
6	Ergebnisse und Diskussion: Weiterführende Untersuchungen der potenziellen ADAMTS7-Inhibitoren	85
6.1	Wechselwirkungen zwischen ADAMTS7-Bindetasche und potenziellen Inhibitoren	85
6.2	Inverses Screening	88

6.2.1	Potenziell nicht selektive Hits	92
6.2.2	Potenziell selektive Hits	94
6.3	Wechselwirkungen zwischen ADAMTS7-Bindetasche und potenziell (nicht) selektiven Inhibitoren	96
6.3.1	Analyse des Zusammenhangs zwischen Länge, Volumen und Selektivität von potenziellen ADAMTS7-Inhibitoren	96
6.3.2	Wechselwirkungen der potenziell nicht selektiven Hits	97
6.3.3	Wechselwirkungen der potenziell selektiven Hits	100
6.3.4	MD-Simulationen der potenziell selektiven Hits	108
6.4	Zusammenfassung	113
7	Zusammenfassung und Ausblick	115
A	Anhang	123
A.1	Ergebnisse: Modellierung der ADAMTS7-Proteinstruktur	123
A.1.1	Wahl des/der Template(s)	123
A.1.2	Modellvalidierung der ADAMTS7-Modelle	125
A.2	Ergebnisse: Untersuchungen der ADAMTS-Familie	127
A.2.1	Sequenz- und Strukturvergleich	127
A.2.2	Bindemuster von bekannten Inhibitoren in ADAMTS-Proteinen	131
A.3	Ergebnisse: Virtual Screening	132
A.3.1	Redocking zur Evaluierung von Dockingmethoden	132
A.3.2	Virtual Screening	134
A.4	Ergebnisse: Weiterführende Untersuchungen der potenziellen ADAMTS7-Inhibitoren	142
A.4.1	Inverses Screening	142
A.4.2	Wechselwirkungen zwischen ADAMTS7-Bindetasche und potenziell (nicht) selektiven Inhibitoren	147
	Literaturverzeichnis	151
	Tabellenverzeichnis	160
	Abbildungsverzeichnis	162
	Formelverzeichnis	165
	Publikationen	167

Abkürzungsverzeichnis

ADAM	engl. <i>a disintegrin and metalloproteinase</i>
ADAMTS	engl. <i>a disintegrin and metalloproteinase with thrombospondin motifs</i>
COMP	engl. <i>cartilage oligomeric matrix protein</i>
GWAS	Genomweite Assoziationsstudie
KHK	Koronare Herzkrankheit (engl. <i>coronary artery disease</i>)
MD-Simulation	Moleküldynamik-Simulation
MI	Myokardinfarkt
MMP	engl. <i>matrix metalloproteinase</i>
PDB	Protein Data Bank
RMSD	Mittlere quadratische Abweichung (engl. <i>root-mean-square deviation</i>)
RMSF	Mittlere quadratische Schwankung (engl. <i>root-mean-square fluctuation</i>)
SNP	Einzelnukleotid-Polymorphismus (engl. <i>single nucleotide polymorphism</i>)

1 Einleitung

Jeder Mensch kommt früher oder später in seinem Leben mit Medikamenten in Berührung. Ob bei Impfungen, Kopfschmerzen oder (chronischen) Erkrankungen: Arzneimittel können helfen, Krankheiten vorzubeugen, zu heilen oder bestimmte Symptome zu lindern. Arzneimittel sind nach § 2 des deutschen Arzneimittelgesetzes [1] „Stoffe oder Zubereitungen aus Stoffen,

1. die zur Anwendung im oder am menschlichen oder tierischen Körper bestimmt sind und als Mittel mit Eigenschaften zur Heilung oder Linderung oder zur Verhütung menschlicher oder tierischer Krankheiten oder krankhafter Beschwerden bestimmt sind oder
2. die im oder am menschlichen oder tierischen Körper angewendet oder einem Menschen oder einem Tier verabreicht werden können, um entweder
 - a) die physiologischen Funktionen durch eine pharmakologische, immunologische oder metabolische Wirkung wiederherzustellen, zu korrigieren oder zu beeinflussen oder
 - b) eine medizinische Diagnose zu erstellen.“

Das Erforschen und Finden von neuen Medikamenten und somit neuen Therapien ist eine große Herausforderung. Die Arzneimittelforschung selbst ist ein multidisziplinäres Arbeitsfeld mit dem Ziel neue Substanzen zu finden und daraus ein wirksames Therapeutikum zu entwickeln [2, S. 2].

Der Prozess der Medikamentenentwicklung umfasst viele Stufen. Von der Erforschung eines angreifbaren Targets (häufig Proteine) für eine bestimmten Erkrankung und der Suche nach potenziellen Wirkstoffen (Liganden), über vorklinische Tests und klinische Studien, in denen der Wirkstoff erstmals am Menschen getestet wird, bis zur potenziellen Zulassung vergehen im Durchschnitt etwa 13 Jahre (siehe Abb. 1.1) [3].

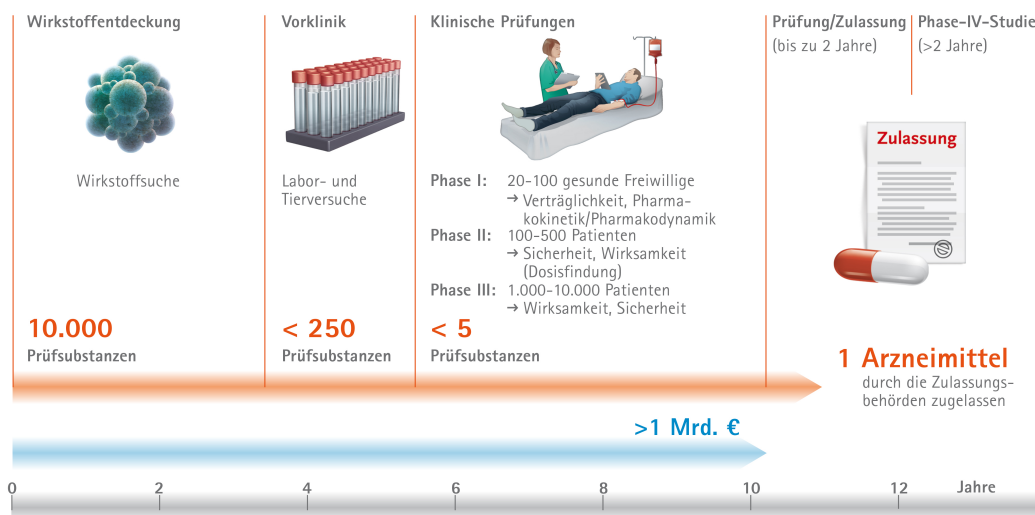


Abbildung 1.1: Stufen der Medikamentenentwicklung (aus [4]).

Der klassische Weg bei der Suche nach neuen Wirkstoffen ist das Testen von vielen Substanzen auf zellulärem Gewebe. Hierzu gibt es Hochdurchsatzverfahren, in denen hunderttausende Substanzen geprüft

werden können [5]. Mit den Methoden der Röntgenkristallographie und der NMR-Spektroskopie ist es heute zudem möglich, die räumliche Gestalt von Proteinen und ihren Liganden aufzuklären und so den Wirkmechanismus genau zu entschlüsseln [5]. Die neusten Methoden in der Medikamentenentwicklung umfassen außerdem computergestützte Methoden, die bei der Suche nach Wirkstoffen oder bei der Entwicklung neuer Wirkstoffe helfen [6].

Neu entwickelte Medikamente sollen gezielt wirken, möglichst keine Nebenwirkungen haben, oder eine bessere Wirksamkeit als bekannte Arzneistoffe aufweisen. Jedoch ist die Anzahl der neu in die Therapie eingeführten Wirkstoffe in den letzten Jahrzehnten stetig gesunken [2, S. 2]. So wurden in den Jahren 1960-1969 70-100 neue Wirkstoffe eingeführt und in den 1990er und 2000er Jahren nur noch jeweils 40-45 [2, S. 2]. Gründe dafür sind ein bereits bestehender, hoher Therapiestandard in vielen Indikationsgebieten, ein deutlich verstärktes Sicherheitsbewusstsein sowie zunehmende Anforderungen vonseiten des Gesetzgebers und die steigenden Kosten für Forschung und Entwicklung [2, S. 2]. Zusammengefasst ist die Entwicklung von Arzneistoffen von der Forschung bis zur Zulassung ein Jahre andauernder, teurer Prozess. Für ein Medikament mit neuem Wirkstoff, welches es bis zur Zulassung schafft, muss mit Kosten von 1,0 bis 1,6 Milliarden US-Dollar gerechnet werden [3]. Außerdem besteht in jedem Schritt das Risiko, dass sich der Wirkstoff als nicht wirksam erweist oder unerwünschte Nebenwirkungen auftreten, was schlussendlich zum Abbruch des Prüfverfahrens führen würde.

Diese Arbeit befasst sich vor allem mit der ersten Stufe in der Medikamentenentwicklung, der Suche nach neuen Wirkstoffen. Ziel ist es, mit *in silico*-Methoden Moleküle zu finden, die eine bestimmte Wirkung auf das Proteintarget haben, welches mit Herz-Kreislaufkrankungen im Zusammenhang steht. Insbesondere sollen im Folgenden die kardiovaskulären Erkrankungen und der Angriffspunkt bzw. das Target, das Protein ADAMTS7, näher erläutert werden.

1.1 Aktueller Stand der kardiovaskulären Genomforschung

Die koronare Herzkrankheit (KHK) und ihre schwerste Komplikation, der Myokardinfarkt (MI), sind weltweit eine der häufigsten Todesursachen [8].

Bei der KHK handelt es sich um eine Durchblutungsstörung der Herzkranzgefäße mit Veränderungen der Gefäßmuskelschicht ausgelöst durch Atherosklerose wie in Abbildung 1.2 dargestellt. Der Gefäßdurchmesser ist durch Plaqueablagerungen, die unter anderem aus Lipiden und Blutbestandteilen bestehen, verringert. Dies führt zu einer verminderten Durchblutung. Kommt es infolgedessen zur Sauerstoffunterversorgung des Herzmuskels durch einen Verschluss einer Arterie, kann dies einen MI zur Folge haben [9].

Die Ursachen von KHK und MI werden seit langem erforscht. Risikofaktoren für KHK und

MI sind unter anderem Rauchen, mangelnde Bewegung sowie die Ernährungsweise, Erkrankungen wie Diabetes oder Hypercholesterinämie und das Lebensalter [10]. Neben diesen allgemein bekannten Risi-

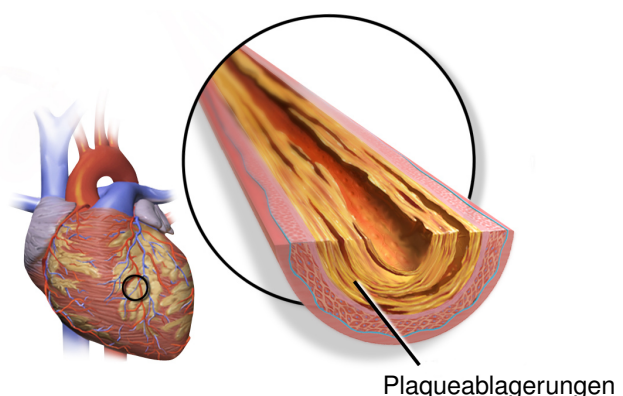


Abbildung 1.2: KHK wird ausgelöst durch die Verkalkung der Arterien, wodurch es zu einem Verschluss und damit zu einem MI kommen kann (aus [7]).

kofaktoren, die ursächlich für diese Art von Erkrankungen sein können, spielen auch genetische Faktoren eine große Rolle [11]. So wurde früher fälschlicherweise angenommen, dass es sich bei KHK und MI um Zivilisationskrankheiten handelt. Jedoch konnten bei Untersuchungen von Mumien aus vier Regionen, die aus Zeiträumen zwischen 3100 v. Chr. und 1500 n. Chr. stammen, in einem Drittel dieser sicher oder wahrscheinlich Atherosklerose in den Arterien nachgewiesen werden [12]. Diese Ergebnisse widerlegen, dass es sich bei der KHK und dem MI um reine Zivilisationskrankheiten handelt. KHK und MI sind komplexe Erkrankungen [13]. Es wird geschätzt, dass eine genetische Veranlagung etwa 40 bis 60 % des KHK-Risikos ausmacht [11].

Durch genomweite Assoziationsstudien (GWAS) konnten bereits eine Vielzahl an Genloci mit bestimmten Erkrankungen assoziiert werden. Dabei werden, anhand von genetischem Material von Probanden, Varianten in Genen mit dem Auftreten von bestimmten Merkmalen in Verbindung gebracht. Hierfür wird genetisches Material von Patienten mit gesunden Kontrollen verglichen, um SNPs (engl. *single nucleotide polymorphism*) mit dem Auftreten von Krankheiten zu verknüpfen. Das Wissen um die Assoziation von SNPs mit einer Erkrankung kann bei der Früherkennung und der Entwicklung neuer Therapien helfen [9]. In den letzten Jahren konnten GWAS beachtliche Ergebnisse erzielen, indem unbekannte Gene oder Genregionen identifiziert wurden, die mit komplexen Erkrankungen, wie z.B. KHK oder Diabetes, assoziiert wurden [9]. Die ersten GWAS zur KHK [13, 14, 15] aus dem Jahre 2007 führten unter anderem zur Identifizierung eines Locus auf Chromosom 9p21 (rs4977574) sowie weiteren Risikoloci. In den folgenden Jahren konnten durch weitere GWAS (z. B. [16]) weitere KHK-Risikoloci aufgedeckt werden.

Der Zusammenschluss von Forschergruppen und die Bildung von großen, internationalen Konsortien ermöglichte im Rahmen von CARDIoGRAM [17], C4D [18] und CARDIoGRAMplusC4D [19, 20] die Untersuchung von Daten hunderttausender KHK-Fälle und Kontrollen. Die Meta-Analyse von 14 GWAS im Rahmen des CARDIoGRAM-Konsortiums, die rund 22 000 KHK-Fälle und etwa 65 000 Kontrollen umfasste, ergab 13 neue KHK-Loci und bestätigte zehn der zwölf bereits bekannten KHK-Loci [17]. Die 13 neuen Risikoloci erhöhen das KHK-Risiko um 6 bis 17 % pro Allel. Innerhalb des C4D-Konsortiums wurden nicht nur Probanden mit europäischer Abstammung untersucht. So konnten, basierend auf den genetischen Daten von rund 15 000 Individuen (Europäer und Probanden mit südasiatischer Herkunft) mit KHK und 15 000 Kontrollen, fünf weitere KHK-Loci identifiziert werden [18]. Im Rahmen des CARDIoGRAMplusC4D-Konsortiums [19] wurde eine noch größere Kohorte (etwa 64 000 KHK-Fälle und 131 000 Kontrollen) untersucht. Die Analysen dieses Konsortiums deckten 15 neue, genomweit signifikante KHK-Loci auf. Des Weiteren gelang es Nikpay und Kollegen [20] weitere zehn Risikoloci mittels einer Meta-Analyse ($\approx 185\,000$ KHK-Fälle und Kontrollen) zu identifizieren. In den aktuellsten Studien von Webb *et al.* [22] (42 335 KHK-Fälle und 78 240 Kontrollen) und Howson *et al.* [23] (88 192 KHK-Fälle und 162 544 Kontrollen) konnten, anhand des Vergleichs des genetischen Materials, insgesamt 20 weitere SNPs mit KHK assoziiert werden. Die GWAS sowie deren Meta-Analysen bestätigen somit bekannte Risikoloci und decken neue unbekannte Regionen und deren KHK-Assoziation auf. Bis Juli 2017 wurden insgesamt 75 Risikoloci genomweit signifikant ($p < 5 \cdot 10^{-8}$) durch GWAS mit KHK assoziiert (siehe Abb. 1.3).

In Tabelle 1.1 sind die bisher aufgedeckten und mit KHK assoziierten GWAS-Loci aufgelistet. Hat ein Individuum eine oder mehrere dieser Varianten, bedeutet das nicht automatisch, dass eine KHK auftreten wird. Es heißt lediglich, dass das Risiko dieser Person für KHK erhöht ist, im Vergleich zu einer Person, die diese Variante(n) nicht trägt. Für die meisten mit den Risikoloci assoziierten Gene ist außerdem die Pathophysiologie noch nicht bekannt [10]. Kritikpunkt an den GWAS bezüglich KHK ist, dass die Effektstärken

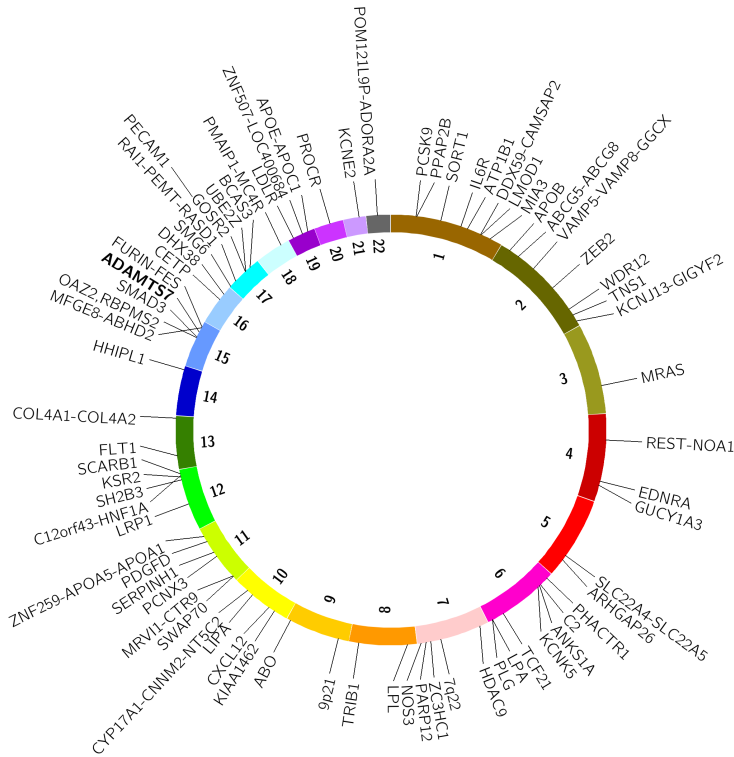


Abbildung 1.3: Circosplot mit bekannten KHK-Risikoloci sortiert nach Chromosom. Grafik erstellt mit Circos [21].

der gefunden Varianten relativ klein sind und sie bis heute wenige diagnostische oder therapeutische Ansätze hervorbrachten [10]. Dabei bietet die Veränderung von genetischen Faktoren ein großes Potenzial zur Risikoreduzierung [10].

Tabelle 1.1: Mit KHK assoziierte GWAS-Genloci ($p < 5 \cdot 10^{-8}$) sortiert nach Chromosom. Zusätzlich aufgeführt sind die assoziierten Gen(e), das Risikoallel mit Häufigkeit sowie die zugehörige Quellenangabe.

Chromosom	SNP	Gen(e) in Region	Risikoallel (Häufigkeit)	Quelle
1	rs1892094	<i>ATP1B1</i>	T (0,50)	[23]
	rs6700559	<i>DDX59-CAMSAP2</i>	T (0,47)	[23]
	rs4845625	<i>IL6R</i>	T (0,47)	[19]
	rs2820315	<i>LMOD1</i>	T (0,30)	[23]
	rs17465637	<i>MIA3</i>	C (0,74)	[13]
	rs11206510	<i>PCSK9</i>	T (0,81)	[24, 25]
	rs17114036	<i>PPAP2B</i>	A (0,91)	[17]
	rs599839	<i>SORT1</i>	A (0,78)	[13]
	2	rs6544713	<i>ABCG5-ABCG8</i>	T (0,30)
rs515135		<i>APOB</i>	G (0,83)	[19]
rs1801251		<i>KCNJ13-GIGYF2</i>	A (0,35)	[22]
rs2571445		<i>TNS1</i>	A (0,39)	[23]
rs1561198		<i>VAMP5-VAMP8-GGCX</i>	A (0,45)	[19]
rs6725887		<i>WDR12</i>	C (0,15)	[17]
rs2252641		<i>ZEB2</i>	G (0,46)	[19]
3	rs9818870	<i>MRAS</i>	C (0,18)	[16]
4	rs1878406	<i>EDNRA</i>	T (0,15)	[19]
	rs7692387	<i>GUCY1A3</i>	G (0,81)	[19]
	rs17087335	<i>REST-NOA1</i>	T (0,21)	[20]
5	rs246600	<i>ARHGAP26</i>	T (0,48)	[23]

Tabelle 1.1 (Fortsetzung): Mit KHK assoziierte GWAS-Genloci ($p < 5 \cdot 10^{-8}$) sortiert nach Chromosom. Zusätzlich aufgeführt sind die assoziierten Gen(e), das Risikoallel mit Häufigkeit sowie die zugehörige Quellenangabe.

Chromosom	SNP	Gen(e) in Region	Risikoallel (Häufigkeit)	Quelle
5	rs273909	<i>SLC22A4-SLC22A5</i>	C (0,14)	[19]
6	rs17609940	<i>ANKS1A</i>	G (0,75)	[17]
	rs3130683	<i>C2</i>	T (0,86)	[22]
	rs10947789	<i>KCNK5</i>	T (0,76)	[19]
	rs12526453	<i>PHACTR1</i>	C (0,65)	[17]
	rs4252120	<i>PLG</i>	T (0,73)	[19]
	rs3798220	<i>LPA</i>	C (0,02)	[17, 27]
	rs12190287	<i>TCF21</i>	C (0,62)	[17]
7	rs10953541	<i>BCAP29</i> (7q22-Lokus)	C (0,80)	[18]
	rs2023938	<i>HDAC9</i>	G (0,10)	[19]
	rs3918226	<i>NOS3</i>	T (0,06)	[20]
	rs10237377	<i>PARP12</i>	T (0,35)	[23]
	rs11556924	<i>ZC3HC1</i>	C (0,62)	[17]
8	rs264	<i>LPL</i>	G (0,86)	[19]
	rs2954029	<i>TRIB1</i>	A (0,55)	[19, 26]
9	rs579459	<i>ABO</i>	C (0,21)	[17]
	rs4977574	<i>CDKN2A/B-ANRIL</i> (9p21-Lokus)	G (0,46)	[13, 14, 15]
10	rs1746048	<i>CXCL12</i>	C (0,87)	[13]
	rs12413409	<i>CYP17A1-CNNM2-NT5C2</i>	G (0,89)	[17]
	rs2505083	<i>KIAA1462</i>	C (0,38)	[18]
	rs1412444	<i>LIPA</i>	T (0,42)	[18]
11	rs11042937	<i>MRVII-CTR9</i>	T (0,49)	[22]
	rs12801636	<i>PCNX3</i>	A (0,23)	[23]
	rs974819	<i>PDGFD</i>	T (0,32)	[18]
	rs590121	<i>SERPINH1</i>	T (0,30)	[23]
	rs10840293	<i>SWAP70</i>	A (0,55)	[20]
	rs964184	<i>ZNF259-APOA5-APOA1</i>	G (0,13)	[17]
12	rs2258287	<i>C12orf43-HNF1A</i>	A (0,34)	[23]
	rs11830157	<i>KSR2</i>	G (0,36)	[20]
	rs11172113	<i>LRP1</i>	C (0,41)	[22]
	rs11057830	<i>SCARB1</i>	A (0,15)	[22, 23]
	rs3184504	<i>SH2B3</i>	T (0,44)	[17]
13	rs4773144	<i>COL4A1-COL4A2</i>	G (0,44)	[17]
	rs9319428	<i>FLT1</i>	A (0,32)	[19]
14	rs2895811	<i>HHIPL1</i>	C (0,43)	[17]
15	rs3825807	<i>ADAMTS7</i>	A (0,57)	[17, 18, 28]
	rs17514846	<i>FURIN-FES</i>	A (0,44)	[19]
	rs8042271	<i>MFGE8-ABHD2</i>	G (0,90)	[20]
	rs6494488	<i>OAZ2, RBPMS2</i>	G (0,18)	[23]
	rs56062135	<i>SMAD3</i>	C (0,79)	[20]
16	rs1800775	<i>CETP</i>	C (0,51)	[22]
	rs1050362	<i>DHX38</i>	A (0,38)	[23]
17	rs7212798	<i>BCAS3</i>	C (0,15)	[20]
	rs17608766	<i>GOSR2</i>	C (0,14)	[23]
	rs1867624	<i>PECAM1</i>	C (0,39)	[23]
	rs12936587	<i>RAI1-PEMT-RASD1</i>	G (0,56)	[17]
	rs216172	<i>SMG6</i>	C (0,37)	[17]
	rs46522	<i>UBE2Z</i>	T (0,53)	[17]
18	rs663129	<i>PMAIP1-MC4R</i>	A (0,26)	[20]
19	rs2075650	<i>APOE-APOC1</i>	G (0,14)	[26]

Tabelle 1.1 (Fortsetzung): Mit KHK assoziierte GWAS-Genloci ($p < 5 \cdot 10^{-8}$) sortiert nach Chromosom. Zusätzlich aufgeführt sind die assoziierten Gen(e), das Risikoallel mit Häufigkeit sowie die zugehörige Quellenangabe.

Chromosom	SNP	Gen(e) in Region	Risikoallel (Häufigkeit)	Quelle
19	rs1122608	<i>LDLR</i>	G (0,75)	[17, 29]
	rs12976411	<i>ZNF507-LOC400684</i>	T (0,09)	[20]
20	rs867186	<i>PROCR</i>	G (0,11)	[23]
21	rs9982601	<i>KCNE2</i>	T (0,15)	[25]
22	rs180803	<i>POM121L9P-ADORA2A</i>	G (0,97)	[20]

1.2 Risikogen *ADAMTS7*

Bei einem der bekannten KHK-Risikogene handelt es sich um *ADAMTS7*. Reilley *et al.* [28] gelang es im Jahre 2011 erstmals einen SNP in diesem Gen mittels GWAS als KHK-Risikolokus zu identifizieren. In ihren Analysen verglichen sie 12 393 KHK-Patienten mit 7 383 Kontrollen und konnten eine Variante in *ADAMTS7* (rs1994016, $p = 4,98 \cdot 10^{-13}$) als Risikolokus für KHK ausmachen. In einem weiteren Vergleich von 3 644 KHK-Patienten mit einer Gruppe von 5 783 KHK-Patienten, welche einen frühen MI (σ : Alter < 50 Jahre, φ : Alter < 60 Jahre) erlitten haben, konnten Reilley *et al.* keine signifikante Assoziation des *ADAMTS7*-Lokus mit MI ($p = 0,08$) herstellen. Sie postulieren, dass das Protein *ADAMTS7* eine Rolle bei der Bildung von koronarem Plaque spielen könnte und damit die Assoziation mit KHK zu erklären ist. Dass der *ADAMTS7*-Lokus mit KHK und nicht mit einem frühen MI assoziiert wurde, begründen Reilley *et al.* damit, dass bestimmte genetische Veranlagungen KHK fördern und andere MI [28].

Nachfolgend wurden in zwei weiteren GWAS Verbindungen zwischen KHK und Varianten (rs4380028 ($p = 3,71 \cdot 10^{-9}$), rs3825807 ($p = 1,07 \cdot 10^{-12}$)) innerhalb und in der Nähe von *ADAMTS7* [17, 18] identifiziert. Im *ADAMTS7*-Gen wird die Hauptallel A, des signifikantesten SNPs rs3825807 mit einem um 8 % erhöhten Risiko für KHK assoziiert [30]. Diese aktuellen Studien zeigen basierend auf genetischen Daten eine signifikante Assoziation von *ADAMTS7* mit KHK. Jedoch konnte der Zusammenhang zwischen der Erkrankung und dem Gen nicht aufgeklärt werden.

Anhand von Tierversuchen konnte ein Einblick in die Funktion des proteinkodierenden *ADAMTS7* erlangt werden. In Ratten konnten Wang *et al.* [31] zeigen, dass nach einer gezielten Verletzung der Carotis mit einem Ballonkatheter, welcher einen MI induziert, *Adams7* parallel zur Neointimabildung deutlich überexprimiert ist. Außerdem demonstrieren ihre Versuche, dass *Adams7* mit vaskulären, glatten Muskelzellen (engl. *vascular smooth muscle cells*) in der neu gebildeten Neointima kolokalisiert ist. Die Neointimabildung ist eine aggressive Antwort auf eine Gefäßverletzung und tritt z.B. nach der Implantation eines Stents auf. Im Einklang dazu sind die Ergebnisse von Kessler *et al.* [32]: Ihre Untersuchungen an *Adams7*-Knockout-Mäusen zeigten einen Verlust der Neointimabildung nach einer gezielten Gefäßverletzung. Diesen Ergebnissen nach scheint *ADAMTS7* eine entscheidende Rolle bei der Intimahyperplasie, der Reaktion eines Gefäßes auf ein Trauma, nach einer Gefäßverletzung zu haben. Diese beruht vermutlich auf einer erhöhten Migration von glatten Gefäßmuskelzellen [30]. Der Effekt des Proteins auf die Migration von glatten Muskelzellen könnte aus der Spaltung des *ADAMTS7*-Substrates COMP (Substrate von *ADAMTS7*: siehe Abschnitt 1.3.3) hervorgehen [33]. So zeigten die Untersuchungen von Wang *et al.* [31] außerdem, dass nach der gezielten Verletzung der Gehalt an COMP-Fragmenten und *Adams7* gleichzeitig ansteigt. Da

die Mobilisierung der glatten Muskelzellen ein wichtiger Prozess bei der Plaquebildung und -stabilität ist, implizieren diese Ergebnisse, dass ADAMTS7 eine wichtige Rolle bei der Atherosklerose spielt [30].

Zur Untersuchung der Entwicklung von Atherosklerose züchteten Bauer *et al.* [34] *Adamts7*-Knockout-Mäuse auf den proatherogenen Mausstämmen *Ldlr*^{-/-} und *ApoE*^{-/-}. Sie versorgten die Tiere mit einer hochkalorischen Diät, die bei den *Ldlr*^{-/-}- und *ApoE*^{-/-}-Mäusen normalerweise atherosklerotische Veränderungen induziert. Außerdem entwickeln diese beiden Mausmodelle für gewöhnlich atherosklerotische Läsionen, die einen ähnlichen Phänotyp wie humane Läsionen dieser Art aufweisen. Atherosklerotische Läsionen sind Veränderungen bei der Atherosklerose und beschreiben die Stufen der KHK. Die Plaqueansammlung in einer Arterie ist demnach eine Läsion. Die Versuche von Bauer *et al.* ergaben, dass die Gendeletion von *Adamts7* zu einer reduzierten Bildung dieser atherosklerotischer Läsionen führt. Somit konnten sie *in vivo* zeigen, dass eine direkte genetische Assoziation von *Adamts7* mit Atherosklerose und KHK besteht [34].

Die Ergebnisse aus GWAS und Tierversuchen implizieren somit eine Rolle von ADAMTS7 in der Krankheitsentstehung der Atherosklerose. Aus diesem Schluss folgt der Ansatz, dass ein Funktionsverlust von ADAMTS7, z. B. über eine Inhibition, eine Möglichkeit zum Schutz vor Atherosklerose und KHK darstellen könnte [30]. Alle folgenden Analysen beziehen sich auf die humanen (ADAMTS-)Proteine.

1.3 ADAMTS-Proteine

Das Protein ADAMTS7 gehört zur Familie der ADAMTS-Proteine. Die Gruppe der humanen ADAMTS-Proteine umfasst insgesamt 19 Mitglieder. ADAMTS steht hierbei für *a disintegrin and metalloproteinase with thrombospondin motifs*. Die ADAMTS-Gene bzw. die ADAMTS-Proteine sind von 1 bis 20 durchnummeriert. Mit dem Namen *ADAMTS11* wurde fälschlicherweise ein Gen bezeichnet, welches schon als *ADAMTS5* bekannt war [35]. Daher handelt es sich bei ADAMTS5 um ADAMTS11 und die Bezeichnung ADAMTS11 wird nicht genutzt. ADAMTS-Proteasen haben eine große Bedeutung bei der Ausbildung und Aufrechterhaltung von Gewebestrukturen [36]. Sie spielen eine Schlüsselrolle beim Aufbau des Bindegewebes, Entzündungen, Angiogenese und der Organentwicklung [37, 38].

In Abbildung 1.4 ist ein phylogenetischer Baum für die Proteinfamilie dargestellt. Dieser beruht auf einem Alignment mit den kompletten Sequenzen der Proteine dieser Familie. Er stellt die Verwandtschaft der einzelnen Familienmitglieder zueinander dar. Die ADAMTS-Proteine lassen sich anhand ihrer Sequenz in acht Untergruppen einteilen, was sich auch in ihrer Domänenstruktur widerspiegelt [38].

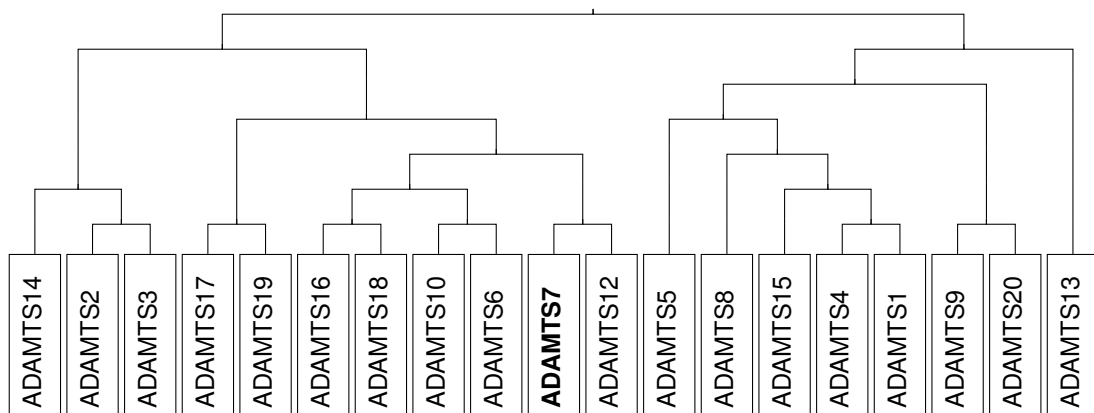


Abbildung 1.4: Phylogenetischer Baum der ADAMTS-Proteine (human, komplette Sequenz, in Anlehnung an Tortorella *et al.* [38]).

1.3.1 Domänenstruktur

Proteine der ADAMTS-Familie gehören unter anderem neben ADAM-Proteinen (engl. *a disintegrin and metalloproteinases*) und MMPs (engl. *matrix metalloproteinases*) zur Gruppe der Metzinkine [39]. Metzinkine bilden unter den Metalloproteinasen eine Superfamilie, die sich durch eine katalytische Domäne auszeichnen. Diese katalytische Domäne, die Metalloproteinase-Domäne, zeigt bei den Metzinkinen einen gemeinsamen Kernaufbau mit einem konservierten Zink-bindenden Motiv und einem sogenannten *Met-turn*, ein Motiv mit einem konservierten Methionin [40]. Die Sequenzfolge des Zink-bindenden Motivs ist HEXGHXXG/N/SXXH, wobei X eine beliebige Aminosäure repräsentiert. Die drei Histidin-Reste in diesem Motiv koordinieren ein katalytisches Zink-Ion in der Bindestelle der Metalloproteinase-Domäne [41]. Acht von 21 ADAM-Proteinen (ADAM2, -7, -11, -18, -22, -23, -29, -32) haben das Zink-Bindemotiv in der Metalloproteinase-Domäne verloren und sind somit proteolytisch inaktiv [42]. Das *Met-turn*-Motiv in C-terminaler Richtung nach dem Zink-Bindemotiv enthält ein stark konserviertes Methionin in der Sequenz V/IMA/S [41]. Die genaue Funktion dieses Motivs ist nicht bekannt.

Strukturell weisen MMP-, ADAM- und ADAMTS-Proteine vor allem in den N-terminalen Domänen Übereinstimmungen auf: Sie beinhalten alle ein Signalpeptid, eine Pro-Domäne und eine Metalloproteinase-Domäne. Die Domänenstrukturen von MMP-, ADAMTS- und ADAM-Proteinen unterscheiden sich wie folgt (siehe Abb. 1.5): Die Proteine der ADAM-Familie haben, genau wie die ADAMTS-Proteine, Disintegrin-ähnliche Domänen und Cystein-reiche Domänen. Weiterhin enthalten die MMP- und ADAM-Proteine C-terminale Domänen, die in ADAMTS-Proteinen nicht vorkommen. Dies sind zum Beispiel Hemopexin-Domänen in MMPs [43] und EGF (engl. *epidermal growth factor*)-ähnliche Domänen und zytosolische Domänen in ADAM-Proteinen [44].

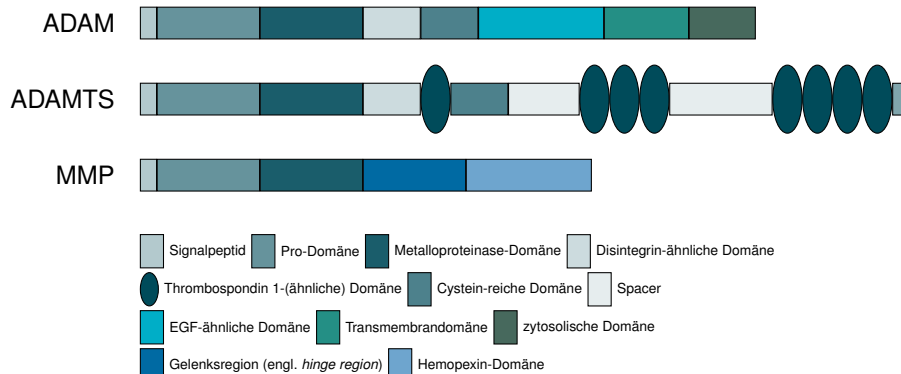


Abbildung 1.5: Schematische Darstellung der Domänenstrukturen von ADAM-, ADAMTS- und MMP-Proteinen im Vergleich (in Anlehnung an Bourboulia *et al.* [43], Liu [45] und Brocker *et al.* [46]).

Der folgende detaillierte Überblick über die ADAMTS-Proteine soll die Gemeinsamkeiten und Unterschiede der Domänenstrukturen dieser Proteinfamilie aufzeigen. In Abbildung 1.6 sind die Domänenstrukturen für alle ADAMTS-Proteine schematisch dargestellt. Die ADAMTS-Proteinstruktur kann in eine Proteinase-Domäne und eine Zusatz-Domäne unterteilt werden. Die Proteinase-Domäne wird bei allen Proteinen dieser Familie in das Signalpeptid, die Pro-Domäne, die Metalloproteinase- sowie die Disintegrin-ähnliche Domäne unterteilt. Die Zusatz-Domäne der ADAMTS-Proteine weist eine höhere Variabilität als die Proteinase-Domäne auf. Anhand der Aufteilung der Zusatz-Domäne kann die Sequenzähnlichkeit wiedererkannt werden (siehe hierzu Abb. 1.4). Sie besteht jeweils aus einer unterschiedlichen Anzahl an Thrombospondin 1-(ähnlichen) Domänen, einer Cystein-reichen Domäne sowie Spacer-Domänen und weiteren

speziellen Domänen in einigen ADAMTS-Proteinen. So haben ADAMTS9 und -20 eine GON-1-Domäne und ADAMTS2, -3 und -14 eine Prokollagen N-Propeptidase-Domäne am C-Terminus, die in ADAMTS7 nicht vorhanden ist. ADAMTS13 hat als einziges Protein dieser Familie C-terminal eine CUB-Domäne (engl. *complement C1r/C1s*). Alle ADAMTS-Proteine, bis auf ADAMTS1, -4, -5, -8, -9, -13, -15 und -20, besitzen eine PLAC-Domäne (engl. *protease and lacunin*) [36].

In allen ADAMTS-Proteinen kommen folgende Domänen vor: Signalpeptid, Pro-, Metalloproteinase-, Disintegrin-ähnliche, Thrombospondin 1-(ähnliche), Cystein-reiche und Spacer-Domäne. Diese Domänen haben jeweils eine spezifische Funktion. Das Signalpeptid legt das Transportziel des Proteins fest und durch die Abspaltung der Pro-Domäne entsteht aktives ADAMTS-Protein. Die Metalloproteinase-Domäne enthält ein konserviertes Zink-Bindemotiv sowie ein *Met-turn*-Motiv. Diese Motive umfassen eine Bindetasche für Inhibitoren [36]. Tortorella *et al.* [38] vermuten, dass die Disintegrin-ähnliche Domäne für diese Proteinfamilie vermutlich falsch benannt wurde. Es ist nicht bekannt, dass ADAMTS-Proteine, im Gegensatz zu ADAM-Proteinen, Integrine über diese Domäne binden. Möglicherweise hat diese Domäne eine Funktion bei der Enzymaktivität und der Substratspezifität. Strukturelle Daten von ADAMTS4 und -5 zeigen, dass es eventuell zwei alternierende Konformationen der Bindestelle in diesem Bereich gibt [47]. Am C-Terminus liegt die Thrombospondin 1-(ähnliche) Domäne, die an Zell-Zell-Interaktionen, Angiogenese und Apoptose beteiligt ist [38]. Außerdem wird postuliert, dass die Cystein-reiche Domäne für die Regulierung der Enzymaktivität verantwortlich sein könnte und die Funktion der Spacer-Domäne in der Substraterkennung und der Proteolyse liegt [38].

Nicht in allen ADAMTS-Proteinen enthalten sind die GON1-, die CUB-, die Prokollagen N-Propeptidase- und die PLAC-Domäne [30, 36, 38, 46]. Die Aufgabe der GON-1-Domäne ist nicht bekannt. Sie kommt nur in ADAMTS9 und -20 vor und enthält zehn konservierte Cysteine [36]. Die CUB-Domäne ist in ADAMTS13 am C-Terminus vorhanden und dient als Bindestelle für das Substrat Von-Willebrand-Faktor [46]. Die Prokollagen N-Propeptidase-Domäne reguliert die Verarbeitung von Prokollagenen und ist in ADAMTS2, -3 und -14 zu finden [38]. Eine weitere Domäne, die nicht in allen ADAMTS-Proteinen vorliegt, ist die PLAC-Domäne. Sie ist in 11 der 19 ADAMTS-Proteine vorhanden, unter anderem auch in ADAMTS7, und liegt immer am C-Terminus mit sechs konservierten Cysteinen vor [36].

1.3.2 Räumliche Gestalt der ADAMTS-Proteine anhand bekannter Kristallstrukturen

In der Protein Data Bank (PDB, www.rcsb.org) [48] sind einige Röntgenkristallstrukturen von humanen ADAMTS-Proteinen enthalten. Von ADAMTS7 selbst gibt es keine bekannte räumliche Struktur. Zum Zeitpunkt der Erstellung der Arbeit (2013-2017) existierte in der PDB außerdem keine komplett aufgelöste Struktur von einem Protein dieser Familie. Bei den vorhandenen ADAMTS-Daten in der PDB handelt es sich um Teilstrukturen. Insgesamt sind 18 Teilstrukturen veröffentlicht, die ADAMTS1, -4, -5 und -13 umfassen.

Die einzelnen ADAMTS-Röntgenkristallstrukturen umfassen jeweils die Metalloproteinase-Domäne, die Metalloproteinase- und die Disintegrin-ähnliche Domäne oder den Bereich der Disintegrin-ähnlichen bis zur Spacer-Domäne. Von ADAMTS1 (z. B. PDB-ID: 2V4B [49]) und ADAMTS4 (z. B. PDB-ID: 2RJP [47], siehe Abb. 1.7a) sind Strukturen, welche die Metalloproteinase- bis zur Disintegrin-ähnlichen Domäne widerspiegeln, bekannt. Die ADAMTS5-Röntgenkristallstrukturen repräsentieren die räumliche Struktur der Metalloproteinase-Domäne (z. B. PDB-ID: 3HYG [50]), wobei eine Struktur (PDB-ID: 2RJQ [47]) auch die Disintegrin-ähnliche Domäne abdeckt. Die Metalloproteinase-Domäne dieser Strukturen besteht aus

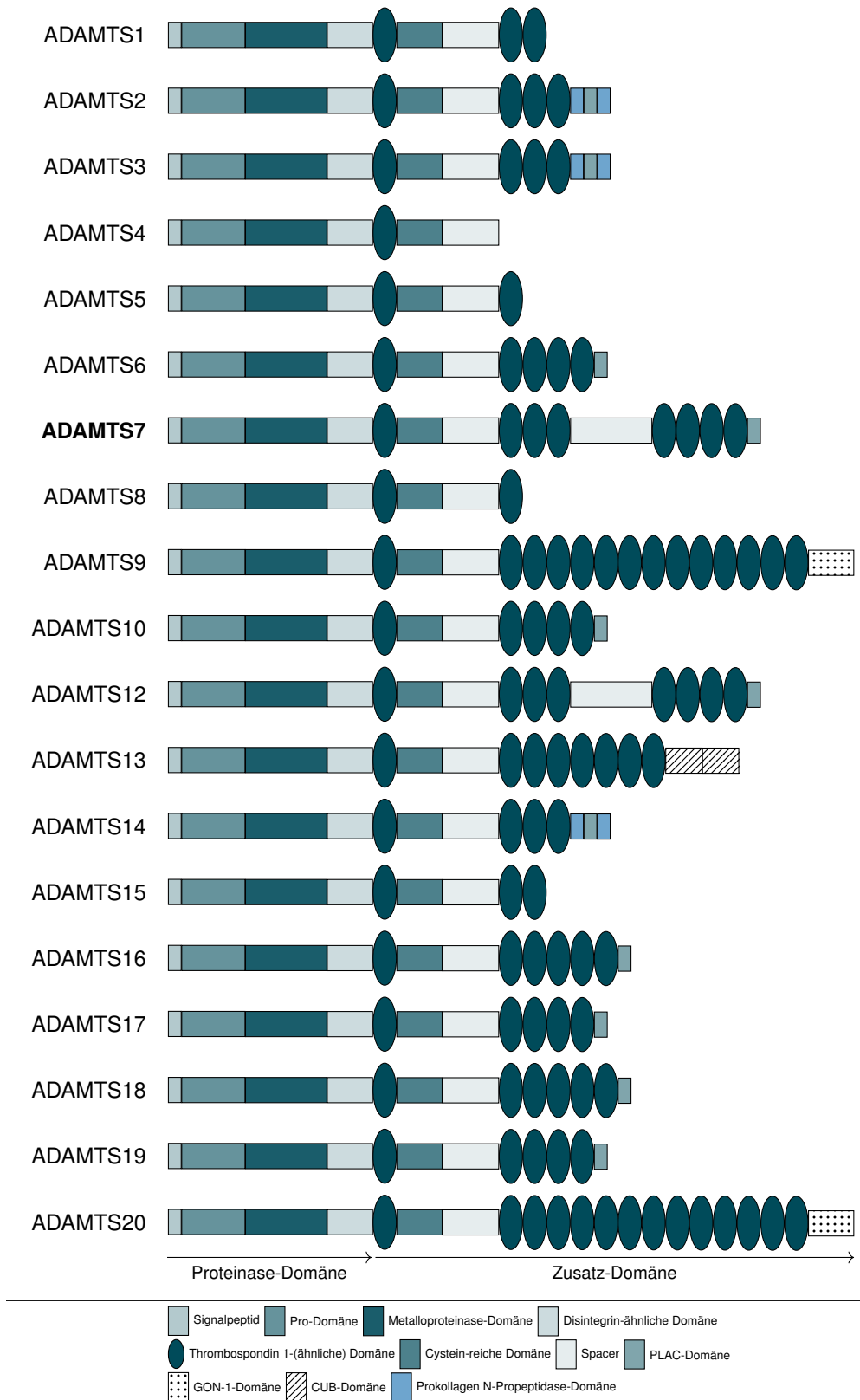


Abbildung 1.6: Schematische Darstellung der Domänen der humanen ADAMTS-Proteine (in Anlehnung an Kelwick *et al.* [36] und Liu [45]). Die ADAMTS-Strukturen können in eine Proteinase-Domäne und eine Zusatz-Domäne unterteilt werden. Die Proteinase-Domäne ist in allen ADAMTS gleich aufgebaut. Die Zusatz-Domäne unterliegt einer höheren Variabilität.

fünf α -Helices und fünf β -Faltblättern, sowie der Bindetasche für das Zink-Ion. Im Vergleich dazu hat die Disintegrin-ähnliche Domäne weniger stabilisierende Sekundärstrukturen (zwei α -Helices und fünf β -Faltblätter), dafür liegen längere Loop-Bereiche vor [47].

Die bekannten räumlichen Strukturen von ADAMTS13 (z. B. PDB-ID: 3GHM [51]) umfassen die Disintegrin-ähnliche bis zur Spacer-Domäne (siehe Abb. 1.7b). Der C-terminale Teil der Cystein-reichen Domäne hat eine ungeordnete Raumstruktur. Wohingegen der N-terminale Teil aus einer α -Helices sowie zwei antiparallelen β -Faltblättern besteht. Dieser ist mit der Spacer-Domäne, welche sich aus zehn β -Faltblättern zusammensetzt, über den C-terminalen Bereich der Cystein-reichen Domäne verbunden. Die Disintegrin-ähnliche Domäne ist über die Thrombospondin 1-Domäne mit den C-terminalen Domänen verknüpft [51].

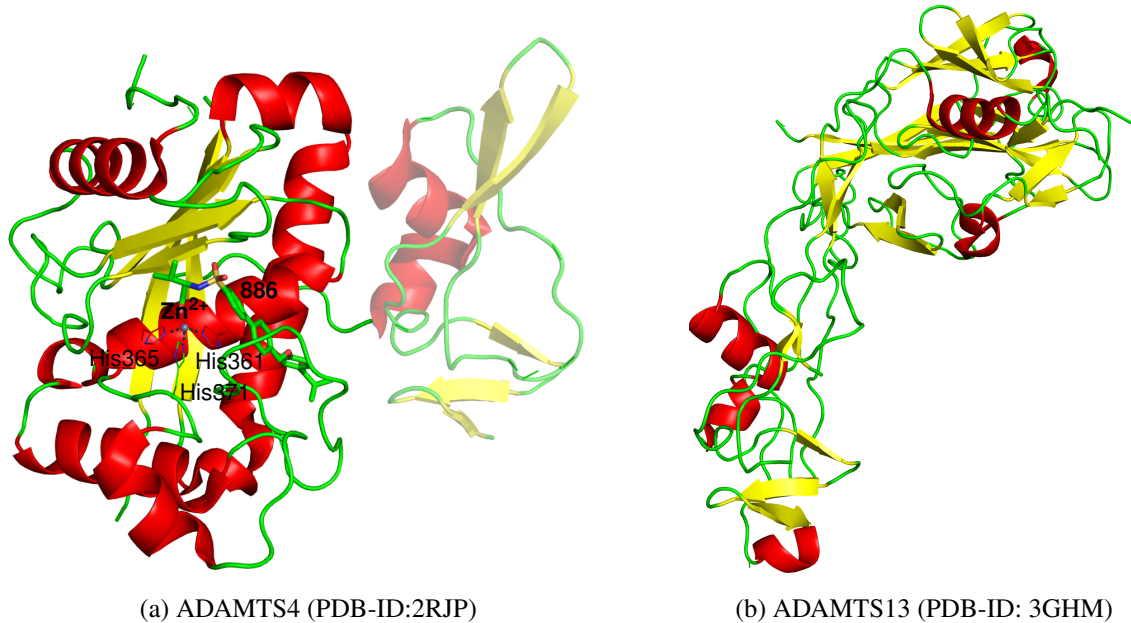


Abbildung 1.7: Röntgenkristallstrukturen von ADAMTS4 und -13. In (a) ist die Metalloproteinase-Domäne von ADAMTS4 samt Bindestelle mit gebundenem Liganden 886 (in Stick-Darstellung) sowie die Disintegrin-ähnliche Domäne (transparent) dargestellt (PDB-ID: 2RJP, Kette A [47]). (b) präsentiert die Disintegrin-ähnliche Domäne bis zu Spacer-Domäne von ADAMTS13 (PDB-ID: 3GHM [51]). Die Proteine sind in Cartoon-Darstellung dargestellt. Sofern nicht anders angegeben, wurden Bilder von Proteinstrukturen mit PyMOL (The PyMOL Molecular Graphics System, Version 1.7.0, Schrödinger, LLC) generiert.

Bei 13 der 15 bekannten ADAMTS-Strukturen, welche die Metalloproteinase-Domäne abdecken, handelt es sich um Protein-Inhibitor-Komplexe. Die Inhibitoren in diesen Strukturen koordinieren alle das bereits erwähnte katalytische Zink-Ion, welches in jeder ADAMTS-Bindetasche gebunden ist. Anhand der ADAMTS5-Röntgenkristallstruktur (PDB-ID: 3HYG [50]), welche mit einem Liganden (099) kokristallisiert ist, wird dies in Abbildung 1.8b beispielhaft verdeutlicht. Das Zink-Bindemotiv umfasst drei Histidine, His410, His414 und His420, die das Zink-Ion koordinieren. Des Weiteren wird das Zink-Ion über eine Zink-bindende Gruppe des Liganden 099 koordiniert. Typische Verbindungen mit Zink-bindenden Gruppen sind Carbonsäuren, Hydroxamsäuren und Thiole [52]. Tortorella *et al.* [38] postulieren, dass die Koordinierung des katalytischen Zink-Ions eine mögliche Strategie für die Inhibition von ADAMTS(7)-Proteinen sein könnte.

Mosyak *et al.* [47] untersuchten Kristalle von ADAMTS4 und -5 (Metalloproteinase-Domäne und Disintegrin-ähnliche Domäne). Sie konnten zeigen, dass die Teilstrukturen dieser beiden Proteine die gleiche Tertiärstruktur (Domänenstruktur und räumlicher Aufbau) aufweisen. Die Bindetaschen beider Proteine

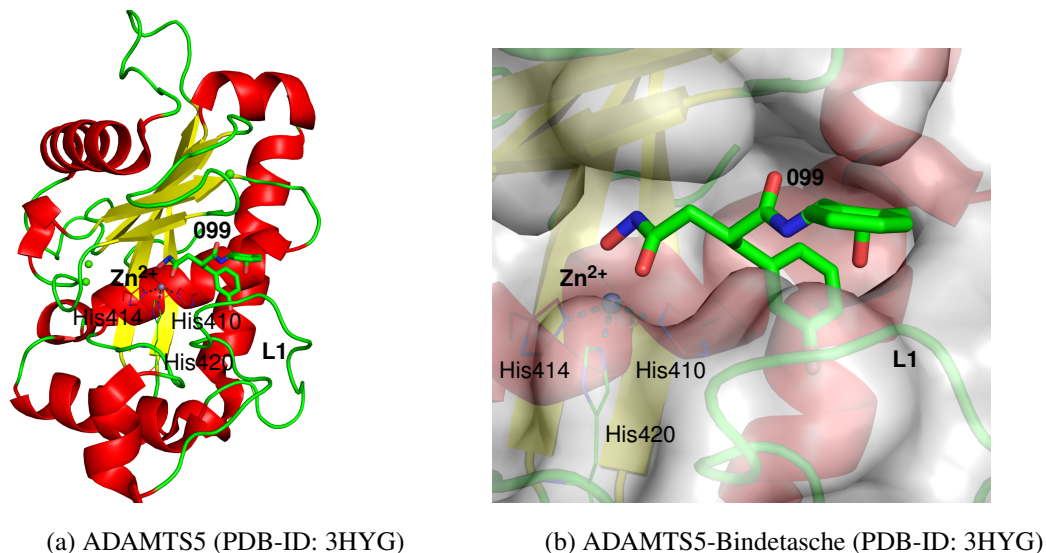


Abbildung 1.8: ADAMTS5-Röntgenkristallstruktur mit Inhibitor 099. In (a) ist die gesamte Metalloproteinase-Domäne (Cartoon-Darstellung) von ADAMTS5 mit gebundenem Liganden (in Stick-Darstellung) dargestellt (PDB-ID: 3HYG, Kette A [50]). (b) zeigt das katalytischen Zink-Ions sowie den Inhibitor 099 in der Bindetasche (Oberflächendarstellung) der Metalloproteinase-Domäne. Das Zink-Ion ist als graue Kugel dargestellt, die Zink-koordinierenden Bindungen als grau gestrichelte Linien.

haben die gleiche Anordnung. Außerdem stellten sie fest, dass die Bindestelle in der Metalloproteinase-Domäne zwei unterschiedliche Konformationen, je nachdem ob ein Ligand gebunden ist oder nicht, annehmen kann. Der L1-Loop (siehe Abb. 1.8b) in der Bindetasche kann sich demnach einem Liganden anpassen.

Die bekannten (Teil-)Strukturen ermöglichen einen Überblick über die räumliche Struktur der Proteine dieser Familie, vor allem im Bereich der Metalloproteinase-Domäne. Anhand dieser können Gemeinsamkeiten im dreidimensionalen Aufbau ausgemacht werden und die Bindung von Inhibitoren in der Proteinbindestasche lokalisiert werden. Das Vorhandensein von ADAMTS-Röntgenkristallstrukturen bietet außerdem eine Grundlage für die Modellierung des Targetproteins ADAMTS7.

1.3.3 Substrate und Inhibitoren der ADAMTS-Proteine und assoziierte Erkrankungen

Alle ADAMTS-Proteine sind proteolytisch aktiv [45]. Für ADAMTS7 und ADAMTS12 ist bekannt, dass sie das Protein COMP (engl. *cartilage oligomeric matrix protein*), welches auch als Thrombospondin-5 bezeichnet wird, spalten können [33, 53]. ADAMTS7 bindet weiterhin α_2 -Makroglobulin [45] und CTGF (engl. *connective tissue growth factor*) [54]. Für die meisten der anderen ADAMTS-Proteine konnten ebenfalls Substrate identifiziert werden, sodass eine weitere Unterteilung der ADAMTS-Proteine anhand der Substrate vorgenommen werden kann: Die Aggrecanasen oder Proteoglykanasen (ADAMTS1, -4, -5, -8, -9, -15, -20) bilden eine Untergruppe neben den Prokollagen N-Propeptidasen (ADAMTS2, -3, -14), den COMP-spaltende Proteinase (ADAMTS7- und -12) und der Von-Willebrand-Faktor-spaltenden Proteinase (ADAMTS13) [35]. Für die sechs Proteine ADAMTS6, -10, -16, -17, -18 und -19 ist aktuell kein Substrat bekannt. Schaut man sich die Substrate in Verbindung mit der Sequenzähnlichkeit der ADAMTS-Proteine an, wird deutlich, dass die ähnlicheren Proteine, welche Untergruppen der ADAMTS-Proteine bilden, die gleichen Substrate binden [36] (siehe Abb. 1.9 und Tabelle 1.2).

Inhibitoren sind für 7 der 19 ADAMTS-Proteine bekannt (siehe Tabelle 1.2). Guo *et al.* [55] konnten *in vitro* und an Zellkulturen aus menschlichen Knorpelzellen zeigen, dass das Protein GEP (engl. *granulin-*

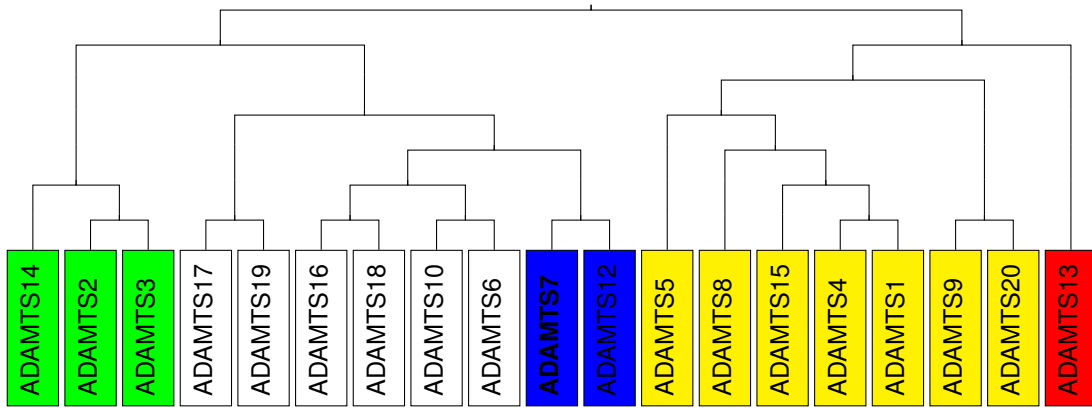


Abbildung 1.9: Phylogenetischer Baum der ADAMTS-Proteine (in Anlehnung an Tortorella *et al.* [38]) sowie deren Unterteilung anhand der bekannten Substrate [36]: Aggrekanasen sind gelb markiert, Prokollagen N-Propeptidasen grün, COMP-spaltende Proteinase blau und Von-Willebrand-Faktor-spaltende Proteinase rot.

epithelin precursor) an ADAMTS7 und -12 bindet und den Abbau von COMP kompetitiv inhibiert sowie inhibitorisch auf die ADAMTS7- und ADAMTS12-Expression wirkt. Sie schlussfolgern, dass GEP an die Substratbindestelle der beiden Proteine und nicht an die Bindetasche in der Metalloproteinase-Domäne bindet. Die Arbeit von Luan *et al.* [56] zeigt, dass α_2 -Makroglobulin nicht nur ein Substrat von ADAMTS7 und -12 ist, sondern ebenfalls einen endogenen Inhibitor dieser beiden Proteine darstellt: α_2 -Makroglobulin inhibiert konzentrationsabhängig die COMP-Spaltung und könnte einen KHK-Therapieansatz darstellen [30, 56]. Die vier TIMP-Proteine (*tissue inhibitors of metalloproteinases*), endogene Inhibitoren der MMP-Familie, binden an die Zusatz-Domäne von vier ADAMTS-Proteinen und wirken auf diese inhibitorisch (ADAMTS1, -2, -4 und -5) [36].

Tabelle 1.2: Humane ADAMTS-Gene/Proteine sowie deren bekannte Substrate und Inhibitoren nach [35, 36, 44, 45]. Zusätzlich sind die Erkrankungen angegeben, welche mit Veränderungen in den *ADAMTS*-Genen assoziiert sind.

Gen/Protein	Bekannte Substrate	Bekannte Inhibitoren	Assoziierte Erkrankungen
ADAMTS1	u. a. Aggrecan, Versikan	TIMP-2, TIMP-3	Krebs
ADAMTS2	Prokollagen Type I-III und V	TIMP-3, Papilin	Dermatosparaxis
ADAMTS3	Prokollagen Typ II, Biglykan		
ADAMTS4	u. a. Aggrecan, Versikan	α_2 -Makroglobulin, TIMP-3, C-terminus von Fibronectin	Arthritis
ADAMTS5	u. a. Aggrecan, Versikan	α_2 -Makroglobulin, TIMP-3	Arthritis
ADAMTS6			
ADAMTS7	α_2 -Makroglobulin, COMP, CTGF	α_2 -Makroglobulin, GEP	KHK KHK
ADAMTS8	Aggrecan		
ADAMTS9	Aggrecan, Versikan		Krebs
ADAMTS10			Weill-Marchesani-Syndrom
ADAMTS12	α_2 -Makroglobulin, Aggrecan, COMP	α_2 -Makroglobulin, GEP	Krebs
ADAMTS13	Von-Willebrand-Faktor	Thrombin	Thrombotisch-thrombozytopenische Purpura
ADAMTS14	Prokollagen Typ I		
ADAMTS15	Aggrecan, Versikan		Krebs
ADAMTS16			Hypertonie
ADAMTS17			Weill-Marchesani-ähnliches Syndrom
ADAMTS18			Mikrokornea-myopische chorioretinale Atrophie-Telekanthus-Syndrom
ADAMTS19			
ADAMTS20	Aggrecan, Versikan		

Mutationen in *ADAMTS*-Genen oder die Dysregulation dieser sind assoziiert mit einer Reihe von Erkrankungen (siehe Tabelle 1.2). Neben der Assoziation von *ADAMTS7* mit KHK, sind *ADAMTS4* und *-5* mit Arthritis; *ADAMTS1*, *-9*, *-12* und *-15* mit Krebs und *ADAMTS16* mit Hypertonie assoziiert [36]. Die Inaktivierung von *ADAMTS2* führt zu Dermatosparaxis, einer Bindegewebserkrankung [36]. Mutationen in *ADAMTS10* bzw. *ADAMTS17* führen zum Weill-Marchesani-Syndrom bzw. zum Weill-Marchesani-ähnlichen Syndrom [36]. *ADAMTS18*-Mutationen verursachen das Mikrokornea-myopische chorioretinale Atrophie-Telekanthus-Syndrom [36]. Rezessive Mutationen in *ADAMTS13* führen zu Thrombotisch-thrombozytopenischen Purpura. Bei dieser Erkrankung geht die Funktionsweise zur Spaltung des Von-Willebrand-Faktors verloren und es kommt zur Blutplättchenaggregation und zum Gefäßverschluss [57].

Die zunehmende Anzahl an Erkrankungen, die mit *ADAMTS*-Genen assoziiert bzw. durch Mutationen in diesen ausgelöst werden, lässt *ADAMTS*-Proteine als attraktives Target für neue Therapienansätze erscheinen [35]. Bis heute gibt es keine Behandlung, die auf spezifischen *ADAMTS*-Inhibitoren, welche auf die Bindestelle bzw. das aktive Zentrum in der Metalloproteinase-Domäne zielen, beruht [35]. Im Allgemeinen ist die Suche nach spezifischen Inhibitoren für Metalloproteasen, zu denen *ADAM*-, *ADAMTS*- und *MMP*-Proteine gehören, eine große Herausforderung, da sich die Bindetaschen dieser Proteine sehr ähnlich sind [35].

1.4 *ADAMTS7*

Das humane *ADAMTS7*-Gen liegt auf Chromosom 15 und hat fünf Transkripte. Eines dieser Transkripte ist proteinkodierend, besteht aus 5 490 Basenpaaren und 24 Exons. Dieses Transkript kodiert für das Protein *ADAMTS7*, welches aus insgesamt 1 686 Aminosäuren besteht. Strukturelle Daten über dieses Protein sind zum Zeitpunkt der Arbeit (2013-2017) nicht bekannt. *ADAMTS7* wird in fast allen humanen Geweben exprimiert. Es konnte unter anderem in Herz, Gehirn, Haut, Lunge, Leber, Skelettmuskulatur, Niere und Bauchspeicheldrüse nachgewiesen werden [58, 59].

ADAMTS7 wurde zuerst von Hurskainen und Kollegen [60] im Jahre 1999 „entdeckt“. Später stellte sich heraus, dass es sich bei dem von ihnen bestimmte Protein um eine verkürzte Form von *ADAMTS7* handelte. Das komplette *ADAMTS7*-Protein wurde 2004 als *ADAMTS7B* veröffentlicht [61] und stellt das eigentliche funktionelle Protein dar. Einige Jahre später konnten Liu *et al.* [33] zeigen, dass *ADAMTS7* das Protein COMP spaltet. COMP ist ein nicht-kollagenes Matrix-Protein des Knorpels bestehend aus 757 Aminosäuren. Das Glykoprotein COMP setzt sich aus fünf identischen Untereinheiten zusammen. Liu *et al.* [33] postulierten weiterhin, dass COMP an den C-terminalen Thrombospondin 1-ähnlichen Domänen von *ADAMTS7* bindet und dass die Angriffsstelle für die COMP-Spaltung die EGF-Domäne von COMP ist. Zusätzlich zeigen ihre Versuchen, dass die *ADAMTS7*-Enzymaktivität pH-abhängig (pH: 7,5 – 9,5) und das Vorhandensein von Zink-Ionen nötig ist. *ADAMTS7* und *ADAMTS12* teilen den identischen Aufbau der Domänenstruktur (siehe Abb. 1.6) und bilden eine Untergruppe der *ADAMTS*-Proteine, aufgrund der gleichen Struktur und des gleichen Substrats, COMP. Im Unterschied zu *ADAMTS7* kann *ADAMTS12* *in vitro* auch Aggrekan spalten [45].

Bai *et al.* [62] untersuchten die proteolytische Aktivität von *ADAMTS7* in Abhängigkeit von der Mutation zweier der drei konservierten Histidine (p.H388L und p.H392G) im Zink-bindenden Motiv der Metalloproteinase-Domäne. Sie konnten nachweisen, dass COMP vom Protein, mit den beiden mutierten Aminosäuren, nicht gespalten wird. Folglich ist dieses konservierte Motiv essenziell für den Umsatz von COMP

durch ADAMTS7, auch wenn das Substrat nicht an diese Domäne bindet. Des Weiteren untersuchten Bai *et al.* [62] die Abhängigkeit der enzymatischen Aktivität bezogen auf die Proteinlänge. Der Verlust der vier C-terminalen Thrombospondin 1-ähnlichen Domänen, die notwendig für die COMP-Bindung sind, führt zu einer drastischen Reduktion der Enzymaktivität (auf etwa 20 %). Das ADAMTS7-Teilprotein bestehend aus Signalpeptid, Pro-Domäne und Metalloproteinase-Domäne zeigt schließlich nur noch eine Aktivität von etwa 10 % im Vergleich zum gesamten ADAMTS7-Protein. Demzufolge ist die Zusatz-Domäne von ADAMTS7 an der Regulation der Enzymaktivität beteiligt und wichtig für diese [62].

Der mit KHK-assoziierte, nicht synonyme SNP rs3825807 in ADAMTS7 resultiert in einer Serin-zu-Prolin-Substitution an der 214. Stelle des Proteins [17]. Diese Mutation liegt in der Pro-Domäne des Proteins, welche bei der Aktivierung des Proteins abgespalten wird. Pu *et al.* [63] postulierten in ihrer Arbeit, dass der nicht synonyme SNP rs3825807 einen Einfluss auf die Proteinreifung von ADAMTS7, die COMP-Spaltung sowie die Migration der glatten Gefäßmuskelzellen hat.

1.5 Zielsetzung

ADAMTS7 ist nach den aktuellen Erkenntnissen ein potentiell Target für die Suche nach neuen therapeutischen Ansätzen in der KHK-Therapie [30]. Experimentelle Versuche an Tieren [34] lassen vermuten, dass der Verlust der Proteinfunktion und somit auch die Inhibition dieses Proteins protektiv gegen Atherosklerose und KHK wirken könnte. Bisher ist weder eine 3D-Struktur von ADAMTS7 bekannt noch sind Wirkstoffe publiziert, welche potenzielle ADAMTS7-Inhibitoren darstellen [30]. Folglich war das Hauptziel der vorliegenden Arbeit die Suche nach und die Identifizierung von potenziell wirksamen ADAMTS7-Inhibitoren.

Ein Ziel war daher die Berechnung einer räumlichen ADAMTS7-Struktur mit computerbasierten Methoden, um die Funktionsweise und die Struktur-Wirkungs-Beziehungen im Target zu untersuchen. Aus diesen Ergebnissen sollten Kriterien für affine und selektive ADAMTS7-Inhibitoren abgeleitet werden. Ist es möglich anhand eines ADAMTS7-Modells, Verbindungen zu finden, die eine hohe Bindeaffinität zu ADAMTS7 aufweisen?

Ein weiteres Ziel bei der Suche nach ADAMTS7-Inhibitoren war es, möglichst selektive Verbindungen zu identifizieren. Mit dieser Bedingung sollte garantiert werden, dass Inhibitoren möglichst spezifisch wirken und bestenfalls keine Nebenwirkungen auslösen. Zu diesem Zweck sollten sequenzielle und strukturelle Untersuchungen der ADAMTS-Familie nicht konservierte bzw. strukturell differenzielle Bereiche aufdecken, die selektive Wechselwirkungen mit möglichen Inhibitoren eingehen könnten. Können Bereiche dieser Art gefunden werden und können, basierend auf den bekannten Bindemustern von ADAMTS-Protein-Ligand-Komplexen, Interaktionen aufgedeckt werden, welche für die Bindung an ADAMTS7 essenziell sein könnten? Diese Erkenntnisse sollten als Filter für die Suche nach Inhibitoren in großen Datenbanken dienen.

So stellte sich weiter die Frage, ob es möglich ist, potenzielle Inhibitoren von ADAMTS7 mithilfe eines Virtual Screenings zu identifizieren. Kann aus den erlangten Ergebnissen ein Pharmacophore für ADAMTS7 erstellt werden, welches essenzielle Eigenschaften von potenziellen Inhibitoren abdeckt und zur Filterung einer großen Datenmenge an Liganden genutzt werden kann? Gelingt es anhand von Filterkriterien die Protein-Liganden-Interaktion adäquat zu bewerten? Lässt sich ein Muster bei den Wechselwirkungen zwischen Bindetasche und potenzielle Inhibitoren im ADAMTS7-Protein erkennen? Im Weiteren sollte die Bindeaffinität von potenziellen ADAMTS7-Inhibitoren in den anderen ADAMTS-Proteinen untersucht werden. Lassen sich mittels dieses Ansatzes Aussagen über die Selektivität dieser Moleküle tätigen?

Zusammengefasst ist das Ziel dieser Arbeit über die Strukturmodellierung von ADAMTS7 einen räumlichen Einblick in die Bindetasche und somit in die Funktion dieses Proteins sowie die Möglichkeiten der Hemmung dieser zu bekommen. Die Analysen sollten mögliche selektive ADAMTS7-Inhibitoren aufdecken, die in experimentellen Untersuchungen überprüft werden sollen. Die erlangten Ergebnisse sollen zukünftige *in vitro*- bzw. *in vivo*-Versuche unterstützen und langfristig bei der Entwicklung einer neuen Therapie für KHK helfen.

2 Methoden

Mittels der computergestützten Analyse von Proteinen und deren Bindetasche(n) ist es möglich, eine Vorhersage über potentielle Liganden dieser zu treffen. Hierbei ermöglichen *in silico*-Untersuchungen z. B. die Abschätzung der Bindeaffinität von einer Vielzahl an Verbindungen zu einem Protein und benötigen im Vergleich zu experimentellen Versuchen wenig Zeit. Bei der Suche nach Liganden für die Arzneimittelentwicklung können *in silico*-Methoden helfen, die Auswahl an Verbindungen, die potenzielle Wirkstoffe sein könnten, einzugrenzen. Mit *in silico*-Methoden sollten in dieser Arbeit potenzielle Inhibitoren für ADAMTS7 identifiziert werden. Ziel dieser Methoden ist es, die *in vitro*- und *in vivo*-Versuche zu unterstützen bzw. deren Aufwand zu verkleinern [6].

Wichtige Methoden dieses Gebiets umfassen unter anderem die Berechnung von Proteinstrukturen sowie die Modellierung von Protein-Protein- bzw. Protein-Ligand-Wechselwirkungen. Die Proteinmodellierung und die Abschätzung der Bindeaffinität von Liganden werden in dieser Arbeit eine übergeordnete Rolle einnehmen. In diesem Kapitel sollen daher die für diese Arbeit wichtigen Methoden erläutert werden. Bevor die Erstellung eines Proteinmodells aus einer Aminosäuresequenz, dessen Validierung sowie die Erstellung verschiedener Konformationen dargestellt wird (3D-Strukturmodellierung, siehe Abschnitt 2.3 sowie Abschnitt 2.4), werden Methoden des Sequenzvergleichs (siehe Abschnitt 2.1) und die Sekundärstrukturvorhersage für Proteine (siehe Abschnitt 2.2) dargelegt. Anschließend werden in Abschnitt 2.5 die Methoden der Molekülmechanik – Kraftfelder (siehe Abschnitt 2.5.1) und Energieminimierung (siehe Abschnitt 2.5.2) – erläutert. Außerdem wird die Moleküldynamik (MD)-Simulation (siehe Abschnitt 2.5.3) beschrieben sowie die Validierungsmöglichkeiten dieser Methode (siehe Abschnitt 2.5.4).

Anschließend wird auf die *in silico*-Methode der Pharmacophoremmodellierung (siehe Abschnitt 2.6) sowie das Docking (Abschnitt 2.7) und das Virtual Screening (siehe Abschnitt 2.8) eingegangen. Diese Methoden dienen der Charakterisierung von Proteinbindetaschen und Liganden sowie der Abschätzung von Bindeaffinitäten.

2.1 Sequenzvergleich

Miteinander verknüpfte Aminosäuren bilden die Aminosäuresequenz eines jeden Proteins und damit die Primärstruktur dieses. Anhand dieser Sequenzinformation können Vergleiche zwischen Proteinen berechnet werden, die als Alignment bezeichnet werden. Anhand eines Alignments wird die Ähnlichkeit von Sequenzen bewertet. Diese Bewertung spielt eine wichtige Rolle, da im Allgemeinen davon ausgegangen wird, dass eine hohe Sequenzidentität oder -ähnlichkeit bei Proteinen eine ähnliche Funktion und/oder Struktur impliziert [64]. So ermöglicht der Sequenzvergleich in bioinformatischen Untersuchungen phylogenetische Analysen, d. h. Aussagen über die Verwandtschaft von Proteinen sowie funktionelle und strukturelle Vorhersagen, basierend auf der Aminosäuresequenz [65].

Zur Bewertung der Sequenzidentität oder -ähnlichkeit mit einem Alignment werden Bewertungsmatrizen (z. B. PAM, BLOSUM) genutzt [66]. Der Austausch von Aminosäuren ist nicht für jede Aminosäure

gleich wahrscheinlich. Mit den Bewertungsmatrizen wird für jede Kombination von Aminosäuren eine Bewertung festgelegt: ein positiver, hoher Wert für identische Aminosäuren oder Aminosäuren mit ähnlichen chemischen oder physikalischen Eigenschaften, welche aufgrund dieser Eigenschaften eine hohe Mutationswahrscheinlichkeit haben. Im Gegensatz dazu werden unähnliche Aminosäuren niedriger oder negativ bewertet. Des Weiteren erfolgt eine Bewertung von Lücken (engl. *gaps*) im Alignment sowie der Länge dieser Lücken [66].

Es existieren globale und lokale Alignmentmethoden, die je nach Anwendung Vor- und Nachteile haben. Mit den lokalen Methoden werden ähnliche Teilsequenzen in den Primärstrukturen gesucht, in den globalen Algorithmen werden Sequenzen einander über die gesamte Länge angepasst. Lokale Methoden haben den Vorteil, dass z. B. bei der Suche nach homologen Proteinen nur in begrenzten Bereichen nach ähnlichen Teilsequenzen gesucht wird. Proteine sind meist aus strukturellen bzw. funktionellen Einheiten (Domänen) aufgebaut, sodass diese Methode das Auffinden solcher Strukturmerkmale erlaubt [67]. Ein Beispiel für eine globale Methode ist der Needleman-Wunsch-Algorithmus [68]. Der Smith-Waterman-Algorithmus [69] stellt eine lokale Methode dar. Beide Alignmentmethoden sind sehr rechenaufwendig [70]. Zur Approximation des Smith-Waterman-Algorithmus wurden die Methoden BLAST [71] und FASTA [72] entwickelt, die einfachere, heuristische Verfahren für lokale Alignments darstellen [70]. Beide Methoden eignen sich für die Suche nach ähnlichen Sequenzen z. B. in Datenbanken.

Der PSI-BLAST-Algorithmus (engl. *Position-Specific Iterated BLAST*) [71] stellt eine Erweiterung des BLAST-Algorithmus dar. Bei diesem Algorithmus handelt es sich um eine multiple Sequenzalignmentmethode. Bei dieser Methode werden z. B. bei der Suche nach homologen Proteinen, iterativ aus Treffern einer BLAST-Suche positionsspezifische Bewertungen anhand der Treffer extrahiert und so ein Profil erstellt. Im nächsten Suchdurchlauf wird dieses Profil gegen die Sequenzdatenbank abgeglichen. So können die Informationen von mehreren Sequenzen bei der Suche genutzt werden, was im Gegensatz zu einer „einfachen“ BLAST-Suche, genauere Ergebnisse nach sich zieht, da durch diese Vorgehensweise eine höhere Empfindlichkeit erreicht wird [71]. In dieser Arbeit wird PROMALS3D [65], ein multiples Sequenzalignmentsprogramm, welches auf PSI-BLAST beruht, verwendet.

2.2 Sekundärstrukturvorhersage

Die Primärstruktur eines Proteins bildet durch räumliche Wechselwirkungen, z. B. Wasserstoffbrückenbindungen, eine Sekundärstruktur aus [73, S. 5]. Diese besteht aus bestimmten Strukturmotiven: den Helices, Faltblättern, Schleifen und “Random-Coil”-Strukturen. Letzteres Motiv wird auch als Loop bezeichnet und stellt eine flexible Schleife dar, die aufgrund fehlender geordneter Wechselwirkungen keine definierte Sekundärstruktur aufweist. Sowohl Helices als auch Faltblätter und Schleifen zeichnen sich durch geordnete Wasserstoffbrücken im Proteinrückgrat aus, welche für die speziellen Geometrien dieser geordneten Strukturen sorgen. Die räumliche Anordnung der Sekundärstrukturen bilden die Tertiärstruktur des Proteins. Ein erster möglicher Schritt in der Vorhersage der Tertiärstruktur einer bekannten Sequenz besteht darin, die Sekundärstruktur des Proteins abzuschätzen [74].

Die Vorhersage der Sekundärstruktur eines Proteins basiert auf bekannten räumlichen Daten von Proteinen sowie dem Wissen über das gehäufte Auftreten von bestimmten Aminosäuren in den jeweiligen Sekundärstrukturmotiven. Im Allgemeinen wird anhand von homologen Proteinen mit bekannter Struktur, die Wahrscheinlichkeit von Sekundärstrukturen über die Sequenz bestimmt [75]. Die Sekundärstrukturvorher-

sage ermöglicht, auch bei geringeren Sequenzidentitäten Strukturhomologien zwischen der Targetsequenz und bekannten 3D-Strukturen zu erkennen. Dabei handelt es sich um eine wichtige Information z. B. bei der Modellierung von Proteinstrukturen, wenn keine homologen Proteinstrukturen vorliegen [74]. Jedoch ist es mit dieser Methode nicht möglich, die räumliche Anordnung der Sekundärstrukturen zu schätzen. Hiermit ist nur eine lokale Bewertung der Interaktionen zwischen den Aminosäuren realisierbar [76].

Es gibt eine Vielzahl an Programmen zur Sekundärstrukturvorhersage. In dieser Arbeit wird genau wie für die Alignments PROMALS3D [65] genutzt, welches für die Vorhersage der Sekundärstrukturen auf PSIPRED [74], einen Maschinelernalgorithmus, welcher künstliche neuronale Netze nutzt, zurückgreift. Die Genauigkeit von Sekundärstrukturvorhersage kann je nach Methode bei bis zu 80 % liegen [77]. Für PSIPRED konnte in zwei Studien eine Genauigkeit von 76,4 % [78] bzw. 81,4 % [79] nachgewiesen werden.

2.3 3D-Proteinstrukturen und deren Modellierung

Wie im vorigen Abschnitt erwähnt, bezeichnet die Tertiärstruktur eines Proteins die räumliche Anordnung der Sekundärstruktur, die durch bestimmte Wechselwirkungen, wie z. B. Disulfidbrücken oder Wasserstoffbrücken stabilisiert wird. Die Tertiärstruktur eines Proteins wird aus einer zusammenhängenden Aminosäurekette gebildet. Die Zusammenlagerung von mehreren (eventuell verschiedenen) Proteinketten wird als Quartärstruktur bezeichnet. Die Kenntnis von räumlichen Proteinstrukturen schafft die Voraussetzung für vielfältige (*in silico*-) Analysen, unter anderem über den Einfluss von Mutationen auf die Struktur oder die Identifikation von Bindetaschen und somit der Entwicklung von Liganden für diese. Strukturelle Daten ermöglichen somit das Verständnis der Proteine, z. B. in Bezug auf deren Funktionsweise [48]. Mit diesem Wissen kann die Proteinfunktion, z. B. durch einen Inhibitor, beeinflusst werden.

Es besteht die Möglichkeit die 3D-Struktur eines Proteins experimentell zu bestimmen. Bei der Röntgenkristallografie kann anhand eines Kristalls des Zielproteins die räumliche Struktur berechnet werden. Mittels NMR-Spektroskopie kann die Proteinstruktur in Lösung bestimmt werden. Ist es möglich, die 3D-Struktur mittels der Kristallstrukturanalyse oder der NMR-Spektroskopie aufzuklären, können diese Daten für *in silico*-Analysen genutzt werden. Hierbei hängt die Güte der Struktur von vielen Faktoren, wie z. B. der Auflösung bei Röntgenkristallstrukturen oder der Abdeckung im Vergleich zur Sequenz des Proteins, ab [80]. NMR- oder Röntgenkristallstrukturen liefern somit ein Bild über die räumliche Gestalt des Proteins. In der Proteindatenbank PDB (www.rcsb.org) [48] sind die dreidimensionale Strukturen von Proteinen, Nukleinsäuren sowie deren Komplexen hinterlegt, welche vor allem mittels der Kristallstrukturanalyse oder NMR-Spektroskopie aufgelöst wurden. Aktuell beinhaltet die PDB etwa 132 500 Proteinstrukturen (Stand: August 2017), wovon der Großteil Röntgenkristall- ($\approx 118\,500$ Strukturen) und NMR-Strukturen ($\approx 12\,000$ Strukturen) sind.

Es besteht jedoch eine große Lücke zwischen der großen Anzahl an bekannten Proteinsequenzen und der geringen Anzahl an bekannten räumlichen Daten dieser [81]. Die experimentelle Bestimmung von Proteinstrukturen unterliegt einigen Einschränkungen. Viele Proteine sind zu groß für eine NMR-Analyse und können nicht für die Röntgenkristallografie kristallisiert werden [64]. Außerdem sind diese Verfahren sehr kosten- und zeitaufwendig. Ist man dennoch an der Struktur eines Proteins interessiert, kann diese unter bestimmten Voraussetzungen mittels *in silico*-Methoden modelliert werden. Diese Methoden basieren auf der Annahme, dass die 3D-Struktur des Proteins durch die Aminosäuresequenz vorgegeben ist und die native räumliche Struktur das Energieminima dieser darstellt [82]. Grundvoraussetzung für die Modellierung

ist somit das Vorliegen der Sequenz des Zielproteins, auch Targetprotein genannt. Es existieren drei verschiedene Ansätze für die Modellierung von Proteinstrukturen: *ab initio*-Strukturvorhersagen, *Threading* und *Homology Modeling*, was auch komparative Modellierung genannt wird.

Die drei Methoden unterscheiden sich grundlegend anhand ihrer Anforderungen. Ein Homology Modeling zur Modellierung eines Zielproteins kann nur vorgenommen werden, wenn homologe Proteine, die eine gemeinsame Abstammung mit dem Target haben, mit bekannter räumlicher Struktur vorliegen. Dies basiert auf der Annahme, dass sequenziell ähnliche Proteine auch einen ähnlichen räumlichen Aufbau besitzen [83]. Gleichzeitig gilt, dass die räumliche Struktur in einer Familie konservierter ist als die Proteinsequenz [84]. Bei dieser Methode wird das 3D-Modell basierend auf einem Alignment zwischen Zielprotein und dem bzw. den sequenziell homologen Protein(en), die Template(s) genannt werden, erstellt [85]. Liegen für eine Proteinsequenz keine homologen, aber strukturell ähnliche Proteine (z. B. bestimmt durch eine Sekundärstrukturvorhersage) vor, kann ein Threading vorgenommen werden. Die Sequenz des Zielproteins wird auf die Rückgrat-atome des Templates aufgefädelt. Diese Methode ist jedoch durch die Konformationen der vorhandenen Strukturen sowie die Unsicherheit durch die geringe Sequenzähnlichkeit limitiert [86, 87]. Die *ab initio*-Strukturvorhersage findet Anwendung, wenn keine homologen oder annähernd ähnliche Proteine mit bekannter 3D-Struktur vorliegen. Nur anhand der Sequenz wird ein Modell des Targetproteins erstellt und das Energieminimum anhand einer Energiefunktion gesucht [84]. Die *ab initio*-Strukturvorhersage ist jedoch die unsicherste und ungenaueste Methode, da bis heute nicht genau geklärt werden konnte, wie nur aus der Proteinsequenz auf die Faltung dieser und somit auf die Proteinstruktur geschlossen werden kann [86].

Das Homology Modeling stellt von den drei Methoden zur Proteinmodellierung, angesichts der zugrunde liegenden experimentelle Daten sowie Annahmen, die zuverlässigste dar [84]. In dieser Arbeit wurde zur Modellierung von Proteinstrukturen die Methode des Homology Modelings genutzt. Daher wird im Folgenden diese Methode genauer erläutert.

2.3.1 Homology Modeling

Die Grundlage für ein Homology Modeling bilden zum Targetprotein homologe Proteine, die eine bekannte 3D-Struktur haben. Diese können in dem Verfahren als Template genutzt werden und sind somit jene Strukturen, auf denen die Modellierung basiert. Für das Homology Modeling muss daher ein Template ausgewählt werden. Dieses wird basierend auf der Sequenzidentität zum Target bestimmt. Zusätzlich können mehrere 3D-Strukturen für unterschiedliche Bereiche des Targetproteins als Template genutzt werden. Dies kann z. B. mit der unterschiedlichen Abdeckung der Proteinsequenz oder strukturellen Parametern begründet sein. Bei einem solchen Modell spricht man von einem zusammengesetzten Modell (engl. *composite model*). Dabei wird empfohlen, nur homologe Proteine als Template zu nutzen, die eine Sequenzidentität von mindestens 30% zum Target haben [86]. Unter diesem Prozentsatz ist ein Alignment unzuverlässig ebenso wie eine Aussage über die gleiche Faltung von Target und Template.

Die Güte des Homologiemodells hängt neben der Sequenzidentität auch von der Template-Struktur ab. Daher werden mögliche Templates (z. B. bei gleicher Sequenzidentität) anhand verschiedener Kriterien bewertet, um das beste Template auszuwählen. Hierbei spielen strukturelle Parameter eine Rolle, wie z. B. die Auflösung, die R- und R_{free} -Werte sowie der B-Faktor. Diese Parameter werden für Röntgenkristallstrukturen erhoben und ermöglichen eine Aussage über deren Qualität. Die Auflösung einer Röntgenkristallstruktur enthält die Information über die Genauigkeit der Position der Atome in der erhaltenen Struktur. Je niedriger die Auflösung des Templates ist desto besser sind die experimentellen Daten der Struktur, was wiederum

auf ein besseres bzw. genaueres Modell schließen lässt. Die R- und R_{free} -Werte geben Aufschluss über die Zuverlässigkeit der Röntgenkristallstruktur im Vergleich zu den experimentellen Daten. Typisch sind Werte um 0,2 für R und 0,26 für R_{free} . Der B-Faktor gibt einen Hinweis auf Ungenauigkeiten in den Atompositionen. So deuten hohe Werte (über 50 \AA^2) auf eine hohe Flexibilität und Ungenauigkeit hin. Der Mittelwert des B-Faktors für eine Struktur wird mit dem durchschnittlichen B-Faktor beschrieben [88]. Die Gesamtheit dieser strukturellen Parameter sowie die Sequenzidentität ermöglichen die Abschätzung der Güte von bekannten Proteinstrukturen im Hinblick auf die Nutzung dieser als (mögliche) Templates beim Homology Modeling.

Es gibt eine Vielzahl an Homology Modeling-Programmen, welche frei verfügbar oder kommerziell erhältlich sind. Für den Nutzer ist dabei oft nicht auf den ersten Blick klar, was diese Programme voneinander unterscheidet. Wichtig ist vor allem, dass zuverlässige Modelle z. B. für Docking-Studien erstellt werden können, wenn keine Röntgenkristall- oder NMR-Daten des Targets vorliegen. Die Studie von Nayeem [89], in der verschiedene Homology Modeling-Programme (u. a. Schrödinger Prime und SWISS-MODEL) verglichen wurden, zeigte, dass die Ergebnisse der getesteten Programme bei einer Sequenzidentität über 40 % zwischen Target und Template alle ähnlich gute Resultate hervorbrachten. Bei einer Identität unter 19 % zeigte nur Schrödinger Prime zufriedenstellende Ergebnisse. Ähnliche Resultate deckte die Studie von Dolan *et al.* [90] auf. YASARAs Homology Modeling Programm zeigte bei CASP8, einem Wettbewerb bei dem die aktuellen Methoden zur Proteinstrukturvorhersage getestet werden, gute Vorhersageergebnisse und realistische Homologiemodelle [91]. Oshiro *et al.* [92] untersuchten, inwieweit sich die Ergebnisse von Docking-Studien zwischen Modellen und Röntgenkristallstrukturen unterscheiden. Dabei konnten sie zeigen, dass mit Homologiemodellen bei einer Sequenzidentität über 50 % zwischen Target und Template im aktiven Zentrum, fünfmal mehr aktive Verbindungen identifiziert werden als durch Zufall. Die Ergebnisse mit Homologiemodellen sind somit vergleichbar mit den erlangten Resultaten mit Röntgenkristallstrukturen und folglich ein guter Ansatz für das Target ADAMTS7 in dieser Arbeit. Aus diesen Studien geht zudem hervor, dass es nicht DAS beste Programm für die Erstellung von Homologiemodellen gibt.

Für diese Arbeit standen drei verschiedene Softwarepakete für das Homology Modeling zur Verfügung. Sie werden im Folgenden jeweils genauer erläutert.

SWISS-MODEL

SWISS-MODEL (<http://swissmodel.expasy.org>) [93] ist ein automatischer Server für das Homology Modeling und gleichzeitig von den drei genutzten Programmen das einzig kostenfreie. Als Startpunkt der Modellierung dient entweder nur die Proteinsequenz des Targets oder das Alignment zwischen Target und Template. Dann wird das Rückgrat und die Seitenketten mit dem Programm PROMOD3 [94] erstellt. Anschließend erfolgt die Erstellung der Loops basierend auf einer Datenbank, mit welcher passende „Templates“ für die Erstellung der Loop-Regionen gefunden werden. Fehlende Seitenketten werden auf Basis des SCWRL4-Algorithmus [95] modelliert. Die Modellqualität wird anhand der Parameter GMQE [96] (Abk. *Global Model Quality Estimation*) und QMEAN [97, 98] (Abk. *Global Model Quality Estimation*) geprüft. Das finale Modell wird mit dem Kraftfeld CHARMM27 [99, 100] energiminimiert.

Es ist mit SWISS-MODEL nicht möglich, dass Target basierend auf verschiedenen Templates zu modellieren.

YASARA

Ein weiteres Programm, welches in dieser Arbeit Anwendung fand, ist YASARA (www.yasara.com) [64, 91]. Innerhalb des Programms kann entweder auf eine PSI-BLAST-Suche [71] nach Template-Strukturen genutzt werden oder ein Template sowie Alignment vorgegeben werden. Die Modellerstellung läuft folgendermaßen ab: Anhand des Alignments wird das Proteinrückgrat modelliert, welches auf den Koordinaten des Templaterückgrats basiert. Für zwischen Target- und Templatesequenz identische Aminosäuren werden die Koordinaten der Seitenketten übernommen. Anschließend werden die fehlenden Seitenketten des Modells mit dem SCWRL-Algorithmus [101] berechnet. Lücken im Alignment werden als Loops modelliert [101]. Die Modellierung der Loops erfolgt wissensbasiert. Hierfür wird in der PDB nach bekannten Loops mit maximaler Sequenzähnlichkeit als Template gesucht. Es folgt die Optimierung des Modells durch eine Energieminimierung zur Korrektur von Fehlern. Hierbei stehen verschiedene Kraftfelder zur Verfügung. Das Modell wird anhand des Z-Scores (WHAT_CHECK) [102] validiert und bei einer negativen Validierung werden vorangegangene Schritte der Modellierung wiederholt [64, 85].

Standardmäßig kann beim YASARA Homology Modeling nicht mehr als ein Template ausgewählt werden. Jedoch können Modelle mit den einzelnen Templates erstellt werden und diese im Nachhinein aligniert und manuell zu einem zusammengesetzten Modell verknüpft werden.

Schrödinger Prime

Einen weiteren Algorithmus zur Proteinstrukturvorhersage bietet die Software Schrödinger Prime [103]. Für das Homology Modeling können Alignments zwischen Template und Target basierend auf Sequenz und Sekundärstrukturinformationen erstellt werden oder vorgegeben werden. Das Prime Homology Modeling beinhaltet die folgenden Schritte: Für das Modell werden die Koordinaten des Proteinrückgrats von den zugeordneten Aminosäuren sowie die Seitenketten von konservierten Aminosäuren kopiert. Die Seitenketten werden optimiert und die nicht konservierten Seitenketten basierend auf der Auswertung eines Datensatzes der PDB abgeschätzt und modelliert. Die nicht konservierten Reste werden zusätzlich energieminimiert. Das dabei genutzte Kraftfeld kann individuell gewählt werden, wobei in dieser Arbeit das Kraftfeld OPLS2005 [104] zur Berechnung der Energie genutzt wurde. Insertionen im Alignment werden *ab initio* modelliert. Loops, welche nicht Template-basiert sind, können außerdem einer Verfeinerung unterzogen werden (Loop-Refinement).

2.3.2 Conformational Sampling

Proteinstrukturen sind (in bestimmten Bereichen) flexibel und können für verschiedene Funktionalitäten unterschiedliche Konformationen einnehmen [105]. Ein Beispiel hierfür sind flexible Regionen in Bindetaschen, die der Anpassung an verschiedene Liganden dienen. Dies kann z. B. in NMR-Studien beobachtet werden, in denen die 3D-Struktur in Lösung untersucht wird und verschiedene Konformationen aufgedeckt werden können [106]. Im Gegensatz dazu stellt ein Modell oder eine Röntgenkristallstruktur nur eine Konformation der Proteinstruktur dar. Die Charakterisierung aller möglichen Konformationen eines Proteins oder der Bindetasche ist wichtig für die Vollständigkeit der folgenden Analysen. Einen Lösungsansatz zur Modellierung der verschiedenen Konformationen bestimmter Bereiche im Protein bietet das *Conformational Sampling*. Daher wird hier diese Methode erläutert. In dieser Arbeit wurden *Conformational Samplings* von Loop-Regionen mit Schrödinger Prime [103] durchgeführt.

Das *Conformational Sampling* mit Prime basiert auf einer *ab initio*-Vorhersage und umfasst vier Schritte: Zuerst wird der Loop modelliert bzw. viele Kandidaten-Konformationen dieses Loops. Dabei wird der Konformationsraum des Loops mithilfe einer Diederwinkel-basierten Prozedur umfassend abgetastet. Nachfolgend wird durch das Clustern dieser Loop-Konformationen eine Auswahl von repräsentierenden Konformationen des Loops ermittelt. Im Anschluss folgt eine Optimierung der Seitenkette sowie eine Energieminimierung der verschiedenen Loop-Konformationen. Die vorhergesagten Loop-Konformationen werden nach ihrer Energie geordnet. Das Ergebnis des *Samplings* sind demnach unterschiedliche Konformationen, welche auch als *Conformational Ensemble* bezeichnet werden.

2.4 3D-Proteinstrukturen und deren Validierung

Um die Güte von (Homologie-)Modellen von Proteinen zu bewerten, gibt es eine Vielzahl an Methoden. Diese haben den Zweck, die Qualität der Modelle zu beurteilen. Dies hilft eventuelle Fehler zu erkennen sowie abzuschätzen, ob es sich bei den Proteinmodellen um native räumliche Strukturen handelt.

Eine mögliche Fehlerquelle beim Homology Modeling sind Kompatibilitätsfehler, die durch ein falsches Alignment entstehen können und auf eine Diskrepanz zwischen Proteinsequenz und -struktur hindeuten. Zur Bestimmung dieser Fehler kann mithilfe von wissensbasierten Energiefunktionen die Kompatibilität zwischen der Aminosäuresequenz und der räumlichen Struktur abgeschätzt werden [107]. Hierbei erfolgt eine Unterscheidung in lokale und globale Kompatibilitätsfehler für die Analyse über die Sequenz und die Proteinstruktur. In dieser Arbeit wird das Programm ProSa2003 [107, 108] und der Meta-Server MetaMQAPII [109] zur Erkennung und Visualisierung dieser möglichen Fehler genutzt.

ProSa2003 gibt einen lokalen Aufschluss über die Qualität des Modells anhand der Darstellung der Energie über die Sequenz. Weiterhin kann mittels des Z-Scores eine globale Abschätzung der Qualität von räumlichen Proteindaten in Abhängigkeit von der Sequenzlänge vorgenommen werden. Die Z-Scores umfassen paarweise Interaktionen/Energien (engl. *pair*), Oberflächenenergien (engl. *surface*) und einen kombinierten (engl. *combined*) Wert [108]. Dies ermöglicht im Vergleich die Abschätzung, ob die Struktur der nativen Faltung der jeweiligen Aminosäuresequenz entsprechen kann.

MetaMQAPII ist ein Meta-Server, welcher ebenso die Qualität von Proteinmodellen abschätzt. Zwei Werte sind dabei wichtig: GDT_TS und RMSD [109]. RMSD ist in diesem Fall nicht die konkrete Differenz zwischen zwei Strukturen, wie sie in Abschnitt 2.5.4 erläutert wird, sondern die Abschätzung der Abweichung des Modells zur nativen Struktur des Proteins in Ångström. Der GDT_TS ist eine globale Abschätzung der Qualität des Modells im Vergleich zur (unbekannten) räumlichen Struktur. Je höher dieser Wert und je niedriger der RMSD, desto besser ist das Modell zu bewerten.

Eine weitere Fehlerquelle in Modellen sind sterische Fehler. Hiermit sind Unstimmigkeiten in der räumlichen Struktur gemeint, die nicht auf dem Alignment beruhen, wie z. B. falsch orientierte Seitenketten oder Loop-Regionen, die nicht der nativen Konformation entsprechen können. Mithilfe eines Ramachandran Plots, wie er im Programm PROCHECK [110] enthalten ist, können solche Fehler identifiziert werden. Der Ramachandran Plot ist eine Darstellung der Dieder- bzw. Torsionswinkel des Proteinrückgrats. Diese Winkel der Aminosäuren kommen nur in bestimmten Kombinationen vor und erlauben daher eine Aussage über die Qualität eines Proteinmodells. So lassen sich Fehler aufdecken, wenn diese Winkel im Modell in nicht erlaubten Bereichen vorkommen. Mithilfe von Energieminimierungen und/oder Moleküldynamik-Simulation

besteht die Möglichkeit sterische Fehler zu korrigieren. Diese Methoden werden in Abschnitt 2.5.2 bzw. Abschnitt 2.5.3 erläutert.

2.5 Molekülmechanik

Ein Ansatz der molekularen Modellierung ist es, ausgehend von den physikalischen Naturgesetzen mit möglichst genauen Rechenverfahren realistische 3D-Molekülstrukturen zu erstellen und zu untersuchen [2, S. 226]. Dabei werden zwei Verfahren unterschieden: quantenmechanische Verfahren und Kraftfeldrechnungen. Quantenmechanische Verfahren berechnen die elektronische Struktur von Molekülen anhand der Schrödingergleichung. Mit diesen Verfahren ist es z. B. möglich chemische Reaktionen zu simulieren. Quantenmechanische Verfahren sind sehr rechenintensiv und wesentlich aufwendiger als Kraftfeldverfahren [86]. Daher werden sie in dieser Arbeit nicht weiter behandelt. Erste Entwicklungen der Kraftfeldverfahren, auch Molekülmechanik genannt, gab es Mitte des 20. Jahrhunderts. Der Grundgedanke war, dass die Berechnung der Van-der-Waals-Wechselwirkungen, Bindungsstreckung und Winkeldeformation ausreichen, um die Struktur und Energie von Molekülen zu beschreiben [2, S. 228]. Die Molekülmechanik gewann erst durch die Verfügbarkeit von Computern an Bedeutung [111, 112].

2.5.1 Kraftfelder

Kraftfelder sind Funktionen, welche die potenzielle Energie von Molekülen beschreiben. Dabei vernachlässigen Methoden, die auf Kraftfeldern basieren, die Elektronenbewegung (im Gegensatz zu quantenmechanischen Verfahren) und die Energie eines Systems wird nur anhand der Position der Atomkerne kalkuliert. Kraftfelder sind demnach Funktionen mit anpassbaren Parametern, welche die kovalenten und nicht-kovalenten Kräfte, die zwischen Atomen wirken, darstellen. Dabei gilt die Annahme, dass Bindungslängen und -winkel in Molekülen Standardwerte annehmen [113].

Die Energie eines Moleküls setzt sich aus mehreren Komponenten zusammen: der Bindungslänge l_i , dem Bindungswinkel θ_i , dem Torsionswinkel ω sowie den nicht-kovalenten Wechselwirkungen [113]. Diese wesentlichen Bestandteile eines Kraftfeldes sind in Abbildung 2.1 dargestellt.

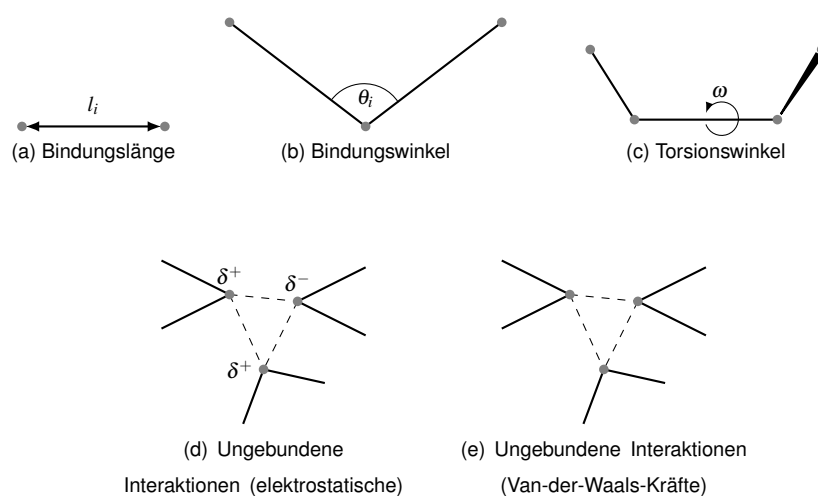


Abbildung 2.1: Schematische Darstellung der auf ein Kraftfeld wirkenden, wesentlichen Bestandteile nach Leach [114, S. 166]: (a) Bindungslänge l_i , (b) Bindungswinkel θ_i , (c) Torsionswinkel ω sowie (d) und (e) ungebundene Interaktionen (elektrostatisch, Van-der-Waals-Kräfte).

Anhand einer beispielhaften Potenzialfunktion eines Kraftfeldes (siehe Formel (2.1)) [112, 113] soll ein Überblick über den generellen Aufbau eines Kraftfeldes gegeben werden. Die potenzielle Energie $E_{pot}(\mathbf{R})$ eines Moleküls mit N Atomen, deren Positionen $\mathbf{r}_i \in \mathbb{R}^3$ mit $1 \leq i \leq N$ und $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$, $\mathbf{R} \in \mathbb{R}^{3 \times N}$ kann folgendermaßen abgebildet werden:

$$\begin{aligned}
 E_{pot}(\mathbf{R}) &= E_{\text{Bindungslänge}} + E_{\text{Bindungswinkel}} + E_{\text{Torsion}} + E_{\text{nicht kovalent}} \\
 &= \sum_{\text{Bindungen}} \frac{k_d}{2} (l_i - l_{i,0})^2 + \sum_{\text{Bindungswinkel}} \frac{k_\theta}{2} (\theta_i - \theta_{i,0})^2 \\
 &\quad + \sum_{\text{Torsionswinkel}} \frac{V_n}{2} (1 + \cos(n\omega - \gamma)) \\
 &\quad + \sum_{i=1}^N \sum_{j=i+1}^N \left(4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right)
 \end{aligned} \tag{2.1}$$

Der erste Term der Formel (2.1) modelliert die Interaktion zwischen Paaren gebundener Atome mittels eines harmonischen Potentials, welches ansteigt, wenn die Bindungslänge l_i von der Referenzlänge $l_{i,0}$ abweicht. Der zweite Term stellt einen Vergleich über alle im Molekül vorkommenden Winkel θ_i zwischen drei Atomen, die miteinander verbunden sind, und einem Referenzparameter $\theta_{i,0}$ dar. Im dritten Term wird die Veränderung der Energie modelliert, wenn um eine Bindung gedreht wird, wobei ω den Torsionswinkel darstellt. Der vierte Term in Formel (2.1) stellt die Energie der ungebundenen Interaktionen dar. Sie wird zwischen allen Atompaaren, die in unterschiedlichen Molekülen oder im gleichen Molekül durch mindestens drei andere Bindungen getrennt sind, modelliert. Gewöhnlich wird dieser Term durch ein Lennard-Jones-Potential für Van-der-Waals-Interaktionen (erster Summand im vierten Term) und ein Coulomb-Potential für elektrostatische Interaktionen (zweiter Summand im vierten Term) dargestellt. r_{ij} beschreibt den Abstand zwischen dem i -ten und dem j -ten Atomen, q_i Ladung und ϵ_0 die Dielektrizitätskonstante des Vakuums. Bei $k_d, k_\theta, l_{i,0}, \theta_{i,0}, V_n, \epsilon_{ij}$ und σ_{ij} in Formel (2.1) handelt es sich um Parameter.

Die erläuterten Komponenten der Energiefunktion spiegeln die intra- und intermolekularen Kräfte eines Moleküls wider. Alle in *in silico*-Analysen genutzten Kraftfelder bestehen aus mindestens diesen vier Termen von (2.1). Zusätzlich können sie noch weitere Terme, welche z. B. die atomare Polarisierbarkeit berücksichtigen, beinhalten [113]. In dieser Arbeit wird beim Homology Modeling mit YASARA das Kraftfeld YASARA2 [91] genutzt, mit SWISS-MODEL das Kraftfeld CHARMM27 [99, 100]. Für alle Methoden mit der Schrödinger Software wurde das Kraftfeld OPLS2005 [104] verwendet.

2.5.2 Energieminimierung

Eine Proteinstruktur kann mehrere lokale Energieminima annehmen, jedoch existiert nur ein globales Energieminimum für jede Struktur. Modelle eines Proteins stellen nicht automatisch die energetisch günstigste Konformation dieser Struktur dar. Jedoch wird bei der Proteinmodellierung nach jener Struktur mit der geringsten potenziellen Energie gesucht. Dies beruht auf der Annahme, dass diese den nativen Zustand des Proteins darstellt. Mit einer Energieminimierung kann das Proteinmodell optimiert werden. Dies geschieht über eine iterative Anpassung der Atompositionen, wodurch sterische Fehler im Modell reduziert werden. Dadurch wird die Struktur verfeinert, was eine Minimierung der potenziellen Energie nach sich zieht. Die Methode der Energieminimierung kann eine optimierte Struktur ermitteln, welche aber nicht zwangsläufig das globale Minimum darstellt [115].

In dieser Arbeit wurden Minimierungsalgorithmen von YASARA, SWISS-MODEL und Schrödinger verwendet. Diese sollen im Folgenden kurz erläutert werden. Das Ergebnis der Energieminimierung ist bei allen Algorithmen maßgeblich abhängig von der Startstruktur, da es sich um lokale Optimierungsmethoden handelt [2, S. 229].

Mit dem Minimierungsalgorithmus von YASARA [116] wird die 3D-Struktur einer simulierten Abkühlung (engl. *simulated annealing minimization*), einem heuristischen Minimierungsverfahren, unterzogen. Der Minimierungsprozess startet mit dem Verfahren des steilsten Abstiegs (engl. *steepest decent*), eine grobe Näherungsmethode, welches schnell in die Nähe eines Energieminimums E des Modells führt. Anschließend erfolgt die simulierte Abkühlung, bei der die Simulationsbox schrittweise auf 0 K herunter gekühlt wird, wobei die Atomgeschwindigkeiten iterativ mit dem Faktor 0,9 skaliert werden. Dieser Schritt wird bis zur Konvergenz von E_{pot} fortgeführt [117].

Die Minimierung der Homologiemodelle im SWISS-MODEL-Verfahren geschieht iterativ [118]. Hierbei wird in jedem Schritt, genau wie bei YASARA, das Verfahren des steilsten Abstiegs genutzt, gefolgt von einem konjugierten Gradientenverfahren (engl. *conjugate gradient*). Beim Verfahren des steilsten Abstiegs wird iterativ in Richtung des negativen Gradienten der Energiefunktion gegangen bis ein Minimum erreicht wird. Das konjugierte Gradientenverfahren hingegen ist eine effizientere Methode, da in jeden Schritt eine Abstiegsrichtung bestimmt wird, sodass alle alten Richtungen keine Verbesserung mehr bringen. Die Optimierung der Struktur wird fortgeführt bis keine stereochemischen Fehler mehr vorliegen.

Die Minimierungen mit der Software Schrödinger wurden mit MacroModel durchgeführt. Als Minimierungsmethode wurde die *Polak-Ribiere Conjugate Gradient* (PRCG)-Methode genutzt (Abbruchkriterium: Norm des Gradienten $\leq 0,05$).

2.5.3 Moleküldynamik-Simulation

Proteinstrukturen, unabhängig davon ob es Modelle oder experimentell erlangte Daten sind, spiegeln eine starre Konformation wider. MD-Simulationen erlauben es die Dynamik dieser Strukturen zu untersuchen. Diese rechnergestützte Methode dient der Abschätzung von Molekülbewegungen und ermöglicht das Verhalten von Molekülen oder größeren Systemen unter verschiedenen Bedingungen nachzuahmen. MD-Simulationen liefern somit Informationen über die zeitliche Abhängigkeit von Eigenschaften des Systems und über strukturelle Veränderungen der Konformation [119]. Dabei werden die Bewegungen der Atome bzw. Moleküle unter Einwirken eines Kraftfeldes verfolgt. Es gilt die Annahme, dass die Wechselwirkungen zwischen den Teilchen den Regeln der klassischen Mechanik, also den Newtonschen Gesetzen, gehorchen [120]. Die Newtonschen Gesetze [114, S. 353] können folgendermaßen spezifiziert werden:

1. Ein Körper bewegt sich geradlinig in eine Richtung mit konstanter Geschwindigkeit, es sei denn eine Kraft wird auf ihn ausgeübt.
2. Kraft ist gleich der zeitliche Änderung des Impulses. Bei konstanter Masse gilt daher:

$$\mathbf{F} = \frac{d}{dt} \mathbf{p} = m \frac{d}{dt} \mathbf{v} = m \mathbf{a}.$$
3. Jede Aktion erzeugt gleichzeitig eine gleich große Reaktion (Gegenkraft), die auf den Verursacher der Aktion zurückwirkt.

Die MD-Simulation des zu analysierenden Moleküls erfolgt, wie bei der Energieminimierung, in einer Simulationsbox, welche mit einer Lösung, z. B. Wassermolekülen, gefüllt ist und unter periodischen Randbedingungen [2, S. 233]. Zur Neutralisierung der Ladung endständiger Aminosäuren werden dem Modell ungeladene Gruppen, sogenannte *cap termini*, angefügt. Die initialen Geschwindigkeiten der Atome in der

Simulation werden anhand einer Boltzmannverteilung bestimmt [2, S. 233]. Die Position \mathbf{r}_i und die Geschwindigkeit eines Atoms $\frac{d\mathbf{r}_i}{dt}$ kann durch die Lösung der gewöhnlichen Differentialgleichung, die sich aus dem zweiten Newtonschen Gesetz ($\mathbf{F} = m\mathbf{a}$) ergibt, berechnet werden:

$$\frac{d^2 r_i^j}{dt^2} = \frac{F(r_i^j)}{m_i}. \quad (2.2)$$

Die Formel (2.2) beschreibt die Bewegung eines Teilchens mit Masse m_i entlang einer Koordinate r_i^j mit $j = (x, y, z)$, wobei $F(r_i^j)$ die Kraft beschreibt, die auf das Teilchen in diese Richtung wirkt. Während der MD-Simulation wird in einem fest definierten Zeitintervall mithilfe der Newtonschen Bewegungsgleichung (2.2) die nächste Position jedes Teilchens berechnet [120]. Die Entwicklung der Bewegung wird über mehrere Nanosekunden verfolgt und in Form einer Trajektorie aufgezeichnet. In dieser Arbeit wurde das Programm Desmond [121] für MD-Simulationen genutzt.

Eine MD-Simulation besteht dabei aus zwei Phasen. Die erste Phase wird auch Äquilibrationsphase genannt und bezeichnet den Zeitraum, in welchem das System beginnend mit der Startkonfiguration in einen Gleichgewichtszustand gebracht wird. Anschließend startet die Produktionsphase [122]. Eine 10 ns lange MD-Simulation kann bereits ausreichend sein, um Bewegungen der Seitenketten oder sogar Proteindomänen zu verfolgen [112].

Proteine sind keine starren Komplexe wie es den Anschein durch Röntgenkristallstrukturen oder die Erstellung eines Modells haben könnte. Anhand von Simulationen kann ein realistisches Bild über die Flexibilität von Proteinen, d. h. ihre native Bewegung bestimmt werden. Zu starke strukturelle Abweichungen vom Startmodell deuten hierbei auf Unstimmigkeiten in diesem hin [123]. Erhebliche Veränderungen in der Proteinstruktur wären ein Hinweis darauf, dass das Modell keiner native Konformation entspricht. Nach einer kurzen Äquilibrationsphase sollte das Protein einen stabilen Zustand annehmen und keine großen strukturellen Veränderungen auftreten. Die Güte einer MD-Simulation wird mittels verschiedener Methoden beurteilt. Eine Möglichkeit zur Bewertung bietet der im Folgenden beschriebene RMSD.

2.5.4 Methoden zum Strukturvergleich: RMSD und RMSF

Zum Vergleich der Veränderung einer Struktur über die Zeit bei einer MD-Simulation oder zweier Docking-Posen eines Liganden gibt es eine globale und eine lokale Vergleichsmethode. Der RMSD stellt dabei die globale, der RMSF die lokale Methode dar [124]. Beide Verfahren werden im Folgenden beschrieben.

Der RMSD steht für die mittlere quadratische Abweichung (engl. *root-mean-square deviation*) und ist eine Methode zum Vergleich zweier Strukturen S und P . Dafür werden die beiden Strukturen übereinander gelegt (engl. *align*) und die Koordinaten ihrer Atome anhand von Formel (2.3) verglichen. Hierbei ist N die Anzahl an Atomen, die sowohl in S als auch in P vorkommen, darstellt. $\mathbf{s}_i = (s_i^x, s_i^y, s_i^z)$ bzw. $\mathbf{p}_i = (p_i^x, p_i^y, p_i^z)$ beinhaltet für $1 \leq i \leq N$ die Koordinaten des jeweils i -ten Atoms der jeweiligen Sequenz. Der RMSD wird in Ångström angegeben, wobei $1 \text{ \AA} = 0,1 \text{ nm} = 10^{-10} \text{ m}$ entspricht.

$$\text{RMSD}(S, P) = \sqrt{\frac{1}{N} \sum_{i=1}^N ((s_i^x - p_i^x)^2 + (s_i^y - p_i^y)^2 + (s_i^z - p_i^z)^2)} \quad (2.3)$$

Zum Vergleich zweier Strukturen kann der RMSD-Wert über die C_α -Atome der Aminosäuren, die als Teil des Proteinerückgrats die Position der Aminosäuren präsentieren, berechnet werden. Für diesen Fall werden in Formel (2.3) nur die Koordinaten der C_α -Atome berücksichtigt.

Von großer Wichtigkeit ist der RMSD bei MD-Simulationen: Wird der RMSD für alle in der Trajektorie gespeicherten Moleküle bezogen auf die Koordinaten des Moleküls zum Zeitpunkt $t = 0$, d. h. der Startstruktur, berechnet, lässt sich die Stabilität bzw. die Schwingung der räumlichen Struktur während der MD-Simulation abschätzen. Je geringer die absolute Differenz der RMSD-Werte zwischen zwei Zeitpunkten t und $t + 1$ ist, desto weniger schwingt das Molekül während der MD-Simulation. Daraus kann geschlossen werden, dass es sich bei der Konformation dieses Moleküls wahrscheinlich um eine native, stabile Konformation handelt. Als stabil wird ein Zeitabschnitt einer MD-Simulation bezeichnet, wenn sich der $\text{RMSD}_{\text{C}\alpha}$ um einen bestimmten Mittelwert in ein Gleichgewicht einpendelt [123]. Dies lässt darauf schließen, dass keine globalen Änderungen in der Konformation geschehen.

Eine weitere Methode zur Beurteilung der Güte der MD-Simulation ist der RMSF. Die mittlere quadratische Schwankung (RMSF, engl. *root-mean-square fluctuation*) ermöglicht einen Vergleich der durchschnittlichen Position der Atome eines Moleküls über die Zeit. Dafür wird die Position von Atomen eines Referenzzeitpunktes (meist $t = 0$) mit der Position der Atome zu bestimmten Zeitpunkten der Simulation verglichen. Der RMSF des Atoms i der Struktur $P(\mathbf{p}_{i_t} = (p_{i_t}^x, p_{i_t}^y, p_{i_t}^z))$ zum (Zeit-)Punkt t wird nach Formel (2.4) berechnet, indem mit der Position des i -ten Atoms zum Referenzzeitpunkt t_{ref} ($\mathbf{p}_{i_{ref}} = (p_{i_{ref}}^x, p_{i_{ref}}^y, p_{i_{ref}}^z)$) gemittelt wird. Dabei sei n , mit $1 \leq t \leq n$, die Anzahl an gemessenen Zeitpunkten.

$$\text{RMSF}_i = \sqrt{\frac{1}{n} \sum_{t=1}^n ((p_{i_t}^x - p_{i_{ref}}^x)^2 + (p_{i_t}^y - p_{i_{ref}}^y)^2 + (p_{i_t}^z - p_{i_{ref}}^z)^2)} \quad (2.4)$$

Wird der durchschnittliche RMSF für die Aminosäurereste eines Moleküls über eine MD-Simulation berechnet, lässt sich eine Aussage über Schwankungen bzw. die Flexibilität in bestimmten Bereichen im Molekül über die Simulationszeit treffen. Je geringer der RMSF eines Aminosäurerestes, desto geringer ist demnach die Schwingung des betrachteten Systems während der MD-Simulation [124]. Hohe Fluktuationen werden in MD-Simulationen vorwiegend in Loop-Regionen beobachtet und sind ein normales Verhalten dieser. Diese Methode ermöglicht somit die Identifizierung von lokalen Veränderungen während der MD-Simulation.

2.6 Pharmacophoremmodellierung

Für die Suche nach Substanzen, die in einer Proteinbindetasche binden könnten, gibt es im Allgemeinen zwei Ansätze: Liegen Informationen über bekannte Liganden und deren Bindemuster vor, kann basierend auf diesen Informationen nach neuen, ähnlichen Liganden gesucht werden. Ist hingegen nur die räumliche Struktur des Proteins oder mindestens die Struktur der Bindetasche vorliegend, kann strukturbasiert nach neuen potenziellen Liganden gesucht werden [5]. Ist nicht bekannt, an welcher Position sich die Bindetasche an der Proteinoberfläche befindet, kann diese computerbasiert identifiziert werden [5].

2.6.1 Definition und Charakterisierung der Proteinbindetasche

Die Bindetasche eines Proteins ist eine Vertiefung in dessen Oberfläche an die ein Ligand binden kann. Bei der strukturbasierten Suche nach Liganden ist das Wissen um die Position der Proteinbindetasche sowie die Charakterisierung der Eigenschaften der Tasche ein wichtiger Schritt. Gleiches gilt für das Finden und Charakterisieren von allosterischen Bindestellen. Allosterische Bindestellen sind Regionen, bei denen es sich nicht um die aktiven Bindetaschen handelt, an die jedoch Liganden binden und die Funktion indirekt (z. B.

durch Veränderung der räumlichen Struktur) inhibieren können. Die Lage von Bindetaschen am Protein kann mittels bestimmter Algorithmen *in silico* aufgedeckt werden [5].

In dieser Arbeit wird für die Definition und die Charakterisierung der ADAMTS-Bindetaschen das Schrödinger-Programm SiteMap [125, 126] genutzt. Bei dieser Methode handelt es sich um einen energiebasierten Ansatz zur Findung von Hohlräumen. Anhand von drei Schritten werden mögliche Bindetaschen identifiziert:

1. Detektion von Höhlen,
2. Charakterisierung der Höhlen und
3. Evaluierung der charakterisierten Höhlen.

Zur Beschreibung der identifizierten Bindestellen werden die Taschen anhand dreidimensionaler, hydrophober und hydrophiler (aufgeteilt in Wasserstoffbrückenakzeptor, -donor und metallbindende Region) Karten dargestellt. Anhand dieser ist es im Anschluss möglich, Eigenschaften potenzieller Liganden abzuleiten.

Es gibt weitere Methoden und Ansätze zur Findung von Proteinbindetaschen, wie z. B. SURFNET [127], ein geometriebasiertes Programm. Auf diese soll hier nicht weiter eingegangen werden, da nur SiteMap in dieser Arbeit genutzt wurde.

2.6.2 Pharmacophore

Ein Pharmacophore beschreibt die räumliche Anordnung funktioneller Gruppen oder Interaktionstypen, welche wesentlich für die biologische Aktivität von Wirkstoffen eines Targets sind [128]. Somit repräsentiert ein Pharmacophore das Interaktionsmuster von bioaktiven Molekülen mit ihrem Target über Interaktionstypen wie z. B. Wasserstoffbrücken oder hydrophobe Wechselwirkungen [129]. Pharmacophore-Modelle werden z. B. für Virtual Screenings (siehe Abschnitt 2.8) oder dem *de novo*-Design von Liganden eingesetzt [130]. Ein Pharmacophore-basiertes Virtual Screening hat das Ziel, aus einer Ligandendatenbank Moleküle herauszufiltern, die wahrscheinlich eine hohe Aktivität (in experimentellen Tests) zeigen [129].

In dieser Arbeit wurde Schrödinger Phase [131] für die Erstellung der Pharmacophore-Modelle genutzt. Es gibt sechs mögliche Pharmacophore-Features, die innerhalb dieser Methode zur Beschreibung der Eigenschaften potenzieller Liganden bestimmt werden können: Wasserstoffbrückendonoren (D) und -akzeptoren (A), aromatischer Ring (R), negativ (N) und positiv (P) ionisierbare (engl. *ionizable*) sowie hydrophobe (H) Features. Das Pharmacophore enthält dazu Informationen über die Distanzen, die Winkel und die Koordinaten der definierten Eigenschaften sowie über das Volumen der Bindetasche.

2.6.3 *e*-Pharmacophore-Methode

Wie oben bereits erwähnt, gibt es zwei Ansätze für die Erstellung von Pharmacophore-Modellen, je nach Informationslage zum Target und Liganden: liganden- und strukturbasierte Pharmacophore-Modelle [132]. Erstere basieren auf bekannten Liganden eines Proteins. Diese werden align und die gemeinsamen Pharmacophore-Punkte identifiziert. Hierbei gilt die Annahme, dass die gemeinsamen Features essenziell für die Bindung zum Target sind [129]. Liegen keine Liganden und keine bekannten Protein-Ligand-Konformationen für ein Targetprotein vor, besteht die Möglichkeit der strukturbasierten Pharmacophoremmodellierung. Es gibt mehrere Möglichkeiten für die Erstellung eines strukturbasierten Pharmacophores. Sind Protein-Ligand-Komplexe veröffentlicht, kann aus diesen das Interaktionsmuster abgeleitet werden. Ist nur die Bindetasche bekannt, kann ein auf den Daten der Bindetasche basierendes Pharmacophore erstellt werden [129].

Das letztere Vorgehen wurde in dieser Arbeit angewendet. Das ADAMTS7-Pharmacophore wurde mit der *e-Pharmacophore*-Methode nach Loving und Salam [133, 134] und Schrödinger Phase [131] erstellt.

Die *e-Pharmacophore*-Methode dient dazu essenzielle Interaktionsbereiche für Liganden in der Bindetasche zu identifizieren. Hierbei wird eine bestimmte Menge an Fragmenten in die Bindetasche des Proteins eingepasst (Docking: siehe Abschnitt 2.7), um die potenziellen Interaktionsbereiche ausfindig zu machen. Die von Schrödinger bereitgestellte Glide Fragmentdatenbank [135], die 667 Fragmente enthält, wird hierfür genutzt. Diese Datenbank enthält Fragmente, welche aus Molekülen aus der medizinischen Chemie abgeleitet sind. Die vorbereiteten Fragmente werden mit Schrödinger Glide [136] in die Bindetasche des Proteins gedockt. Das Docking wird mit *extra precision* (XP) und flexibel gehaltenen Fragmenten durchgeführt. Es werden maximal 100 Posen pro Ligand ausgegeben, welche zusätzlich einer Energieminimierung unterzogen werden. Die daraus resultierenden Docking-Posen werden als Grundlage zur Erstellung des Pharmacophores genutzt.

Mit der *e-Pharmacophore*-Methode wird das Pharmacophore folgendermaßen erstellt: Die gedockten Fragment-Posen bzw. die 2 000 Posen mit den besten Docking-Scores werden in 15 Cluster anhand ihrer Volumenüberlappungen gruppiert. Schrödinger Phase wird anschließend zur Definition der Pharmacophore-Features genutzt. Für jedes Cluster werden Features definiert, die bewertet werden. Die Bewertung basiert auf der Summe der jeweiligen Docking-Scores der Atome, welche zu dem Feature passen. So wird eine Rangfolge der Pharmacophore-Feature basierend auf diesen Energien erstellt. Das Pharmacophore wird mit der gewünschten Anzahl an Features erstellt, wobei die minimale Distanz zwischen unterschiedlichen Features 2 Å und die minimale Distanz zwischen gleichen Features 4 Å beträgt. Weiterhin postulieren Loving *et al.* [134] anhand ihrer Untersuchungen von bekannten Protein-Ligand-Komplexen, dass Pharmacophore mit drei bis sieben Features die angemessensten Hypothesen darstellen.

2.7 Docking und Scoring

Die Bestimmung des Bindungsmodus und der Bindeaffinität zwischen Protein und Liganden bei der Molekülerkennung ist wichtig für das Verständnis der Interaktionsmechanismen und für die Entwicklung von therapeutischen Interventionen [137]. Docking ist eine *in silico*-Methode zur Vorhersage der Bindung eines Liganden an ein Protein sowie der Wechselwirkungen zwischen Ligand und Protein. Diese Methode ermöglicht die Abschätzung der Bindeaffinität sowie der Bindungspose von chemischen Verbindungen in Proteinen bzw. deren Bindetasche sowie die Protein-Ligand-Interaktionen [138]. Eine Voraussetzung für das Docking ist die 3D-Struktur des Proteins bzw. der Bindetasche und des Liganden, ob experimentell bestimmt oder modelliert.

Beim Docking wird der Ligand in die Proteinbindetasche eingepasst. Dazu wird die 3D-Struktur des Liganden, wenn nicht vorhanden, erzeugt und mögliche Konformere generiert, wenn es sich um ein flexibles Molekül handelt. Anschließend wird das Molekül in die Proteinbindetasche eingesetzt. Für den Fall, dass das Molekül aus sterischen Gründen nicht in die Tasche passt, wird das Docking abgebrochen. Standardmäßig werden bei einem Docking eine Vielzahl an Docking-Posen erzeugt. Diese Posen werden anhand ihrer Bindungsgeometrie bewertet, um die Affinität abzuschätzen [137]. Die Evaluation von Docking-Ergebnissen wird anhand von Scoring-Funktionen vorgenommen, welche die Bindungsenergie der Docking-Posen approximiert. Hierbei wird die Protein-Ligand-Interaktion bewertet und diejenige Docking-Pose mit dem bes-

ten Score als bester Treffer oder als Hit bezeichnet [137]. Scoring-Funktionen können in drei Kategorien eingeteilt werden: kraftfeldbasierte, wissensbasierte und empirische Scoring-Funktionen.

Typischerweise sind Protein und Ligand flexibel und können somit ihre Konformation verändern, um aneinander zu binden. In den meisten Docking-Programmen werden die Liganden als flexibel betrachtet, die Proteine hingegen als „starr“. Dies hat den Grund, dass ein flexibles Protein bzw. eine flexible Bindetasche sehr rechen- und damit zeitaufwendig wäre. Eine Möglichkeit ein flexibles Protein zu betrachten ist das *Conformational Ensemble*-Docking. Hierbei werden unterschiedliche Konformationen des Proteins für die Versuche genutzt, die z. B. auf experimentell bestimmten Strukturen oder Modellen (siehe Abschnitt 2.3.2) beruhen [111].

Die verschiedenen Docking-Algorithmen nutzen unterschiedliche Ansätze für das Finden von Protein-Ligand-Posen. Die Auswahl eines Programms zum Docking, welches die besten Ergebnisse für das jeweilige Target liefert, ist nicht trivial. Daher wird zur Testung empfohlen, die Prädiktion der Bindeaffinität bei verwandten Proteinen oder ähnlichen Verbindungen zu vergleichen [138]. Diese Methode zur Evaluation von Docking-Algorithmen wird als Redocking bezeichnet. Unter Nutzung von bekannten Protein-Ligand-Posen, z. B. kokristallisierte Röntgenkristallstrukturen, werden diese Liganden in die Bindetasche des Proteins gedockt. Der strukturelle Vergleich mittels RMSD (siehe Abschnitt 2.5.4) dient dazu, die Ähnlichkeit zwischen gedockter und bekannter Pose zu bestimmen. Je kleiner der RMSD, desto besser wurde die native Konformation des Liganden gefunden [136]. Ziel ist es, diejenige Dockingmethode zu identifizieren, welche die Pose identifiziert, die der nativen Protein-Ligand-Konformation möglichst nahe kommt. Diese Methode kann zum allgemeinen Vergleich der Güte von Docking-Algorithmen sowie zur Feststellung des besten Docking-Algorithmus für ein Protein bzw. eine Proteinklasse für weitere Experimente wie ein Virtual Screening genutzt werden.

In dieser Arbeit standen die folgenden vier Docking-Programme zur Verfügung: AutoDock4 [139], AutoDockVina [140], PLANTS [141] und Schrödinger Glide [136]. Die hier genutzten Docking-Programme, außer AutoDock4, nutzen eine im Vorhinein definierte Bindetasche, welche als Grid bezeichnet wird. Im Folgenden sollen die vier verschiedenen Algorithmen sowie deren Scoring-Funktionen kurz erläutert werden.

2.7.1 Docking mit AutoDock4 und AutoDockVina

AutoDock ist ein Softwaretool, welches die beiden freien Docking-Programme AutoDock4 [139] und AutoDockVina [140] beinhaltet.

AutoDock4 basiert auf einem Lamarck'schen genetischen Algorithmus [142], welcher beim Docking eine Anzahl an Konformationen zwischen Protein und Ligand erzeugt und diese anhand ihrer Bindungsenergie bewertet. In weiteren Schritten werden die Konformationen mit der niedrigsten Energie ausgewählt und iterativ leicht verändert sowie miteinander rekombiniert: Hierbei wird in den einzelnen Konformationen der lokale Konformationsraum nach einem lokalen Minimum abgesucht und diese Information an spätere Generationen weiter geben [139].

AutoDock4 beinhaltet eine kraftfeldbasierte Scoring-Funktion [142], welche auf dem Kraftfeld AMBER beruht [142]:

$$\begin{aligned} \Delta G_{\text{bind}} = & \Delta G_{\text{vdW}} \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + \Delta G_{\text{hbond}} \sum_{i,j} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} + E_{\text{hbond}} \right) + \Delta G_{\text{elec}} \sum_{i,j} \frac{q_i q_j}{\epsilon(r_{ij}) r_{ij}} \\ & + \Delta G_{\text{sol}} \sum_{i_c, j} S_i V_j e^{(-r_{ij}^2/2\sigma^2)} + \Delta G_{\text{tor}} N_{\text{tor}}. \end{aligned} \quad (2.5)$$

Die einzelnen Terme beinhalten die Bewertung der Van-der-Waals-Wechselwirkungen, der Wasserstoffbrückenbindungen, der elektrostatischen Wechselwirkungen, der Desolvation und der Torsionswinkel. Die Summationen erfolgen über die i Ligandatome und die j Proteinatome.

Im Gegensatz zu AutoDock4 basiert die Scoring-Funktion, welche mit AutoDockVina genutzt wird, auf der Kombination einer empirischen und einer wissensbasierten Funktion:

$$g(c_{\text{inter}}) = \frac{c_{\text{inter}}}{1 + \omega N_{\text{rot}}} \quad (2.6)$$

mit $c_{\text{inter}} = w_{\text{gauss1}} f_{\text{gauss1}} + w_{\text{gauss2}} f_{\text{gauss2}} + w_{\text{repulsion}} f_{\text{repulsion}} + w_{\text{hydrophobic}} f_{\text{hydrophobic}} + w_{\text{hbond}} f_{\text{hbond}}$.

Basierend auf den intermolekularen Wechselwirkungen wird die freie Bindeenergie der Docking-Pose abgeschätzt. c_{inter} beschreibt dabei eine Summation über alle Atompaaire, die sich zueinander bewegen können. Wobei w_i die Gewichtung für die verschiedenen Terme und N_{rot} die Anzahl rotierbarer Bindungen im Ligand darstellt [140].

Beim Docking mit AutoDockVina wird ein Quasi-Newton-Verfahren zur lokalen Optimierung verwendet: Der Wert der Scoring-Funktion sowie ihr Gradient werden genutzt, um die Position und Orientierung des Liganden zu bewerten. In mehreren Durchläufen wird anhand einer zufälligen Startkonformation die bestbewertete Protein-Ligand-Konformation ermittelt [140].

Die Docking-Programme AutoDock4 und AutoDockVina unterscheiden sich folglich in der lokalen Suchfunktion sowie der Parametrisierung der Scoring-Funktion zur Bewertung der Docking-Posen. Weiterhin ist der Algorithmus von AutoDockVina im Vergleich zu AutoDock4 zeitoptimiert und zeigt eine höhere Genauigkeit in Docking-Studien. Während bei einem Redocking mit 190 Protein-Ligand-Komplexen mit AutoDockVina 78 % der Liganden mit einem RMSD unter 2 Å im Vergleich zum nativen Komplex gedockt werden konnten, waren es mit AutoDock4 49 % [140].

2.7.2 Docking mit PLANTS

PLANTS steht für *Protein-Ligand ANT System* und stellt einen stochastischen Suchalgorithmus dar [141]. Eine künstliche Ameisenkolonie sucht dabei nach einer minimalen Energiekonformation (*minimum energy conformation*) des Liganden in der Proteinbindetasche [141]. Der PLANTS-Algorithmus stellt zwei verschiedene Scoring-Funktionen bereit: PLANTS_{PLP} (engl. *piecewise linear potential*) und PLANTS_{CHEMPLP}.

Im PLANTS-Suchalgorithmus werden Protein und Ligand als flexible Strukturen behandelt. Es wird ein Minimierungsproblem gelöst, welches die Freiheitsgrade von Ligand (Translation, Rotation und Diederwinkel) und Protein (Diederwinkel) berücksichtigt. Der verwendete Algorithmus ist an das Verhalten von Ameisen(-kolonien) angelehnt, die Pheromone nutzen, um den kürzesten Weg zu einer Nahrungsquelle zu markieren. Innerhalb des Verfahrens wird die Scoring-Funktion minimiert, wobei der Suchraum durch die Freiheitsgrade des Liganden und des Proteins bestimmt wird und die Ameisen ihre Lösungen anhand eines

Simplex-Verfahrens [143] optimieren. In dem Verfahren wird von jeder „Ameise“ ein Wert für jeden Freiheitsgrad gewählt und in jedem Iterationsschritt wird die beste Lösung mit Pheromonen „markiert“ [141].

Im PLANTS-Algorithmus stehen, wie oben erwähnt, die empirischen Scoring-Funktionen $f_{\text{PLANTSPLP}}$ und $f_{\text{PLANTS_CHEMPLP}}$ zur Verfügung. Die PLP-Scoring-Funktion hat die Form:

$$f_{\text{PLP}} = f_{\text{plp}} + f_{\text{tors-lig}} + f_{\text{clash-lig}} + 0,3 \cdot f_{\text{score-prot}} - 20,0. \quad (2.7)$$

Der Term f_{plp} modelliert die sterischen Interaktionen zwischen Protein und Ligand, wobei auch Metall-Ionen in der Proteinbindetasche beachtet werden. Die drei weiteren Summanden bewerten intramolekulare Interaktionen. Ein Kollisionsterm ($f_{\text{clash-lig}}$) verhindert, dass sich die Atome des Liganden zu nahe kommen, und eine Bewertung des Torsionspotenzials wird vorgenommen ($f_{\text{tors-lig}}$). Der Term $f_{\text{score-prot}}$ bewertet die Interaktionen im Protein mit einem Seitenketten-Kollisions-Term. Die Scoring-Funktion CHEMPLP hat folgende Form:

$$f_{\text{CHEMPLP}} = f_{\text{plp}} + f_{\text{chem-hb}} + f_{\text{tors-lig}} + f_{\text{clash-lig}} + 0,3 \cdot f_{\text{score-prot}} - 20,0. \quad (2.8)$$

Der erste Summand der Funktion f_{CHEMPLP} ist der Gleiche wie in der f_{PLP} -Scoring-Funktion, jedoch mit veränderten Parametern. Außerdem beinhaltet diese Scoring-Funktion im Gegensatz zur Funktion in (2.7) den Term $f_{\text{chem-hb}}$, welcher Wasserstoffbrückenbindungen und Metall-Akzeptor-Interaktionen zwischen Protein und Ligand bewertet. Bei den weiteren intramolekularen Termen handelt es sich in beiden Scoring-Funktionen um die gleichen [141].

Im Vorhinein kann nicht abgeschätzt werden, welche der beiden Funktionen die Eigenschaften eines Proteins besser charakterisiert. Daher sollten zur Validierung beide bzw. die zur Verfügung stehenden Scoring-Funktionen hinsichtlich des jeweiligen Targets getestet werden.

2.7.3 Docking mit Schrödinger Glide

Ein weiterer Docking-Algorithmus, welcher in dieser Arbeit Anwendung fand, ist Schrödinger Glide [136]. Dieser stellt drei verschiedene Docking-Präzisionen sowie zwei unterschiedliche Scoring-Funktionen bereit.

Der Glide-Dockingalgorithmus nutzt eine Reihe hierarchischer Filter, um mögliche Bereiche für den Ligand in der Proteinbindetasche zu identifizieren. Es werden initiale Ligandenkonformationen erstellt, die Minima der Liganden (Winkel) darstellen. Diese Ligandenkonformationen werden in der Bindetasche eingepasst, um mögliche Liganden-Posen einzugrenzen. Bezüglich dieser Posen wird der Ligand in der Bindetasche energieminiert. Die so erhaltenen drei bis sechs Posen mit der niedrigsten Energie werden mittels eines Monte Carlo Samplings, in welchem die Torsionswinkel in einer gewissen Umgebung überprüft werden, verbessert. Die erhaltenen Posen werden durch die Bewertung mit einer Scoring-Funktion klassifiziert. Da das Protein in diesem Algorithmus als starr betrachtet wird, werden im letzten Schritt die Van-der-Waals-Radien der Atome reduziert, um den Platz in der Bindetasche zu vergrößern und die Flexibilität des Proteins zu simulieren [136].

Das Programm Glide bietet drei verschiedene Präzisionen: *high throughput virtual screening* (HTVS), *standard precision* (SP) und *extra precision* (XP). HTVS-Docking eignet sich für das schnelle Screening einer großen Anzahl von Liganden, wobei ein eingeschränktes Sampling der Ligandenkonformation stattfindet. SP-Docking ist der Standard beim Glide-Docking. Beim HTVS- und SP-Docking wird der gleiche Algorithmus sowie die gleiche Scoring-Funktion genutzt. Der Unterschied liegt hier im Sampling. XP-

Docking ist (zeit-)aufwendiger als die beiden anderen Verfahren (HTVS, SP). Es wird ein noch umfangreicheres Sampling vorgenommen sowie eine verbesserte Scoring-Funktion zur Bewertung der Docking-Posen genutzt. Ein XP-Docking eignet sich zur Verfeinerung von gut bewerteten SP-Docking-Posen [144, 145].

Die Bindeaffinität wird im Glide-Algorithmus je nach gewählter Präzision mit zwei unterschiedlichen Funktionen geschätzt. Die Präzisionen HTVS und SP haben, wie oben erwähnt, beide die gleiche Scoring-Funktion. Diese hat folgende Form:

$$\begin{aligned} \Delta G_{\text{bind}} = & C_{\text{lipo-lipo}} \sum f(r_{\text{lr}}) + (C_{\text{hb-neut-neut}} + C_{\text{hb-neut-charged}} + C_{\text{hb-charged-charged}}) \sum g(\Delta r)h(\Delta \alpha) \\ & + C_{\text{max-metal-ion}} \sum f(r_{\text{lm}}) + C_{\text{roth}}H_{\text{roth}} + C_{\text{polar-phob}}V_{\text{polar-phob}} \\ & + C_{\text{coul}}E_{\text{coul}} + C_{\text{vdW}}E_{\text{vdW}} + \text{solvation terms.} \end{aligned} \quad (2.9)$$

Der erste Term bewertet Interaktionen zwischen den hydrophoben Atomen von Ligand und Protein und der zweite Term bewertet Wasserstoffbrückenbindungen in der Pose. Anschließend folgen in der Scoring-Funktion (2.9) Bewertungsterme unter anderem für Wechselwirkungen mit Metall-Ionen sowie Van-der-Waals-Interaktionen. Die Scoring-Funktion ist eine Erweiterung der empirischen ChemScore-Funktion von Eldridge *et al.* [146]. Diese Funktion ist adäquat für das Screenen von Datenbanken und reduziert falsch negative Ergebnisse. Die Scoring-Funktion für das Glide XP-Docking setzt sich folgendermaßen zusammen:

$$\begin{aligned} \text{Score} = & E_{\text{coul}} + E_{\text{vdW}} + E_{\text{bind}} + E_{\text{penalty}} \\ \text{mit } E_{\text{bind}} = & E_{\text{hyd_enclosure}} + E_{\text{hb_nn_motif}} + E_{\text{hb_cc_motif}} + E_{\text{PI}} + E_{\text{hb_pair}} + E_{\text{phobic_pair}} \\ \text{und } E_{\text{penalty}} = & E_{\text{desolv}} + E_{\text{ligand_strain}}. \end{aligned} \quad (2.10)$$

Im Vergleich zu (2.9) beinhaltet diese Scoring-Funktion „Bestrafungsterme“ (engl. *penalty*), z. B. für zu geringe Abstände zwischen Atomen ($E_{\text{ligand_strain}}$), und einen Term für π - π -Wechselwirkungen (E_{PI}) [145].

2.7.4 Inverses Screening

Eine weitere Anwendung des Dockings ist die Methode des *Inverse Screenings*, oder auch *Counterscreening* genannt [147]. Die Idee hinter dieser Docking-Anwendung ist, spezielle chemische Verbindungen in eine Gruppe von Proteinstrukturen zu docken. Dies dient unter anderem dazu unbekannte, potenzielle Targets für bekannte Verbindungen aufzudecken oder Liganden innerhalb einer Familie von Proteinen zu identifizieren. In dieser Arbeit spielt das inverse Screening als Methode zur Selektivitätsbestimmung eine Rolle. Diese spezielle Art des Screenings eignet sich zur Abschätzung der Selektivität von potenziellen Hits in einer Proteinfamilie anhand der Docking-Posen. Auf Grundlage des Docking-Scores sowie bestimmter Protein-Ligand-Wechselwirkungen kann postuliert werden, ob Liganden nur an ein bestimmtes Protein binden. In einer Proteinfamilie mit ähnlichen Bindetaschen kann so abgeschätzt werden, ob potenzielle Liganden in homologe Proteine besser oder schlechter binden und es zu Kreuzreaktionen kommen kann. In der Medikamentenentwicklung hilft diese Methode nicht gewollte Nebenwirkungen und Toxizität auszuschließen [147, 148].

Die Bewertung der Docking-Posen erfolgt auch im inversen Screening anhand des Docking-Scores. Die Docking-Scores von Posen einer Verbindung in unterschiedlichen Proteinstrukturen sind nicht notwendig vergleichbar und können einen Bias beinhalten [149]. Liegen experimentelle Daten über die Bindeenergien vor, können anhand dieser die Docking-Scores normalisiert und verglichen werden. Existieren keine experimentellen Daten, kann mittels des durchschnittlichen Docking-Scores für jeden Liganden eine Nor-

malisierung vorgenommen werden [150]. So kann eine Vergleichbarkeit innerhalb des inversen Screenings überprüft und sichergestellt werden. In dieser Arbeit wurde die Normalisierung der Docking-Scores nach dem MACS-Verfahren (engl. *multiple active site correction*) von Vigers und Rizzi [151] durchgeführt. Die MACS-Methode umfasst folgende Berechnungen zur Anpassung des originalen Scores $Score_{ij}$ vom i -ten Liganden in der j -ten Bindetasche:

$$\mu_i = \sum_j Score_{ij} / N \text{ mit } j = 1, \dots, N \quad (2.11)$$

$$\sigma_i^2 = \sum_j (Score_{ij} - \mu_i)^2 / (N - 1) \text{ mit } j = 1, \dots, N \quad (2.12)$$

$$Score'_{ij} = (Score_{ij} - \mu_i) / \sigma_i \quad (2.13)$$

μ_i und σ_i definieren den Mittelwert bzw. die Standardabweichung des Docking-Scores für den i -ten Liganden. $Score'_{ij}$ steht für den normalisierten Docking-Score im inversen Screening.

2.8 Virtual Screening

High-Throughput Screening wird ein experimentelles Verfahren genannt, welches vor allem in der Pharmaforschung genutzt wird und das Auffinden neuer Wirkstoffe zum Ziel hat [5]. Analog zu dieser „wet lab“-Methode gibt es die *in silico*-Methode des Virtual Screenings. Dabei handelt es sich um das computer-gestützte Durchsuchen und Filtern von Datenbanken mit Millionen an Substanzen zur Identifizierung von potenziellen Liganden des Targets [152].

Für das Virtual Screening wird eine selbstgewählte Menge an Liganden genutzt und gefiltert. Hierbei kann z. B. auf interne oder öffentliche Datenbanken mit chemischen Substanzen zurück gegriffen werden. Es stehen mehrere öffentliche Datenbanken, aus denen Verbindungen mit den gewünschten Eigenschaften bezogen werden können, zur Verfügung, so z. B. PubChem (<http://pubchem.ncbi.nlm.nih.gov>), DrugBank (www.drugbank.ca) oder die ZINC-Datenbank (zinc.docking.org) [153]. Letztere wurde in dieser Arbeit genutzt.

Beim Virtual Screening werden im Allgemeinen zwei Ansätze unterschieden: Der ligandenbasierte Ansatz nutzt bekannte Informationen über aktive bzw. inaktive Verbindungen, um daraus eine Struktur-Aktivitäts-Beziehung vorherzusagen, die es ermöglicht, Strukturen mit unbekannter Aktivität in aktive bzw. inaktive Substanzen zu unterteilen [5]. Für diesen Ansatz sind keine Informationen über die Proteinstruktur notwendig. Beim ligandenbasierten Virtual Screening werden lokale sowie globale Methoden unterschieden. Globale Ansätze nutzen z. B. Fingerprints (siehe Abschnitt 2.8.2) und somit die gesamte Struktur der Liganden zur Abschätzung der Ähnlichkeit und der möglichen Bindeaffinität [154]. Lokale Ansätze hingegen, z. B. Screenings mit einem Pharmacophore (siehe Abschnitt 2.6), schätzen die biologische Aktivität anhand bestimmter chemischer Gruppen ab. Im Allgemeinen zeichnet sich der ligandenbasierte Ansatz im Gegensatz zum strukturbasierten Ansatz durch einen geringeren Zeit- und Rechenaufwand aus. Dieser Ansatz ist jedoch nicht geeignet, um neue Strukturmerkmale zu identifizieren. Dies beruht auf der Tatsache, dass Liganden basierend auf der Ähnlichkeit bzw. ähnlicher chemischer Eigenschaften der bereits bekannten Verbindungen gefiltert werden [129].

Der zweite Ansatz wird als strukturbasiertes Virtual Screening bezeichnet. Hierbei wird die Information über die räumliche Struktur des Targets genutzt, um potenzielle Liganden zu identifizieren. Die Grund-

lage bildet eine experimentell bestimmte Struktur oder ein Modell des Targets genutzt werden. Mithilfe eines Dockings werden die Liganden in das Targetprotein eingepasst und die Bindeaffinität (vergleiche Abschnitt 2.7) und somit die Aktivität der Verbindungen abgeschätzt [155]. Ergebnisse eines strukturbasierten Virtual Screenings sind vielfältiger, gleichzeitig ist dieser Ansatz zeitaufwendiger als der ligandenbasierte. Die Vielfältigkeit der Hits ist jedoch maßgeblich abhängig von der genutzten Screening-Datenbank [132]. Je nach vorhandenen Daten zur jeweiligen Fragestellung kann einer dieser beiden Ansätze oder eine Kombination aus beiden Ansätzen realisiert werden.

Des Weiteren ist es möglich, in einem Virtual Screening weitere Vor- und Nachfilterschritte einzuschließen. So kann die Liganden-Datenbank im ersten Schritt mittels verschiedener Methoden, z. B. anhand eines Pharmacophores (siehe Abschnitt 2.6) oder weiterer Kriterien (*Rule of 5* (siehe Abschnitt 2.8.1), bestimmte funktionelle Gruppen etc.), gefiltert werden. Nach dem Docking-Schritt in einem strukturbasierten Virtual Screening können zusätzliche (optionale) Filterungen der Docking-Posen stattfinden. Diese können z. B. auf bekannten Interaktionen zwischen Ligand und Protein basieren, welche essenziell für die Interaktion sind. Die Art, die Anzahl und die Reihenfolge der Filterschritte bei diesem Ansatz kann variieren.

Ziel eines Virtual Screenings ist das Anreichern von aktiven Liganden, sogenannte richtig positive Hits. Sind aktive bzw. inaktive Verbindungen des Targets bekannt, kann die Güte eines Virtual Screenings z. B. mithilfe des Anreicherungsfaktors EF (engl. *enrichment factor*) abgeschätzt werden [156, S. 97]. Der EF bemisst, ob bekannte, aktive Liganden unter den bestbewerteten Hits auftreten und nicht zufällig in der Rangliste des Virtual Screenings verteilt sind. So kann abgeschätzt werden, ob Virtual Screening-Pipeline sowie die Bewertungsmethode, nach welcher die Hits gerankt sind, ein ansprechendes Ergebnis zeigt.

Das Virtual Screening dient folglich dazu Verbindungen mit bestimmten Eigenschaften zu ermitteln, um den Umfang von potenziellen Liganden (Datenbank) um ein Vielfaches zu verkleinern. Der Zweck dieser Methode ist es demnach den Aufwand von experimentellen Screenings zu verringern, da diese *in silico*-Methode in den Aspekten Zeit- und Kostenoptimierung eindeutige Vorteile bietet [132]. In dieser Arbeit wird der strukturbasierte Ansatz des Virtual Screenings genutzt, weil strukturelle Daten des Targets aber keine Informationen über Liganden zur Verfügung standen. Die Ergebnisse solcher Virtual Screenings, die potenziellen Hits, müssen anschließend experimentell validiert werden.

2.8.1 Rule of 5

Die Wahl von Substanzen für ein Virtual Screening unterliegt mehreren Kriterien. Neben der Verfügbarkeit für spätere biologische Testungen muss zudem sicher gestellt sein, dass Verbindungen getestet werden, welche als Wirkstoff dienen könnten. Diese Stoffe müssen vom Organismus aufgenommen werden können. Hier gibt es einige Ansätze, welche die Bioverfügbarkeit von chemischen Verbindungen abschätzen. Einer davon ist die *Rule of 5* von Lipinski [157, 158]. Diese kann als Filterkriterium im Virtual Screening angewendet werden.

Die *Rule of 5* umfasst vier Regeln, welche aus Erfahrungswerten abgeleitet sind und für die Klassifizierung von chemischen Verbindungen genutzt werden können. So haben Moleküle, welche diesen Ansatz erfüllen, folgende Eigenschaften:

1. Nicht mehr als 5 Wasserstoffbrücken-Donorgruppen.
2. Nicht mehr als 10 Wasserstoffbrücken-Akzeptorgruppen.
3. Molekulargewicht ≤ 500 g/mol.
4. Verteilungskoeffizient $\log P \leq 5$.

Die ZINC-Datenbank enthält einen Datensatz an Substanzen, welche die *Rule of 5* erfüllen. Dieser ist als *Drug-Like* gekennzeichnet.

2.8.2 Diversität von Liganden

Ein weiteres mögliches Filterkriterium im Virtual Screening ist die Diversität von Liganden. Um die Ähnlichkeit bzw. Diversität von Liganden zu bestimmen, existieren unterschiedliche Methoden [156, S. 126f.]: Es gibt 2D-Ansätze, welche auf einem Vergleich der Substrukturen basieren, sowie 3D-Methoden. Diese Ansätze bieten die Möglichkeit chemisch ähnliche Verbindungen aufzudecken. Nur ein 3D-Ansatz ermöglicht zusätzlich die Abschätzung der räumlichen Ähnlichkeit.

In dieser Arbeit wurden 2D-Vergleiche genutzt. Hierfür werden für Liganden sogenannte strukturelle Fingerprints [154] generiert. Fingerprints kodieren dabei die chemischen Eigenschaften eines Liganden in einem Bitvektor. Fingerprints wurden in dieser Arbeit mit Schrödinger Canvas [159, 154] erstellt und im MOLPRINT2D-Format [160] genutzt.

Die Ähnlichkeit zwischen Liganden wird anhand der Fingerprints basierend auf Vergleichsmetriken berechnet. In dieser Arbeit wurde der Tanimoto-Koeffizient [154] genutzt:

$$\tau = \frac{N_{AB}}{N_A + N_B - N_{AB}}. \quad (2.14)$$

N_A bzw. N_B ist die Anzahl an Features in Ligand A bzw. B und N_{AB} die Anzahl an Features, welche in beiden Liganden vorkommen. Der Tanimoto-Koeffizient τ nimmt Werte zwischen null und eins an. Je ähnlicher zwei Liganden, desto größer ist dieser Wert.

Der Vergleich von Liganden kann z. B. genutzt werden, um aktive Liganden aus einer Datenbank basierend auf der Ähnlichkeit herauszufiltern oder die Diversität von Datensätzen zu bestimmen.

3 Ergebnisse und Diskussion: Modellierung der ADAMTS7-Proteinstruktur

Bis heute gibt es keine Informationen über Inhibitoren oder Strukturen des ADAMTS7-Proteins. Strukturelle Informationen über ADAMTS7 können dennoch mittels Homology Modeling gewonnen werden. Bereits in Abschnitt 1.3.2 wurden bekannte räumliche Daten von ADAMTS-Proteinen dargestellt. Folglich liegen experimentell bestimmte 3D-Daten von homologen Proteinen vor, welche die Grundlage für die Modellierung von ADAMTS7 bilden. Für die Erstellung des ADAMTS7-Homologiemodells musste dementsprechend ein geeignetes Template (siehe Abschnitt 3.1) ermittelt werden, sodass im Anschluss die Modellierung mit verschiedenen Programmen erfolgen konnte (siehe Abschnitt 3.2). Zur Identifizierung des besten Modells wurde die Qualität anhand mehrerer Validierungskriterien untersucht (siehe Abschnitt 3.3). Um die Flexibilität der Bindetasche zu berücksichtigen, wurden zusätzlich Modellierungsschritte vorgenommen (siehe Abschnitt 3.4).

Die Ergebnisse der Modellierung der räumlichen ADAMTS7-Struktur erlaubten im Anschluss die Identifizierung von möglichen Inhibitoren mittels eines strukturbasierten Ansatzes (siehe Kapitel 5).

3.1 Wahl des/der Template(s)

Zum Zeitpunkt des ADAMTS7-Homology Modelings (Dezember 2013) standen in der PDB [48] insgesamt 15 Teilstrukturen von vier humanen ADAMTS-Proteinen zur Verfügung (siehe Tabelle 3.1). Bei diesen handelt es sich um Röntgenkristallstrukturen der Proteine ADAMTS1, -4, -5 und -13. Es lagen keine NMR-Daten sowie keine komplett aufgelöste ADAMTS-Röntgenkristallstruktur vor. Die bekannten Röntgenkristallstrukturen umfassen die Bereiche der Metalloproteinase-Domäne bis zur Disintegrin-ähnlichen Domäne (ADAMTS1, -4 und -5) bzw. der Disintegrin-ähnlichen bis zur Spacer-Domäne (ADAMTS13). Bis auf jeweils eine Struktur von ADAMTS1 (PDB-ID: 2V4B) und ADAMTS4 (PDB-ID: 3B2Z) sind die Metalloproteinase-Domänen mit einem Inhibitor kokristallisiert. Der Inhibitor befindet sich bei allen Strukturen in räumlicher Nähe zum katalytischen Zink-Ion, welches durch drei Histidin-Reste des Rezeptors und den jeweiligen Inhibitor koordiniert wird.

Zur Berechnung eines ADAMTS7-Homologiemodells wurden die Sequenzen der 15 Röntgenkristallstrukturen mit der Sequenz des Targets mit dem Programm PROMALS3D [65] (Substitutionsmatrix: BLOSUM62) verglichen (siehe Anhang A.1.1, Abb. A.1). Die Identität zwischen den Sequenzen (siehe Tabelle 3.1) wurde mit dem SIAS Webservice [161] bestimmt. Mögliche ADAMTS7-Templates wurden anhand der höchsten Sequenzidentität identifiziert. Die Modellierung des Targets ADAMTS7 hatte zum Ziel, ein Modell mit größtmöglicher Abdeckung der Aminosäuresequenz zu erstellen. Wichtig war in diesem Schritt, dass die Metalloproteinase-Domäne und die darin enthaltene Bindetasche modelliert werden kann. Folglich musste ein Template aus den ADAMTS-Strukturen für das Homology Modeling ausgewählt werden, welches neben sequenziellen und strukturellen Kriterien auch dieses erfüllt.

Tabelle 3.1: Bekannte humane ADAMTS-Röntgenkristallstrukturen zur Zeit des ADAMTS7-Homology Modelings (Dezember 2013). Zusätzlich sind strukturelle Parameter (Domäne, Abdeckung, Auflösung, R , R_{free} und durchschnittlicher D-Faktor) der räumlichen Strukturen angegeben. Des Weiteren ist für jede Röntgenkristallstruktur die PDB-ID, die Domäne, der gebundene Inhibitor sowie die Sequenzidentität zu ADAMTS7 angegeben. ^a - In Röntgenkristallstruktur fehlen Aminosäuren. MPD - Metalloproteinase-Domäne, Disintegrin - Disintegrin-ähnliche Domäne.

Protein	PDB-ID	Domäne	Auflösung (Å)	Abdeckung	Inhibitor	R	R_{free}	durchschnittlicher B-Faktor (Å ²)	Sequenzidentität zu ADAMTS7
ADAMTS1	2JIH [49]	MPD - Disintegrin	2,10	S256-K548 ^a	097	0,224	0,282	35,5	40,8 %
	2V4B [49]	MPD - Disintegrin	2,00	S256-K548 ^a	apo	0,220	0,263	35,1	41,1 %
	3Q2G [163]	MPD - Disintegrin	2,30	S256-K548 ^a	QGF	0,208	0,261	37,7	41,1 %
	3Q2H [163]	MPD - Disintegrin	2,33	S256-K548 ^a	QHF	0,236	0,297	84,9	40,8 %
ADAMTS4	2RJP [47]	MPD - Disintegrin	2,80	A214-H508 ^a	886	0,206	0,267	47,1	41,6 %
	3B2Z [47]	MPD - Disintegrin	2,80	S215-L507 ^a	apo	0,254	0,312	41,9	41,7 %
ADAMTS5	2RJQ [47]	MPD - Disintegrin	2,60	S264-D555	BAT	0,230	0,260	62,8	42,5 %
	3B8Z [164]	MPD	1,40	S264-I480	294	0,185	0,214	18,1	43,3 %
	3HY7 [50]	MPD	1,69	S264-I480	097	0,173	0,224	20,8	43,3 %
	3HY9 [50]	MPD	2,02	S264-I480	098	0,159	0,238	20,5	43,3 %
	3HYG [50]	MPD	1,40	S264-I480	099	0,181	0,210	15,8	43,3 %
ADAMTS13	3LJT [165]	MPD	1,60	S264-I480	LA3	0,159	0,198	10,3	43,1 %
	3GHM [51]	Disintegrin - Spacer	2,60	A299-P682 ^a	-	0,243	0,289	74,6	29,0 %
	3GHN [51]	Disintegrin - Spacer	2,80	A299-P682 ^a	-	0,229	0,280	54,4	28,8 %
	3VN4 [166]	Disintegrin - Spacer	2,80	D298-P682 ^a	-	0,229	0,277	62,4	28,5 %

Die drei ADAMTS13-Röntgenkristallstrukturen weisen jeweils eine Sequenzidentität von unter 30 % zu ADAMTS7 auf. Im Allgemeinen wird empfohlen, nur Strukturen mit einer Sequenzidentität von mindestens 30 % als Template zu nutzen [162], andernfalls kann in den meisten Fällen nicht von dem gleichen strukturellen Aufbau ausgegangen werden. Folglich konnte der Bereich der Thrombospondin 1-Domäne bis zur Spacer-Domäne für ADAMTS7 nicht modelliert werden, weil die ADAMTS13-Röntgenkristallstrukturen nicht als Template genutzt wurden. Die anderen bekannten ADAMTS-Röntgenkristallstrukturen decken diese Domänen nicht ab. Für die weiteren 12 bekannten ADAMTS-Strukturen, die alle auch die Metalloproteinase-Domäne umfassen, ergaben die Analysen jeweils eine Sequenzidentität von über 40 % zu ADAMTS7. Überdies zeigte die Auswertung, dass ADAMTS7 im Vergleich die höchste Übereinstimmung in der Sequenzidentität (zwischen 42,5 % und 43,3 %) zu den sechs ADAMTS5-Röntgenkristallstrukturen aufweist. Die Strukturen von ADAMTS1 und -4 zeigen bis zu 2,5 Prozentpunkte weniger Übereinstimmung auf sequenzieller Ebene. Aufgrund der höchsten sequenziellen Übereinstimmung wurde aus den ADAMTS5-Röntgenkristallstrukturen das passende Template ausgewählt.

Die ADAMTS5-Struktur mit der PDB-ID 2RJQ umfasst als einzige von den sechs ADAMTS5-Strukturen die Disintegrin-ähnliche Domäne und zeigt eine Auflösung von 2,6 Å. Zusätzlich hat 2RJQ unter den bekannten ADAMTS-Röntgenkristallstrukturen, welche die Metalloproteinase- und die Disintegrin-ähnliche Domäne abdecken, die höchste Sequenzidentität zu ADAMTS7. Im Gegensatz dazu beträgt die Auflösung der fünf ADAMTS5-Strukturen, welche nur die Metalloproteinase-Domäne umfassen, zwischen 1,40 Å und 2,02 Å. Die beste Auflösung zeigen die beiden ADAMTS5-Röntgenkristallstrukturen 3B8Z und 3HYG mit 1,40 Å. Von diesen beiden Strukturen weist 3HYG die niedrigeren und damit besseren strukturellen Werte bei den Parametern R, R_{free} und durchschnittlicher B-Faktor (siehe Tabelle 3.1, Erläuterung: siehe Abschnitt sec:HomologyModelling) auf. Die Röntgenkristallstruktur 3HYG enthält zwei Proteinketten, A und B, welche jeweils die gleiche ADAMTS5-Sequenz räumlich darstellen. Anhand von Ramachandran-Plots (PROCHECK [110], Erläuterung: siehe Abschnitt 2.4) konnte gezeigt werden, dass die Kette A des Proteins keine Aminosäuren im nicht erlaubten Bereich aufweist (siehe Tabelle 3.2). In der Kette B hingegen liegen 1,5 % der Aminosäuren in dieser energetisch ungünstigen Region.

Tabelle 3.2: Ergebnisse der Ramachandran Plots (PROCHECK) für die ADAMTS5-Röntgenkristallstrukturen 2RJQ und 3HYG (aufgeschlüsselt nach Kette).

Protein	PDB-ID	Kette	Ramachandran-Plot			
			bevorzugt erlaubter Bereich	zusätzlich erlaubter Bereich	erlaubter Bereich	nicht erlaubter Bereich
ADAMTS5	2RJQ	A	90,7 %	8,5 %	0,4 %	0,4 %
	3HYG	A	89,9 %	10,1 %	0,0 %	0,0 %
	3HYG	B	92,5 %	6,0 %	0,0 %	1,5 %

Neben der Sequenzidentität zu ADAMTS7 und den strukturellen Parametern der Röntgenkristallstrukturen spielt für das Homology Modeling ebenfalls die Sekundärstrukturübereinstimmung zwischen Template und Target eine wichtige Rolle. Für die Sekundärstrukturvorhersage von ADAMTS7 wurde ebenfalls das Programm PROMALS3D [65] verwendet. Es zeigte sich, dass das Targetprotein ADAMTS7 hauptsächlich aus β -Faltblättern und α -Helices aufgebaut ist, welche über Schleifenregionen miteinander verbunden sind (siehe Anhang A.1.1, Abb. A.2). Mittels der Sekundärstrukturvorhersage konnte gezeigt werden, dass sich ADAMTS7 und die beiden Röntgenkristallstrukturen 3HYG und 2RJQ durch eine sehr große Übereinstimmung der Sekundärstrukturen auszeichnen.

Aus den angegebenen Gründen wurde die Röntgenkristallstruktur 2RJQ als Template zum Homology Modeling der Disintegrin-ähnlichen Domäne genutzt. 2RJQ besteht aus einer Proteinkette (vergleiche Tabelle 3.2), sodass diese zur Modellierung genutzt wurde. Für die Modellierung der Metalloproteinase-Domäne wurde die Polypeptidkette A der Röntgenkristallstruktur 3HYG als Template ausgewählt. Ist im Folgenden von 3HYG die Rede, ist immer die Kette A dieser Struktur gemeint. Die Nutzung von zwei (oder mehr) Templates für die Modellierung eines Proteins ist per se möglich, die modellierte Struktur wird dann als zusammengesetztes Modell (siehe Abschnitt 2.3.1) bezeichnet.

Zusammengefasst wurden nach den Vergleichen mit den 15 bekannten ADAMTS-Röntgenkristallstrukturen (Stand: Dezember 2013) zwei als Templates für die Modellierung der ADAMTS7-Struktur ausgewählt: 3HYG zur Modellierung der Metalloproteinase-Domäne und die Disintegrin-ähnliche Domäne von 2RJQ als Template für die korrespondierende Domäne von ADAMTS7. Diese Templates ermöglichten die Erstellung eines ADAMTS7-Modells, was im folgenden Kapitel erläutert wird.

3.2 Modellerstellung

Wie bereits im Abschnitt 2.3.1 erwähnt, gibt es drei verschiedene Methoden (*ab initio*-Strukturvorhersagen, Threading, Homology Modeling) zur Modellierung von räumlichen Proteinstrukturen. Das Homology Modeling stellt beim Vorliegen von homologen Proteinstrukturen für das Targetprotein die sicherste Methode dar [84] und fand daher in dieser Arbeit Anwendung. Das ADAMTS7-Homologiemodell wurde mit den folgenden drei, zur Verfügung stehenden Programmen

- Schrödinger Prime [103],
- SWISS-MODEL [93] und
- YASARA [91]

berechnet. Die Vorgehensweise und die Ergebnisse der Modellerstellung mit diesen Programmen wird im Folgenden beschrieben und diskutiert. Die Grundlage für alle Modellierungen stellte das mit PROMALS3D berechnete Alignment zwischen ADAMTS7 und den Template-Sequenzen (siehe Anhang A.1.1, Abb. A.2) dar. Dies ermöglicht eine Vergleichbarkeit der Ergebnisse.

Sofern es mit dem jeweiligen Programm möglich war, wurden zusammengesetzte Modelle von ADAMTS7 mit beiden gewählten Templates (siehe Abschnitt 3.1) angefertigt. Für den Fall, dass keine zusammengesetzten Modelle erzeugt werden konnten, wurde entweder nur 3HYG oder 2RJQ als Template genutzt. Das Modell mit der Struktur 2RJQ als einziges Template wurde zum Vergleich der strukturellen Parameter der Modelle angefertigt. Es wurde aber davon ausgegangen, dass die Modelle mit 3HYG als Template für die Metalloproteinase-Domäne, aufgrund der besseren Auflösung der Röntgenkristallstruktur, eine bessere Qualität aufweisen würden. Mit dem Programm SWISS-MODEL (vergleiche Abschnitt 2.3.1) war es zum Zeitpunkt des Homology Modelings (Dezember 2013) nicht möglich ein zusammengesetztes Modell zu erstellen. Demzufolge wurden mit diesem Programm zwei ADAMTS7-Modelle, basierend auf dem Sequenzvergleich mit 2RJQ und 3HYG mit Standardparametern berechnet, welche anschließend mit dem Kraftfeld CHARMM27 [99, 100] energieminiert wurden.

Im Gegensatz zu SWISS-MODELL erlaubt das Programm YASARA (vergleiche Abschnitt 2.3.1) Modelle im Anschluss an das Homology Modeling manuell zu verbinden. Aus diesem Grund wurden zwei ADAMTS7-Modelle jeweils basierend auf dem Template 3HYG bzw. 2RJQ mit Standardparametern berechnet. Im Optimierungsschritt wurden beide Modelle mithilfe eines strukturellen Alignments manuell ver-

knüpft. Zur Reduktion sterischer Fehler wurde das zusammengesetzte Modell zusätzlich mit dem Kraftfeld Yasara2 [91] energieminiert. Die Minimierung wurde beendet, wenn sich die Energie in 200 Iterationen um weniger als 0,05 kJ/mol pro Atom verbesserte.

Das dritte genutzte Programm Schrödinger Prime (Version 3.3) (vergleiche Abschnitt 2.3.1) ermöglicht die Berechnung eines zusammengesetzten Modells während der Homology Modeling-Prozedur. Hierbei kann mehr als ein Template gewählt und für die Modellierung der jeweiligen Bereiche des Targetproteins definiert werden. In dieser Arbeit wurde mit Schrödinger Prime ein ADAMTS7-Modell basierend auf den zwei Template-Strukturen modelliert. Es wurden die Standardeinstellungen genutzt, wobei fehlende Aminosäureseitenketten in den Templates im Gegensatz zu den Programmen SWISS-MODEL und YASARA mit einer *ab initio* Prozedur im Targetprotein modelliert wurden. Modellerte Regionen des Targetproteins, welche nicht von den Template-Strukturen abgedeckt werden, wurden einem Loop-Refinement (Prime, Version 3.3) mit Standardeinstellungen unterzogen. Dies betraf die folgenden Regionen: Gln260-Gln262, Ala349-Arg352, Cys405-Pro412, Lys454-Asp458 und Thr494-Thr501. Anschließend wurde das Modell zur Vermeidung sterischer Fehler, wie bereits bei den anderen beiden Programmen erwähnt, energieminiert (Abbruchkriterium: Norm des Gradienten $\leq 0,05$). Dabei wurde das Kraftfeld OPLS2005 [104] verwendet.

Die Koordinierung des Zink-Ions in der Bindetasche der ADAMTS-Proteine mittels dreier Histidin-Reste (über den Imidazolring) ist hinlänglich bekannt [44]. In der Studie von Harding [167], in welcher die Koordinierung von Metallionen in Proteinstrukturen aus der PDB untersucht wurde, wurde beobachtet, dass der durchschnittliche Abstand zwischen Zink-Ionen und dem Stickstoff des Imidazolrings 2,03 Å (mit einer Standardabweichung von 0,05 Å) (siehe Abb. 3.1) beträgt. Aus diesem

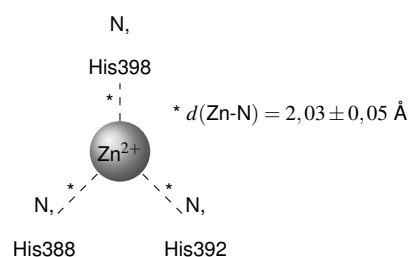


Abbildung 3.1: Darstellung der Bedingung für den Abstand zwischen Zink-Ion und Imidazolring der Histidine in Proteinstrukturen für ADAMTS7 nach Harding [167].

Grund wurden dem Modell beim Homology Modeling mit Schrödinger Prime vor der Energieminimierung *zero-order bonds*, welche die Interaktion und den Abstand zwischen den Zink-koordinierenden Histidin-Resten His388, His392 und His398 zum Zink-Ion definieren, hinzugefügt. Dies sollte die Koordinierung des Zink-Ions in der ADAMTS7-Bindetasche im Modell sicherstellen.

Zusammengefasst konnten mit den drei Programmen SWISS-MODEL, YASARA und Schrödinger Prime vier ADAMTS7-Homologiemodelle erstellt werden, welche entweder die Metalloproteinase-Domäne oder die Metalloproteinase- und die Disintegrin-ähnliche Domäne abdecken. Das Programm SWISS-MODEL erlaubt keine Berechnung von Modellen basierend auf mehreren Röntgenkristallstrukturen. Die erhaltenen SWISS-MODEL-Modelle umfassen die Aminosäuren Ser239 bis Ile457 (Template: 3HYG) und Ser239 bis Pro529 (Template: 2RJQ). Das mit YASARA berechnete Modell umfasst die Aminosäure Pro229 bis Gly539, und das mit Schrödinger Prime erstellte ADAMTS7-Modell deckt die Aminosäuren Lys240 bis Pro529 ab. Somit konnte mit dem Programm YASARA ein ADAMTS7-Homologiemodell erstellt werden, welches 311 Aminosäuren umfasst und somit die größte Abdeckung der berechneten Modelle zeigt (vergleiche Tabelle 3.3).

3.3 Modellvalidierung der ADAMTS7-Modelle

Im Folgenden wurde die Qualität der Modelle untersucht. Die vier erstellten ADAMTS7-Modelle wurden einer Validierung unterzogen, um das „beste“ Modell, welches der nativen ADAMTS7-Proteinstruktur wahrscheinlich am nächsten kommt, zu bestimmen und die Plausibilität der Modelle zu verifizieren. Die Modelle wurden anhand der nachfolgend aufgelisteten Kriterien untersucht und bewertet, um die Güte zu prüfen:

- Abdeckung der Sequenz und räumliche Struktur,
- Disulfidbrücken,
- Zink-Koordinierung und
- die globale und lokale Qualität.

Zur Prüfung der globalen und lokalen Qualität wurden folgende Programme genutzt: ProSa2003 [107], PROCHECK [110] und MetaMQAPII [109] (siehe Abschnitt 2.4).

3.3.1 Abdeckung der Sequenz und Strukturvergleich

In Tabelle 3.3 sind die erstellten Homologiemodelle von ADAMTS7 sortiert nach Template sowie die Größe der Modelle aufgeführt. Alle erstellten Modelle erhalten der Übersichtlichkeit wegen ein Kürzel aus genutzter Software (YA steht für YASARA, SM für SWISS-MODEL und SP für Schrödinger Prime) und fortlaufender Nummer. Das mit dem Template 3HYG erstellte Modell (SM1) weist von allen die geringste sequenzielle Abdeckung (12,9 %) zum ADAMTS7-Proteins auf, was auf die Größe der Röntgenkristallstruktur zurückzuführen ist. Das nur basierend auf dem Template 2RJQ (291 Aminosäuren lang) erstellte Modell (SM2) und die zusammengesetzten Modelle (SP1, YA1) decken C-terminal je nach Modell weitere 60-80 Aminosäuren des Targetproteins ab und bilden somit zwischen 17,1 % (SP1) und 18,4 % (YA1) des Proteins ab.

Tabelle 3.3: ADAMTS7-Homologiemodelle sowie (prozentuale) Abdeckung der Sequenz mitsamt Vergleich der Bindetaschen der Modelle anhand des $RMSD_{C\alpha}$. Die Modelle erhalten zur Übersicht ein Kürzel aus Abkürzung der genutzten Software (YA - YASARA, SM - SWISS-MODEL, SP - Schrödinger Prime) sowie fortlaufender Nummerierung. AS - Aminosäure(n).

Template (PDB-ID)	Modell	Abdeckung	Modellierter Anteil des Proteins (%)	Bindetaschenvgl. ($RMSD_{C\alpha}$ (Å))			
				SM1	SM2	SP1	YA1
3HYG	SM1	Ser239-Ile457	12,9 % (219 AS)	0	1,53	3,85	1,98
2RJQ	SM2	Ser239-Pro529	17,2 % (291 AS)	1,53	0	3,17	2,08
3HYG + 2RJQ	SP1	Lys240-Pro529	17,1 % (290 AS)	3,85	3,17	0	3,63
3HYG + 2RJQ	YA1	Pro229-Gly539	18,4 % (311 AS)	1,98	2,08	3,63	0

Der Vergleich der globalen Struktur der Modelle zeigte, dass diese sich in den Sekundärstrukturen in der Metalloproteinase-Domäne nicht unterscheiden (Modelle nicht abgebildet). Alle Modelle bilden fünf α -Helices und fünf β -Faltblätter aus. In der Disintegrin-ähnlichen Domäne zeigen die Modelle SM2, SP1 und YA1 alle eine α -Helix, die Anzahl an β -Faltblättern variiert. Während SP1 drei Faltblätter aufweist, enthalten SM2 sowie YA1 jeweils fünf β -Faltblätter.

Die Bindetasche des Proteins spielt in dieser Arbeit eine übergeordnete Rolle, daher wurde der strukturelle Unterschied der Bindetasche in den Modellen analysiert. Die ADAMTS7-Bindetasche wurde dafür durch das β -Faltblatt S4 (ADAMTS7: Met350-Met365) sowie die α -Helix H4 (ADAMTS7: Leu380-Gly400) und

die Loop-Region L1 (ADAMTS7: Ser401-Arg430) definiert (siehe Abb. 3.2). Bei diesen drei Regionen handelt es sich um die Bereiche, welche in räumlicher Nähe zum Zink-Ion und somit die Bindetasche von ADAMTS7 bilden. Vergleicht man die Strukturen anhand des RMSD der C_{α} -Atome in diesem Bereich (siehe Tabelle 3.3) zeigen alle Modelle eine Differenz zwischen 1,53 Å und 3,85 Å. Dieser Unterschied ist durch die Loop-Region in der Bindetasche zu erklären, welche in allen Modellen aufgrund der Flexibilität unterschiedlich modelliert wurde. Somit zeigen die Modelle bis zu einem gewissen Grad strukturelle Unterschiede.

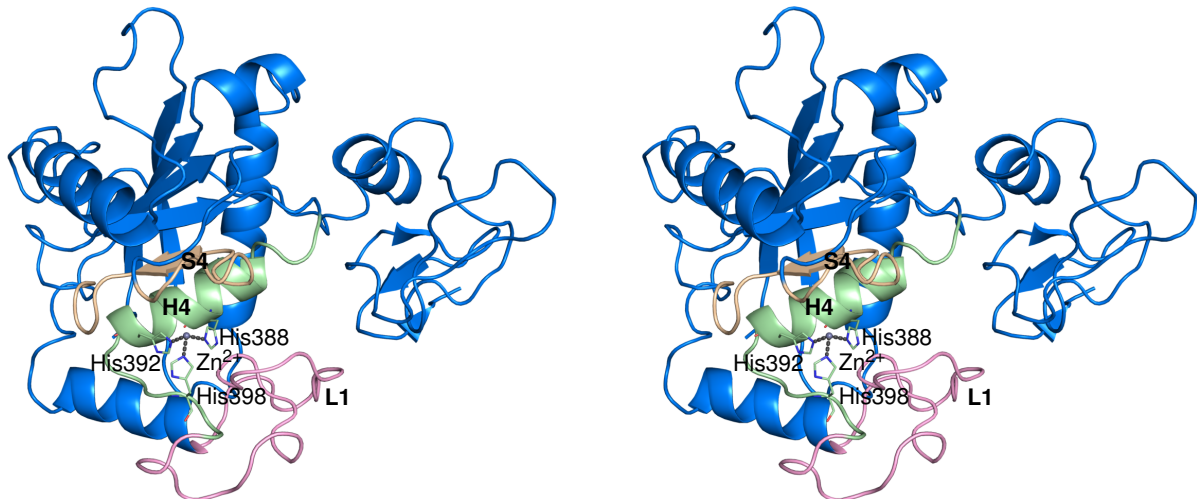


Abbildung 3.2: Beispielhafte Darstellung und Charakterisierung der ADAMTS7-Bindetasche in *cross-eye*-Darstellung anhand des Modells SP1. Die drei Bereiche der Tasche sind farblich kodiert: der Bereich um β -Faltblatt S4 (Met350-Met365) beige, die Region um α -Helix H4 (Leu380-Gly400) pastellgrün und die Loop-Region L1 (Ser401-Arg430) rosé. Das Protein ist in Cartoon-Darstellung, die Zink-koordinierenden Histidine als Linien veranschaulicht. Das Zink-Ion ist als graue Kugel zu erkennen und die Zink-koordinierenden Bindungen als grau gestrichelte Linien.

3.3.2 Disulfidbrücken

In den Röntgenkristallstrukturen der ADAMTS7-Templates 3HYG und 2RJQ wurde beobachtet, dass die Strukturen durch Disulfidbrücken stabilisiert werden. Diese kovalente Bindung ist in den Röntgenkristallstrukturen von 3HYG und 2RJQ zwischen den Seitenketten folgender Cystein-Reste ausgebildet: Cys342-Cys394, Cys371-Cys376, Cys388-Cys471 und Cys426-Cys455. Die Röntgenkristallstruktur 2RJQ bildet, aufgrund der höheren sequenziellen Abdeckung und somit größeren Proteinstruktur, zusätzlich noch vier weitere Disulfidbrücken aus: Cys497-Cys519, Cys508-Cys529, Cys514-Cys548 und Cys542-Cys553. Anhand des Alignments der Metalloproteinase- und Disintegrin-ähnlichen Domänen der ADAMTS-Proteine (siehe Anhang A.1.2, Abb. A.3) wird deutlich, dass insgesamt fünf der acht Cystein-Aminosäurepaare, die in den ADAMTS5-Strukturen Disulfidbrücken bilden, in den ADAMTS-Proteinen konserviert sind. Diese korrespondieren in ADAMTS7 zu: Cys318-Cys372, Cys347-Cys354, Cys366-Cys447, Cys485-Cys503 und Cys516-Cys527. Aus diesem Grund wurden alle erstellten Modelle auf das Vorhandensein von Disulfidbrücken zwischen den konservierten Cysteinen untersucht. Alle konservierten Disulfidbrücken sind in allen vier berechneten ADAMTS7-Modellen ausgebildet. Dieses Kriterium wurde demzufolge in allen Modellen, unabhängig von der genutzten Software und gewähltem Template, erfüllt und stellte somit kein Entscheidungsmerkmal dar.

3.3.3 Zink-Koordinierung

Das katalytische Zink-Ion in der ADAMTS7-Bindetasche ist ein wichtiges Merkmal, genau wie die Koordinierung dieses Metall-Ions über drei konservierte Histidine (His388, His392, His398). In den mit SWISS-MODEL erstellten Modellen wurden keine Zink-Ionen in die Homologiemodelle eingefügt, da dies zum Zeitpunkt der Modellierung mit dem Programm nicht möglich war. Daher ist die Zink-Koordinierung in den Modellen SM1 und SM2 nicht vorhanden. In den Modellen SP1 und YA1 wurde die Zink-Koordinierung beobachtet.

3.3.4 Globale und lokale Qualität

Anschließend wurde die globale sowie die lokale Qualität der ADAMTS7-Modelle untersucht. In Tabelle 3.4 sind die Ergebnisse des Ramachandran-Plots (PROCHECK) für die erstellten Homologiemodelle von ADAMTS7 sortiert nach Template aufgeführt. Anhand des Ramachandran-Plots wird deutlich, dass SM1 und SM2 Aminosäuren mit Torsionswinkeln in energetisch ungünstigen Regionen aufweisen. Die beiden zusammengesetzten Modelle YA1 und SP1 zeigen hingegen keine Auffälligkeiten im Ramachandran-Plot.

Tabelle 3.4: Modellvalidierung der erstellten ADAMTS7-Modelle anhand des mit PROCHECK erstellten Ramachandran-Plots.

Template (PDB-ID)	Modell	Ramachandran-Plot			
		bevorzugt erlaubter Bereich	zusätzlich erlaubter Bereich	erlaubter Bereich	nicht erlaubter Bereich
3HYG	SM1	84,2 %	12,6 %	2,1 %	1,1 %
2RJQ	SM2	86,6 %	11,9 %	0,4 %	1,2 %
3HYG + 2RJQ	SP1	83,3 %	15,5 %	1,2 %	0,0 %
3HYG + 2RJQ	YA1	81,5 %	17,8 %	0,7 %	0,0 %

Die Validierung mit ProSa2003, einem Programm, welches die Kompatibilität der Proteinmodelle mit der Aminosäuresequenz untersucht, zeigt in allen erstellten ADAMTS7-Modellen keine Auffälligkeiten (siehe Tabelle 3.5). Die ProSa2003-Ergebnisse bewerten somit alle Modelle als plausibel.

Die Analysen mit MetaMQAPII (Tabelle 3.5) zeigen nicht für alle ADAMTS7-Modelle optimale Werte (siehe Abschnitt 2.4 für Erläuterung der MetaMQAPII-Werte). Für das YASARA-Modell YA1 ergab die Analyse einen GDT_TS-Wert unter 45. Auch der RMSD-Wert dieses Modells (4,3 Å), welcher die Abweichung zur wirklichen Struktur schätzt, deutet auf strukturelle Unstimmigkeiten in dem Modell hin. Die GDT_TS- ($\geq 64,7$) und die RMSD-Werte ($< 2,5$ Å) der Modelle SM1, SM2 und SP1 hingegen weisen auf plausible Modelle hin.

Tabelle 3.5: Modellvalidierung der erstellten ADAMTS7-Modelle mit MetaMQAPII und ProSa2003.

Template (PDB-ID)	Modell	MetaMQAPII		Z-Scores (ProSa2003)		
		GDT_TS-Wert	RMSD (Å)	Combined	Pair	Surf
3HYG	SM1	71,6	2,3	-7,71	-6,12	-5,56
2RJQ	SM2	64,7	2,4	-8,02	-5,94	-5,97
3HYG + 2RJQ	SP1	65,4	2,4	-7,80	-5,83	-5,78
3HYG + 2RJQ	YA1	43,4	4,3	-7,48	-6,16	-5,29

Zusammengefasst zeigen die Untersuchungen mit PROCHECK im Vergleich die besseren strukturellen Werte für die Modelle YA1 und SP1. In den Modellen SM1 und SM2 wurden nicht nur die Zink-Ionen nicht

modelliert, sondern es konnten auch energetisch ungünstige Regionen anhand des Ramachandran-Plots aufgezeigt werden (siehe Tabelle 3.4). Des Weiteren konnte mit den MetaMQAPII-Analysen (siehe Tabelle 3.5) gezeigt werden, dass das YASARA-Modell YA1 schlechtere Werte (RMSD, GDT_TS) für die Wahrscheinlichkeit einer guten Prädiktion der ADAMTS7-Struktur als die anderen drei Modelle aufweist. Der RMSD ist bis zu 2,0 Å höher als bei SP1, SM1 und SM2. Weiterhin ist der GDT_TS-Wert im Vergleich um über 20 Punkte niedriger als bei den anderen drei Modellen. Die Kombination aller Ergebnisse brachte zum Vorschein, dass drei (SM1, SM2, YA1) der vier Modelle in mindestens einer Analyse Defizite aufweisen. Somit zeigte das Schrödinger Prime-Modell SP1 als einziges Modell keine Defizite in der Modellvalidierung und kommt der nativen ADAMTS7-Struktur wahrscheinlich am nächsten.

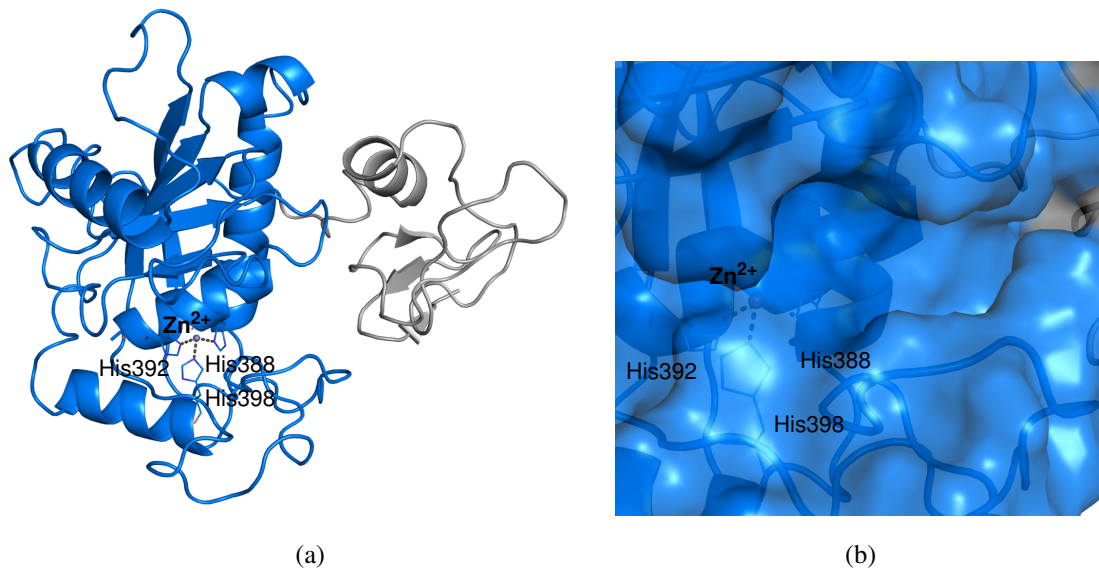
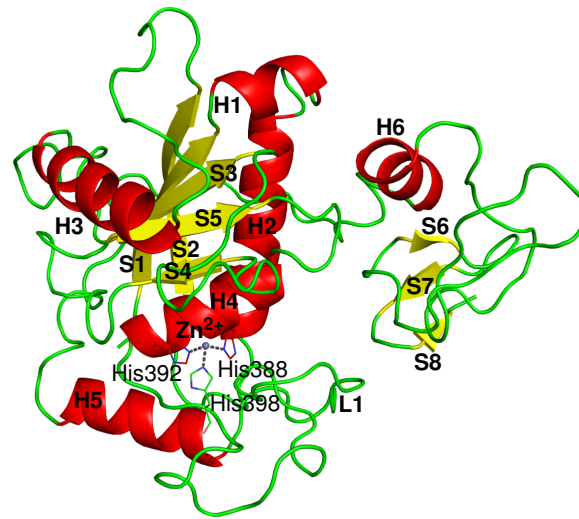


Abbildung 3.3: ADAMTS7-Modell (SP1) basierend auf den ADAMTS5-Templates 3HYG und 2RJQ erstellt mit Schrödinger Prime. (a) Modell anhand von Metalloproteinase- (blau) und Disintegrin-ähnlicher Domäne (grau) verschieden eingefärbt. Das Protein ist in der Cartoon-Darstellung, die Zink-koordinierenden Histidine als Linien veranschaulicht. Das Zink-Ion ist als graue Kugel zu erkennen, die Zink-koordinierenden Bindungen als grau gestrichelte Linien. (b) Ist eine Vergrößerung von (a) um 15° nach oben gedreht mit Darstellung der Oberfläche zur Verdeutlichung der ADAMTS7-Bindetasche.

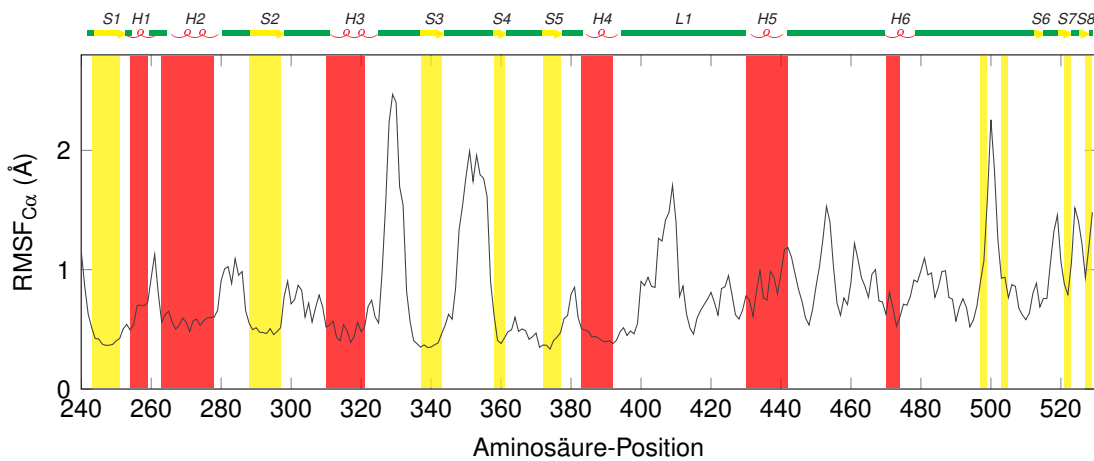
Abbildung 3.3a zeigt das ADAMTS7-Homologiemodell SP1 mit unterschiedlicher Kolorierung der zwei modellierten Domänen. Deutlich zu erkennen ist die Bindetasche mit dem, von den drei Histidinen koordinierten, Zink-Ion. In Abbildung 3.3b ist die ADAMTS7-Bindetasche vergrößert veranschaulicht und durch die Oberflächendarstellung entsteht ein räumlicher Eindruck der Höhle.

3.3.5 MD-Simulation des ADAMTS7-Modells

Um das Verhalten der ADAMTS7-Proteinstruktur unter physiologischen Bedingungen besser zu verstehen, wurde das „beste“ Modell, SP1 (siehe Abb. 3.4a), einer MD-Simulation unterzogen. Die MD-Simulation eines Proteins bietet die Möglichkeit, die Wechselwirkungen innerhalb der Struktur über einen definierten Zeitraum bei bestimmten Bedingungen (z. B. konstantem Druck und konstanter Temperatur) zu untersuchen (siehe Abschnitt 2.5.3). Die MD-Simulation des erstellten Modells der Metalloproteinase- und Disintegrin-ähnlichen Domäne von ADAMTS7 sollte folglich sicherstellen, dass es sich um eine realistische Struktur des Proteins handelt.



(a) ADAMTS7-Modell SP1



(b) Auswertung der lokalen Schwingungen in MD-Simulation

Abbildung 3.4: Darstellung der Sekundärstrukturen im ADAMTS7-Homologiemodell und Untersuchung der Stabilität dieser über die MD-Simulation. (a) Sekundärstrukturen im ADAMTS7-Homologiemodell. Das Modell SP1 ist in der Cartoon-Darstellung, die Zink-kordinierenden Histidine als Linien veranschaulicht. Das Zink-Ion ist als graue Kugel zu erkennen, die Zink-kordinierenden Bindungen als grau gestrichelte Linien. Die Sekundärstrukturen des Modells SP1 sind vom N- zum C-Terminus durchnummeriert und eingefärbt: α -Helices rot, β -Faltblätter gelb und Loop-Regionen grün. (b) Stabilität der Sekundärstrukturen in der MD-Simulation. $RMSF_{C\alpha}$ über die stabile Zeit der ADAMTS7-Simulation (14,5 ns bis 20 ns). Stabile Sekundärstrukturen über den Auswertungszeitraum sind als gelbe (β -Faltblatt) bzw. rote (α -Helix) Markierungen dargestellt. Zum Vergleich ist die Sekundärstruktur des Modells SP1 aus (a) über der Grafik verdeutlicht.

Die Simulation erfolgte mit dem Desmond-Paket [121] (Desmond Molecular Dynamics System, Version 3.6; Maestro-Desmond Interoperability Tools, Version 3.6) mit Standardeinstellungen und dem Kraftfeld OPLS2005 [104]. Das ADAMTS7-Modell SP1 als Startstruktur wurde in einer orthorhombischen Box mit jeweils mindestens 10 Å Abstand zum Protein simuliert. Die Lösung in der Box wurde mittels des SPC (Abk. für *single point charge*)-Wassermodells [168], zur Beschreibung der Wassermoleküle, definiert und physiologische Bedingungen wurden durch eine 0,15 M NaCl-Konzentration erzeugt. Des Weiteren wurde das System mit elf Natrium-Ionen neutralisiert. Die 20 ns lange MD-Simulation wurde bei einer Temperatur von 300 K und 1 bar Druck durchgeführt. Sie erfolgte somit unter thermodynamischen Gleichgewichtsbedingungen bei konstantem Druck und konstanter Temperatur (*NPT-Ensemble*).

Betrachtet man den $\text{RMSD}_{\text{C}\alpha}$ über die Simulationszeit (siehe Abschnitt 2.5.4 und Abb. A.4 im Anhang) wird ersichtlich, dass in der ersten Nanosekunde der Simulation der $\text{RMSD}_{\text{C}\alpha}$ auf etwa 2 Å ansteigt. Die MD-Simulation nimmt nach 14,5 ns einen stabilen Gleichgewichtszustand ein und der $\text{RMSD}_{\text{C}\alpha}$ konvergiert gegen $2,8 \pm 0,5$ Å. Folglich zeigen die $\text{C}\alpha$ -Atome des Proteinmodells zwischen 14,5 ns und 20 ns keine signifikanten Änderungen im Vergleich zur Startstruktur. Daher kann davon ausgegangen werden, dass das Modell global einen stabilen Zustand eingenommen hat.

Um zu untersuchen, ob Bereiche des Proteins großen Fluktuationen unterliegen, wurde der $\text{RMSF}_{\text{C}\alpha}$ über alle Aminosäuren des Modells für den stabilen Auswertungszeitraum (14,5 - 20 ns) berechnet. In Abbildung 3.4b ist der $\text{RMSF}_{\text{C}\alpha}$ über diesen Zeitraum sowie die stabilen Sekundärstrukturen des Modells, welche in mehr als 70 % des Auswertungszeitraums ausgebildet sind, dargestellt. Der Vergleich zwischen den Sekundärstrukturen des Modells und der MD-Simulation zeigen, dass diese speziellen Motive (S1 bis S5, H1 bis H5) in der Metalloproteinase-Domäne stabil sind. Im Bereich der Disintegrin-ähnlichen Domäne kommt es während der MD-Simulation zur Ausbildung zweier zusätzlicher β -Faltblätter und zur Entfaltung des S6-Faltblattes. Nach Takeda [35] sind die Sekundärstrukturen der Disintegrin-ähnlichen Domäne so wie im ADAMTS7-Homologiemodell vorliegend konserviert. Daher wurde davon ausgegangen, dass die Veränderungen in der Sekundärstruktur während der MD-Simulation auf fehlende Interaktionen mit den nicht modellierbaren Domänen begründet sein könnten.

Weiterhin konnten in Loop-Regionen stärkere Schwankungen ($\text{RMSF}_{\text{C}\alpha} > 1,5$ Å) beobachtet werden. Dies betrifft die Loops zwischen Helix H3 und Faltblatt S3, zwischen den Faltblättern S3 und S4 (Bindetasche), zwischen den Helices H4 und H5 (L1-Loop, Bindetasche) sowie zwischen H6 und S6. Die Regionen mit ausgebildeten Sekundärstrukturen weisen keine hohen Abweichungen ($\text{RMSF}_{\text{C}\alpha} > 1,5$ Å) auf. In Loop-Regionen liegt kein „festes“ Interaktionsschema über Wasserstoffbrücken vor. Daher sind höhere RMSF-Werte in diesen Regionen im Vergleich zu Regionen mit α -Helix- oder β -Faltblattstruktur, die durch ein Wasserstoffbrücken-Muster stabilisiert werden, nicht ungewöhnlich (siehe Abschnitt 2.5.4). Loop-Regionen zeigen im Allgemeinen erhöhte RMSF-Werte aufgrund ihrer Flexibilität, die mit dem Fehlen der Stabilisierung des Proteinrückgrats begründet ist. Die Beobachtungen in der Simulation deuten auf ein natives Verhalten der jeweiligen Regionen des Proteins hin. Jedoch sollte die beobachtete Flexibilität der ADAMTS7-Bindetasche in den weiteren Untersuchungen beachtet werden, da sie eine entscheidende Rolle bei der Ligandenbindung, wie bereits in Abschnitt 1.3.2 erwähnt, einnehmen könnte.

Anhand der MD-Simulation konnte gezeigt werden, dass das ausgewählte ADAMTS7-Modell SP1 eine stabile Konformation aufweist und somit ein potenziell natives Modell für das hier untersuchte Zielprotein darstellt.

3.4 Conformational Ensembles von flexiblen Regionen in ADAMTS7-Bindetasche

Die MD-Simulation des ADAMTS7-Modells (siehe Abschnitt 3.3.5) hat gezeigt, dass in der Bindetasche sehr flexible Bereiche in den Loop-Regionen vorliegen. Dies könnte entscheidend für die Ligandenbindung und damit für die Suche nach potenziellen ADAMTS7-Inhibitoren sein. Um dies genauer zu untersuchen, wurden die bekannten ADAMTS-Proteine sowie die zugehörige Literatur hinsichtlich der Flexibilität der Bindetasche analysiert, um dieser gegebenenfalls Rechnung zu tragen. Flexible Regionen in der ADAMTS7-Bindetasche müssten unter Umständen zur Verbesserung der Proteinstruktur modelliert und in die weiteren Analysen mit einbezogen werden.

In der Literatur [47, 164, 165] wurde bereits gezeigt, dass ADAMTS-Proteine in der Bindetasche flexible Loop-Regionen aufweisen, die sich der Ligandenstruktur anpassen können. Moysak *et al.* [47], die die (nicht vollständigen) Röntgenkristallstrukturen der Proteine ADAMTS4 (PDB-ID: 2RJP, 3B2Z) und -5 (PDB-ID: 2RJQ) experimentell bestimmten, konnten zeigen, dass die Inhibitor-Bindestellen dieser Proteine unterschiedliche Konformationen einnehmen. Die miteinander eng verwandten Proteine zeigen grundsätzlich den gleichen strukturellen Aufbau (Sekundärstruktur) [47], jedoch sind Unterschiede in den S1'- und S2'-Loops zu erkennen. Die Aminosäurereste His389-Pro403 in ADAMTS4 bzw. Arg437-Pro451 in ADAMTS5 umfassen den S1'-Loop. Der S2'-Loop enthält die Aminosäuren Cys322-Leu330 in ADAMTS4 und Cys371-Leu379 in ADAMTS5. Die Unterschiede im S2'-Loop werden durch den Vergleich zwischen der apo (3B2Z) und der Inhibitor-gebundenen (2RJP) ADAMTS4-Struktur deutlich. In Abbildung 3.5 sind 2RJP und 3B2Z strukturell aligniert dargestellt: Es ist zu erkennen, wie sich der S2'-Loop in der apo Struktur

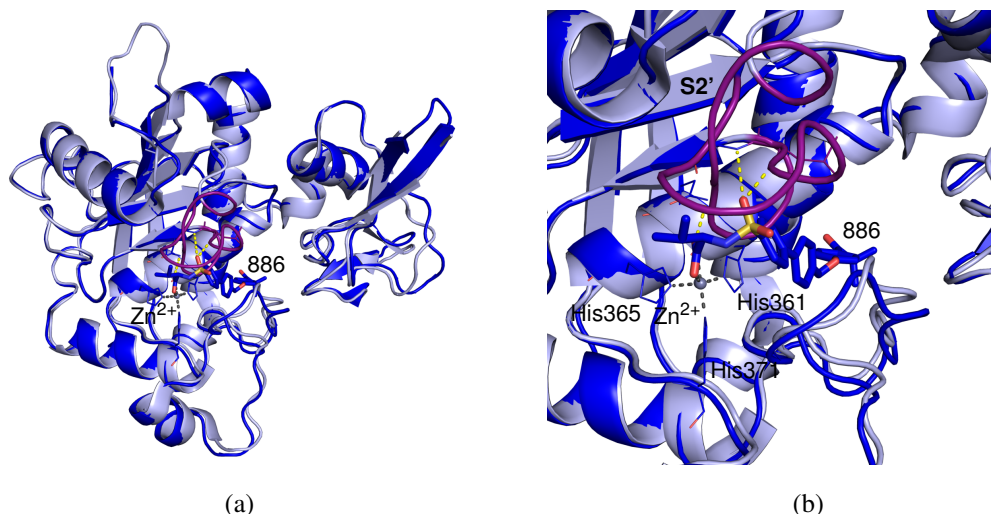


Abbildung 3.5: Vergleich des S2'-Loops anhand der ADAMTS4-Röntgenkristallstrukturen 2RJP (Inhibitor 886 gebunden, dunkelblau) und 3B2Z (apo, hellblau). Der S2'-Loop ist lila eingefärbt. Die ADAMTS4-Röntgenkristallstrukturen sind in Cartoon-Darstellung, die Zink-koordinierenden Histidine als Linien dargestellt, genau wie die mit dem Inhibitor wechselwirkenden Aminosäuren (in 2RJP). Der Inhibitor 886 wird durch Stäbchen dargestellt und das Zink-Ion ist als graue Kugel zu erkennen. Zink-koordinierende Bindungen sind durch graue gestrichelte Linien dargestellt, Wasserstoffbrücken durch gelb gestrichelte Linien. (a) Gesamte Röntgenkristallstrukturen. (b) Zoom auf Bindetasche.

in den Bereich des Inhibitors 886 verschoben hat. Moysak *et al.* [47] postulieren, dass dieser Loop im apo Zustand (3B2Z) eine geschlossene Konformation einnimmt. Wohingegen sich der S2'-Loop in der Rönt-

genkristallstruktur 2RJQ dem Liganden anpasst. In der restlichen Konformation unterscheiden sich die zwei Strukturen nicht.

Sieh *et al.* [164] zeigten in ihren Untersuchungen der ADAMTS5-Struktur (PDB-ID: 3B8Z) im Vergleich mit den ADAM33- (PDB-ID:1R55) und MMP3- (PDB-ID: 1D7X) Strukturen, dass sich die Metalloproteinasen-Domänen in der globalen Struktur ähneln, aber sich der S1'-Loop visuell zwischen den drei Proteinen stark unterscheidet und flexibel ist. In einer weiteren Arbeit von Sieh *et al.* [165], in welcher die Bindetaschen der ADAMTS5-Röntgenkristallstrukturen 3B8Z und 3LJT verglichen wurden, konnte bestätigt werden, dass der S1'-Loop in diesem Protein flexibel ist und seine Konformation größeren Inhibitoren anpassen kann. In Abbildung 3.6 ist diese konformelle Änderung zwischen den zwei Strukturen dargestellt. Es ist zu erkennen, dass sich die Konformation des S1'-Loops in Abhängigkeit vom Inhibitor an diesen (hier: LA3 oder 294) anpasst. Zusammenfassend lassen die Ergebnisse vermuten, dass die Flexibilität des S1'-

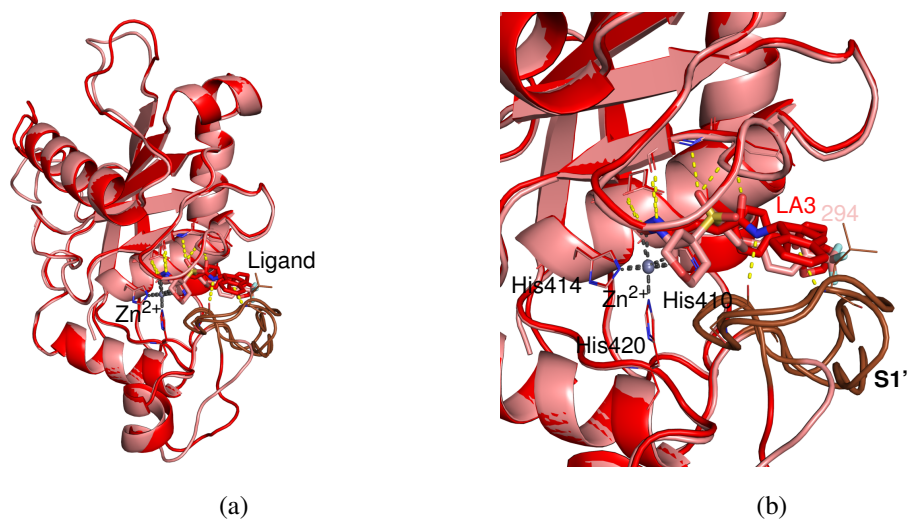


Abbildung 3.6: Vergleich des S1'-Loops in den ADAMTS5-Röntgenkristallstrukturen 3LJT (rot) und 3B8Z (rosé). Der S1'-Loop ist braun eingefärbt. ADAMTS5-Röntgenkristallstrukturen in Cartoon-Darstellung, die Zinkkoordinierenden Histidine als Linien dargestellt, genau wie die mit dem Inhibitor wechselwirkenden Aminosäuren. Der Inhibitor (in 3LJT: LA3, in 3B8Z: 294) wird durch Stäbchen dargestellt und das Zink-Ion ist als graue Kugel zu erkennen. Zink-koordinierende Bindungen sind durch graue gestrichelte Linien dargestellt, Wasserstoffbrücken durch gelb gestrichelte Linien. (a) Gesamte Röntgenkristallstrukturen. (b) Zoom auf Bindetasche.

und S2'-Loops einen Einfluss auf die Bindung sowie auf die Selektivität von ADAMTS-Inhibitoren haben könnte.

Aufgrund der Homologien in der ADAMTS-Proteinfamilie konnte angenommen werden, dass die erwähnten Bereiche im ADAMTS7-Protein ebenfalls flexibel sind. Betrachtet man die entsprechenden S1'- und S2'-Loops (ADAMTS7: S1'-Loop: Phe413-Thr427; S2'-Loop: Cys347-Leu357) im Targetprotein (siehe Abb. 3.7), liegt der S1'-Loop in räumlicher Nähe zum Zink-Ion in der Bindetasche von ADAMTS7 und ist Teil des L1-Loops. Der S2'-Loop liegt im oberen Bereich der Bindetasche, ragt aus der Bindetasche heraus und liegt vor dem S4-Faltblatt. Die MD-Simulation des ADAMTS7-Modells (siehe Abschnitt 3.3.5) zeigte, dass der S1'- und S2'-Loops hohe strukturelle Abweichungen während der Simulationszeit aufweisen (vergleiche Abb. 3.4b). Dies bestätigt somit die Vermutung, dass es sich bei diesen beiden Bereichen um flexible Regionen in der ADAMTS7-Bindetasche handelt. Das erstellte ADAMTS7-Modell könnte somit nur eine mögliche Konformation der Bindetasche darstellen.

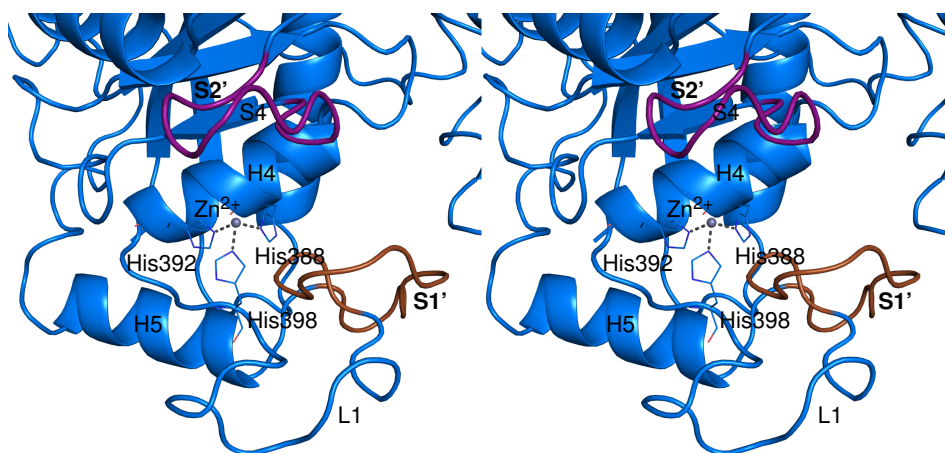


Abbildung 3.7: S1'-Loop (braun) und S2'-Loop (lila) im ADAMTS7-Homologiemodell in *cross-eye*-Darstellung. Protein in Cartoon-Darstellung, die Zink-koordinierenden Histidine als Linien veranschaulicht. Das Zink-Ion ist als graue Kugel zu erkennen. Zink-koordinierende Bindungen sind durch graue gestrichelte Linien dargestellt.

Um die Flexibilität des S1'- und S2'-Loops in die weiteren Analysen integrieren zu können, wurde ein Loop-Sampling durchgeführt. Die dadurch erhaltenen Konformationen sollten helfen den gesamten Konformationsraum der Bindetasche abzudecken. Dies sollte sicherstellen, dass die Flexibilität der ADAMTS7-Bindetasche bei den weiteren Analysen, vor allem bei einem Virtual Screening zur Geltung kommt.

Für das Loop-Sampling des S1'-Loops wurde ein Loop-Refinement (Prime, Version 3.3, Schrödinger, LLC, New York, NY, 2013) [103] (siehe Abschnitt 2.3.2) durchgeführt. Aus dieser Analyse resultierten zwölf verschiedene Konformationen des S1'-Loops, die mit einem $\text{RMSD}_{\text{C}\alpha}$ von 5,15 Å – 7,12 Å oder mehr von der S1'-Loop-Region des Homologiemodells abwichen. Mittels Clustern wurden Konformationen dieser Region zusammengefasst, die sich um weniger als 2 Å im $\text{RMSD}_{\text{C}\alpha}$ unterschieden. Bildeten mehrere Konformationen ein Cluster, wurde eine Durchschnittsstruktur ermittelt, die anschließend mit dem Kraftfeld OPLS2005 energieminiert wurde. Somit konnten durch die Clusteranalyse, ausgehend von zwölf S1'-Loop-Konformationen, vier weitere Konformationen dieses Loops vorhergesagt werden. In Tabelle 3.6 sind die strukturellen Abweichungen zwischen dem ADAMTS7-Homologiemodell und den jeweiligen Sampling-Konformationen aufgeführt. Die Unterschiede zwischen den einzelnen Loop-Konformationen liegen zwischen 2,57 Å und 7,12 Å. Mittels dieser Vorgehensweise kann die Flexibilität der ADAMTS7-Bindetasche modelliert und ein großer Bereich des Konformationsraums abgedeckt werden.

Tabelle 3.6: Strukturvergleich der fünf S1'-Loop-Konformationen von ADAMTS7 anhand des $\text{RMSD}_{\text{C}\alpha}$ in Å. HM - ADAMTS7-Homologiemodell, K1-K4 bezieht sich auf die vier S1'-Loop-Konformationen.

$\text{RMSD}_{\text{C}\alpha}$ (Å)	HM	K1	K2	K3	K4
HM	0	7,12	5,15	5,56	6,69
K1	7,12	0	5,88	5,65	2,57
K2	5,15	5,88	0	6,19	4,63
K3	5,56	5,65	5,57	0	5,33
K4	6,69	2,57	6,19	5,33	0

Ein Sampling des S2'-Loop nach gleichem Vorgehen wie für den S1'-Loop ergab keine weitere Konformationen, die mehr als 2 Å vom ADAMTS7-Homologiemodell abwichen. So konnte keine Veränderung in dieser Region nachgewiesen werden, welche die Flexibilität der Bindetasche beeinflusst. Aus diesem Grund wurde das Loop-Sampling des S2'-Loops verworfen.

In Abbildung 3.8 ist das ADAMTS7-Homologiemodell sowie die vier weiteren Konformationen des S1'-Loops dargestellt. Die verschiedenen Konformationen unterscheiden sich im Volumen der Bindetasche, welches mit Schrödinger SiteMap (Version 2.9) [125, 126] berechnet wurde. Das Bindetaschenvolumen des Homologiemodells ist dabei mit rund 319 \AA^3 am Größten. Die Volumen der vier weiteren Konformationen sind geringer und liegen im Bereich von 135 \AA^3 bis 213 \AA^3 . Die Flexibilität des S1'-Loops könnte für die Anpassung an mögliche Liganden von Bedeutung sein.

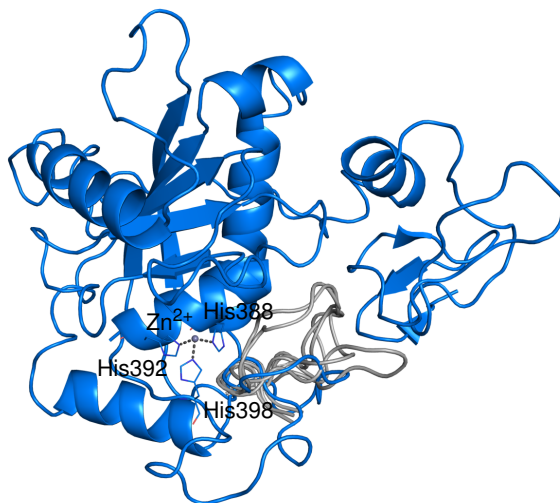


Abbildung 3.8: Sampling des S1'-Loops von ADAMTS7 (Phe413-Thr427). ADAMTS7-Homologiemodell (blau) und vier Konformationen des S1'-Loops (grau) in Cartoon-Darstellung. Die Zink-koordinierenden Histidine sind als Linien veranschaulicht. Das Zink-Ion ist als graue Kugel zu erkennen.

Insgesamt konnten neben dem ADAMTS7-Homologiemodell vier weitere Konformationen der Bindetasche vorhergesagt werden, die die Flexibilität des S1'-Loops berücksichtigen. Diese verschiedenen Konformationen der Proteinbindetasche werden auch als *Conformational Ensemble* bezeichnet. Die nun insgesamt fünf Strukturen der ADAMTS7-Bindetasche sollten für die weiteren Analysen bei der Suche nach ADAMTS7-Inhibitoren Anwendung finden.

3.5 Zusammenfassung

Ausgehend von 15 bekannten ADAMTS-Röntgenkristallstrukturen konnten zwei mögliche Template-Strukturen für das Homology Modeling von ADAMTS7 identifiziert werden. Unter Nutzung dreier verschiedener Programme wurden vier Modelle erstellt. Durch die Beurteilung der Qualität der Homologiemodelle mittels einer Modellvalidierung war es möglich, ein ADAMTS7-Modell zu identifizieren, welches der nativen ADAMTS7-Struktur vermutlich am nächsten kommt. Dabei handelt es sich um das Modell SP1, welches mit Schrödinger Prime erzeugt wurde. Eine MD-Simulation des Modells bestätigte die Ergebnisse der Modellvalidierung. Wird im Folgenden das ADAMTS7-Modell betitelt, ist immer das ausgewählte Modell SP1 gemeint.

Die MD-Simulation des ADAMTS7-Homologiemodells zeigte zudem, dass die Bindetasche flexible Bereiche besitzt. Publierte Arbeiten [47, 164, 165] zu ADAMTS4 und ADAMTS5 und strukturelle Vergleiche der Konformationen bestätigten diese Ergebnisse: Es wurde postuliert, dass sich zwei Loop-Bereiche (S1'- und S2'-Loop) in den ADAMTS4- und ADAMTS5-Bindetaschen dem jeweiligen Inhibitor bei der

Bindung anpassen können. Daher wurde angenommen, dass sich ein ähnliches Verhalten auch bei ADAMTS7 zeigen könnte. Ein Loop-Sampling für ADAMTS7 ergab vier weitere Konformationen des S1'-Loops und keine weitere Konformation des S2'-Loops.

Das ADAMTS7-Homologiemodell sowie die vier weiteren Konformationen (aus dem Loop-Sampling) charakterisieren somit die für die folgenden Analysen essenzielle ADAMTS7-Bindetasche. Das Modell des Targetproteins ADAMTS7 sowie das Conformational Ensemble bilden die Grundlage für die anschließenden Screening-Analysen. Diese sollten die Suche nach potenziellen Inhibitoren, die in die ADAMTS7-Bindetasche binden, ermöglichen.

4 Ergebnisse und Diskussion: Untersuchungen der ADAMTS-Familie

Die ADAMTS-Familie zeichnet sich durch einen ähnlichen Aufbau aus und alle Proteine beinhalten eine Metalloproteinase-Domäne mit einer Bindetasche (siehe hierzu Abschnitt 1.3). Um bei der Suche nach ADAMTS7-Inhibitoren Kreuzreaktionen und unerwünschte Nebenwirkungen zu vermeiden, war das Ziel dieser Arbeit potenzielle, selektive Inhibitoren vorherzusagen. Aus diesem Grund sollten Sequenz- und Strukturanalysen durchgeführt werden, um Gemeinsamkeiten und vor allem Unterschiede innerhalb der ADAMTS-Familie aufzudecken. Diese Ergebnisse sollten im Anschluss im ADAMTS7-Virtual Screening (siehe Kapitel 5) Anwendung finden.

Zum Zeitpunkt der Arbeit waren nur Röntgenkristallstrukturen der Metalloproteinase-Domäne von drei ADAMTS-Proteinen bekannt. Somit sollte im ersten Schritt die Diskrepanz an fehlenden experimentellen Daten der Metalloproteinase-Domäne bei 16 der 19 ADAMTS-Proteine durch Homologiemodelle (siehe Abschnitt 4.1) kompensiert werden. Ein Sequenzvergleich der ADAMTS-Proteine (siehe Abschnitt 4.2) sowie im Speziellen der Metalloproteinase-Domäne (siehe Abschnitt 4.2.1) sollte Aufschluss über die Konservierung der Proteine untereinander und vor allem über den Konservierungsgrad der Binderegion von Inhibitoren geben. Anhand der Sequenz- und Strukturanalysen der ADAMTS-Bindetaschen (siehe Abschnitt 4.2.2) sollten Eigenschaften von potenziellen, selektiven ADAMTS7-Inhibitoren abgeleitet werden. Außerdem wurden bekannte ADAMTS-Protein-Inhibitor-Komplexe untersucht (siehe Abschnitt 4.3), um mögliche essenzielle Wechselwirkungen mit Aminosäuren in den ADAMTS-Bindetaschen aufzudecken. So sollten erste Folgerungen darüber, welche Eigenschaften selektive Inhibitoren ausmachen könnten bzw. wie selektive Inhibitoren gefunden werden könnten, gezogen werden. Diese Informationen sollten in ein Virtual Screening einfließen und eine Anreicherung von potenziellen, selektiven ADAMTS7-Inhibitoren ermöglichen.

4.1 ADAMTS-Modelle

Ein Strukturvergleich der Bindetaschen der ADAMTS-Proteine setzt voraus, dass für alle Proteine dieser Familie räumliche Daten vorhanden sind. Jedoch liegen nur für ADAMTS1, -4 und -5 experimentelle Daten für die Metalloproteinase-Domäne vor. Daher wurden zur Analyse und für den Vergleich Modelle der Proteine ohne bekannte 3D-Struktur, insbesondere deren Metalloproteinase-Domäne, benötigt, um das Defizit an fehlenden räumlichen Daten auszugleichen.

Bevor die ADAMTS-Proteine ohne bekannte räumliche Struktur der Metalloproteinase-Domäne modelliert werden sollten, galt es für ADAMTS1, -4 und -5 repräsentative Strukturen auszuwählen. Kriterium für die Wahl der repräsentativen experimentellen Struktur war zuallererst die Auflösung. Bei gleicher Auflösung wurde jene Struktur, die als Inhibitor-Protein-Komplex vorliegt, gewählt. Beinhaltet die ausgewählte

Röntgenkristallstruktur mehrere Proteinketten, wurde anhand der sequenziellen Abdeckung und struktureller Parameter (B-Faktor und Ramachandran Plot) jene mit den besseren Werten bestimmt.

Für die ADAMTS1-Struktur wurde die Röntgenkristallstruktur mit dem PDB-ID 2V4B, aufgrund der genauesten Auflösung von 2,0 Å aller Strukturen dieses Proteins, ausgewählt. Die Röntgenkristallstruktur beinhaltet zwei Proteinketten, wobei die besseren strukturellen Werte für Kette B bestimmt wurden (siehe Tabelle 4.1). Diese Proteinkette beinhaltet keine Aminosäuren in energetisch ungünstigen Bereichen. Folglich ist diese 3D-Struktur als stellvertretende ADAMTS1-Struktur genutzt worden. Als ADAMTS4-

Tabelle 4.1: Ergebnisse des Ramachandran-Plots (PROCHECK) für die ADAMTS1-Röntgenkristallstruktur 2V4B (aufgeschlüsselt nach Kette).

Protein	PDB-ID	Kette	Ramachandran-Plot			
			bevorzugt erlaubter Bereich	zusätzlich erlaubter Bereich	erlaubter Bereich	nicht erlaubter Bereich
ADAMTS1	2V4B	A	93,3 %	6,3 %	0,0 %	0,4 %
	2V4B	B	90,0 %	9,2 %	0,8 %	0,0 %

Vergleichsstruktur wurde 2RJP ausgewählt. Beide zum Zeitpunkt der Analysen bekannten ADAMTS4-Strukturen haben eine Auflösung von 2,8 Å und beinhalten jeweils eine mutierte Aminosäure, sodass sie nicht dem Wildtyp entsprechen. Schlussendlich wurde die mit einem Inhibitor kokristallisierte Struktur 2RJP für die weiteren Analysen genutzt. Die Röntgenkristallstruktur 2RJP hat insgesamt vier Ketten. Die größte Abdeckung der Sequenz haben die Ketten A und C, wobei Kette C den niedrigeren B-Faktor aufweist. Die strukturellen Untersuchungen mit PROCHECK zeigten, dass in Kette C nur eine Aminosäure (Cys467) im nicht erlaubten Bereich liegt (siehe Tabelle 4.2). Diese befindet sich nicht in der Bindetasche.

Tabelle 4.2: Ergebnisse des Ramachandran-Plots (PROCHECK) für die ADAMTS4-Röntgenkristallstruktur 2RJP (aufgeschlüsselt nach Kette).

Protein	PDB-ID	Kette	Ramachandran-Plot			
			bevorzugt erlaubter Bereich	zusätzlich erlaubter Bereich	erlaubter Bereich	nicht erlaubter Bereich
ADAMTS4	2RJP	A	89,6 %	10,0 %	0,0 %	0,4 %
		B	89,4 %	9,8 %	0,4 %	0,4 %
	2RJP	C	87,5 %	11,7 %	0,4 %	0,4 %
	2RJP	D	90,5 %	8,7 %	0,9 %	0,0 %

Die mutierte Aminosäure Gln362 in dieser Röntgenkristallstruktur wurde mit dem Programm Schrödinger Maestro (Version 9.5) in die Wildtyp-Aminosäure Glu362 umgewandelt und die Struktur anschließend mit dem Kraftfeld OPLS2005 energieminimiert (MacroModel, Version 10.3). Somit konnte die erhaltene Struktur für die weiteren Analysen verwendet werden.

Als Modell-Struktur für ADAMTS5 wurde 3HYG (Kette A) gewählt. Diese Struktur hat eine Auflösung von 1,4 Å auf, ist mit einem Inhibitor kokristallisiert und weist die genauesten strukturellen Parameter der ADAMTS5-Strukturen auf (siehe Tabelle 3.1). Daher wurde diese, analog zur Wahl des Templates für das ADAMTS7-Homology Modeling (siehe Abschnitt 3.1), als repräsentative ADAMTS5-Struktur für die weiteren Analysen genutzt. In Tabelle 4.3 sind die ADAMTS-Proteine mit bekannten 3D-Strukturen der Metallproteinase-Domäne und die jeweils für die folgenden Analysen der Bindetasche genutzte 3D-Struktur aufgeführt und dementsprechend markiert.

Für die 15 ADAMTS-Proteine ohne bekannte 3D-Struktur wurden, analog zum Vorgehen beim Homology Modeling von ADAMTS7 (siehe Kapitel 3), Modelle mit Schrödinger Prime (Version 3.3) erstellt. In diesem Schritt wurde auf die Anfertigung von zusammengesetzten Modellen verzichtet, weil lediglich die Metalloproteinase-Domänen sowie die Bindetaschen der Proteine dieser Familie verglichen werden sollten. Aus diesem Grund spielte eine größtmögliche Abdeckung der Proteinsequenz keine Rolle. Zusätzlich ist aus dem vorigen Kapitel bereits bekannt, dass keine komplette Struktur eines ADAMTS-Proteins vorliegt und die komplette Struktur von ADAMTS7 nicht modelliert werden konnte, was demzufolge auch für die anderen ADAMTS-Proteine gilt.

Als potenzielle Templates (siehe Tabelle 3.1 für die Eigenschaften der Röntgenkristallstrukturen) kamen, wie beim Homology Modeling von ADAMTS7, die ADAMTS1-Strukturen 2JIH, 2V4B, 3Q2G und 3Q2H; die ADAMTS4-Strukturen 2RJP und 3B2Z; sowie die bekannten ADAMTS5-Strukturen 2RJQ, 3B8Z, 3HY7, 3HY9, 3HYG und 3LJT infrage. Die publizierten ADAMTS13-Röntgenkristallstrukturen waren auch bei

diesen Modellierungen als mögliche Templates nicht geeignet, da sie nicht die gesuchte Domäne umfassen. Die Wahl des jeweiligen Templates erfolgte anhand der Analyse der Sequenzidentität zwischen Protein und Röntgenkristallstrukturen, die wie bei der Auswahl des ADAMTS7-Templates über ein multiples Sequenzalignment mit PROMALS3D berechnet wurde. Aus Gründen der Übersichtlichkeit sind die einzelnen Alignments hier nicht aufgeführt. Es konnte für jede zu modellierende ADAMTS-Struktur ein Template mit einer Sequenzidentität von mindestens 30 % gefunden werden. Lagen mehrere mögliche Templates mit identischer, maximaler Sequenzidentität vor, wurden als weitere Kriterien die Auflösung der Röntgenkristallstrukturen sowie der durchschnittliche B-Faktor berücksichtigt. Beinhaltende gewählte Template-Strukturen mehrere Proteinketten wurde jene Kette mit den besseren strukturellen Parametern bestimmt. Mögliche fehlende Aminosäuren sowie eine Evaluation mit ProSa2003 und PROCHECK wurden dann zur Analyse herangezogen. Die Auflösung der ADAMTS4-Röntgenkristallstrukturen ist mit 2,8 Å vergleichsweise niedrig, sodass ein Template mit einer besseren Auflösung möglicherweise ein genaueres Modell liefert. Daher wurde zusätzlich ein Sonderfall berücksichtigt: Handelte es sich bei einer der ADAMTS4-Röntgenkristallstrukturen um jene mit der höchsten Sequenzidentität, wurde die Struktur mit der zweithöchsten Sequenzidentität als Template gewählt (unter der Bedingung: Sequenzidentität $\geq 30\%$). Diese Sonderregel wurde bei ADAMTS12, -13, -17, -18 und -19 angewendet. Für diese fünf Targetstrukturen wurde die ADAMTS1-Röntgenkristallstruktur 2V4B als Template gewählt, welche eine Auflösung von 2,0 Å hat.

Tabelle 4.3: ADAMTS-Proteine (außer ADAMTS7) und genutzte Template-Strukturen für das Homology Modeling sowie die Sequenzidentitäten zwischen Target und Template. * - Für diese Proteine liegen strukturelle Daten der Metalloproteinase-Domäne vor.

Protein	Template (PDB-ID)	Kette	Sequenzidentität
ADAMTS1*	2V4B	B	99 %
ADAMTS2	3LJT (ADAMTS5)	A	30 %
ADAMTS3	3HYG (ADAMTS5)	A	30 %
ADAMTS4*	2RJP	C	99 %
ADAMTS5*	3HYG	A	99 %
ADAMTS6	3HYG (ADAMTS5)	A	40 %
ADAMTS8	2V4B (ADAMTS1)	B	56 %
ADAMTS9	2V4B (ADAMTS1)	B	51 %
ADAMTS10	2V4B (ADAMTS1)	B	39 %
ADAMTS12	2V4B (ADAMTS1)	B	41 %
ADAMTS13	2V4B (ADAMTS1)	B	31 %
ADAMTS14	2V4B (ADAMTS1)	B	31 %
ADAMTS15	2V4B (ADAMTS1)	B	65 %
ADAMTS16	2V4B (ADAMTS1)	B	36 %
ADAMTS17	2V4B (ADAMTS1)	B	35 %
ADAMTS18	2V4B (ADAMTS1)	B	35 %
ADAMTS19	2V4B (ADAMTS1)	B	33 %
ADAMTS20	2V4B (ADAMTS1)	B	47 %

Der Unterschied in der Sequenzidentität zwischen Target und der ADAMTS4-Röntgenkristallstruktur bzw. 2V4B lag jeweils zwischen 0,3 % (ADAMTS12) und 1,1 % (ADAMTS17).

Die für die jeweiligen Modelle genutzten Templates sind in Tabelle 4.3 aufgelistet. Die Templatesuche für ADAMTS3 und ADAMTS6 ergab die höchste Sequenzidentität mit den ADAMTS5-Röntgenkristallstrukturen 3HYG und 3B8Z, welche die gleiche Auflösung haben. Als Template wurde 3HYG (Kette A), analog zur ADAMTS7-Modellierung gewählt (siehe Abschnitt 3.1). Alle weiteren ADAMTS-Modelle der Metalloproteinase-Domäne, außer ADAMTS2, -3 und -6, wurden mit der ADAMTS1-Röntgenkristallstruktur 2V4B (Kette B) als Template berechnet. Von 2V4B ist, analog zur Wahl der repräsentativen Struktur von ADAMTS1, die Kette B aufgrund der strukturellen Parameter (siehe Tabelle 4.1) als Template zur Modellierung herangezogen worden. Die ADAMTS2-Domäne hat 3LJT, eine ADAMTS5-Röntgenkristallstruktur, als Template.

Alle ADAMTS-Modelle wurden einer Modellvalidierung mit PROCHECK unterzogen. In Tabelle 4.4 sind die Ergebnisse des Ramachandran-Plots für jedes Homologiemodell sowie der Bereich, den jedes Modell für das jeweilige Protein abdeckt, aufgeführt. 14 der 15 Modelle zeigen einen geringen Anteil (bis zu

Tabelle 4.4: Modellvalidierung der erstellten ADAMTS-Homologiemodelle anhand des mit PROCHECK erstellten Ramachandran-Plots. Zusätzlich ist die sequenzielle Abdeckung der Modelle angegeben.

Homologiemodell des Proteins	Abdeckung	Ramachandran-Plot			
		bevorzugt erlaubter Bereich	zusätzlich erlaubter Bereich	erlaubter Bereich	nicht erlaubter Bereich
ADAMTS2	Asp263-Pro476	88,8 %	10,1 %	0,5 %	0,5 %
ADAMTS3	Asn254-Pro466	89,4 %	10,1 %	0,0 %	0,5 %
ADAMTS6	Ile248-Asp472	87,4 %	10,6 %	1,0 %	1,0 %
ADAMTS8	Ala218-Glu519	85,8 %	11,8 %	1,2 %	1,2 %
ADAMTS9	Pro292-Pro584	87,4 %	11,9 %	0,4 %	0,4 %
ADAMTS10	Glu238-Phe539	83,0 %	16,3 %	0,4 %	0,4 %
ADAMTS12	Glu245-Pro538	82,5 %	16,7 %	0,4 %	0,4 %
ADAMTS13	Ile79-Thr378	83,8 %	15,4 %	0,8 %	0,0 %
ADAMTS14	Ser258-Gln547	87,0 %	10,9 %	1,6 %	0,4 %
ADAMTS15	Pro217-Arg507	84,2 %	13,8 %	0,8 %	1,2 %
ADAMTS16	Glu289-Gly581	85,3 %	13,1 %	1,2 %	0,4 %
ADAMTS17	Glu231-Pro539	82,6 %	14,4 %	2,2 %	0,7 %
ADAMTS18	Gly292-Gly584	81,2 %	15,6 %	2,4 %	0,8 %
ADAMTS19	Glu324-Pro632	81,8 %	15,3 %	1,8 %	1,1 %
ADAMTS20	Pro258-Thr552	88,2 %	9,8 %	1,6 %	0,4 %

1,2 % bei ADAMTS8 und -15) an Aminosäuren mit Torsionswinkeln in energetisch ungünstigen Regionen. Die wenigen Aminosäuren mit einer möglicherweise nicht nativen Konformation sind jedoch unerheblich für die weiteren Untersuchungen der Bindetasche, da sie nicht in diesem Bereich liegen. Die Validierung mit MetaMQAPII (siehe Tabelle 4.5) zeigte für alle Modelle RMSD-Werte um 4 Å und GDT_TS-Werte > 40. Das heißt, die Modelle weichen vermutlich um rund 4 Å von den nativen Strukturen ab (siehe Abschnitt 2.4). Die Validierung mit ProSa2003, einem Programm, welches die Kompatibilität der Proteinmodelle mit der Aminosäuresequenz untersucht, zeigt in allen erstellten ADAMTS-Modellen keine Auffälligkeiten (siehe Tabelle 4.5). Aus diesem Grund wurden die Ergebnisse als plausibel bewertet, sodass die 15 ADAMTS-Modelle bzw. 19 räumliche ADAMTS-Strukturen für den strukturellen Vergleich zur Verfügung standen.

Die räumlichen Daten der Metalloproteinase-Domänen und insbesondere deren Bindetasche sollten vor allem dazu dienen, Bereiche in der ADAMTS7-Bindetasche zu identifizieren, die Unterschiede aufweisen. Diese Informationen galt es, bei der Suche nach selektiven ADAMTS7-Inhibitoren gezielt zu nutzen. Aus

Tabelle 4.5: Modellvalidierung der erstellten ADAMTS-Homologiemodelle mit MetaMQAPII und ProSa2003.

Homologiemodell des Proteins	MetaMQAPII		Z-Scores (ProSa2003)		
	GDT_TS-Wert	RMSD (Å)	Combined	Pair	Surf
ADAMTS2	45,9	4,1	-6,48	-6,74	-4,25
ADAMTS3	44,2	4,5	-6,30	-4,65	-4,75
ADAMTS6	44,7	4,0	-7,60	-4,27	-6,15
ADAMTS8	46,4	4,1	-9,34	-7,60	-6,55
ADAMTS9	46,5	3,8	-9,09	-7,44	-6,46
ADAMTS10	43,6	4,2	-8,15	-6,10	-5,97
ADAMTS12	44,2	4,0	-7,62	-6,32	-5,54
ADAMTS13	46,4	4,1	-8,27	-6,82	-5,87
ADAMTS14	47,3	3,7	-7,56	6,45	-5,27
ADAMTS15	45,0	4,1	-8,82	-7,98	-6,03
ADAMTS16	45,8	3,9	-8,03	-6,48	-5,96
ADAMTS17	45,3	4,5	-9,07	-6,82	-6,75
ADAMTS18	45,3	4,2	-8,07	-6,16	-6,16
ADAMTS19	44,4	4,5	-8,34	-7,11	-5,87
ADAMTS20	44,6	4,0	-8,69	-7,06	-6,33

diesem Grund war die Modellierung der unbekanntenen ADAMTS-Metalloproteinase-Domänen ein weiterer Schritt auf der Suche nach selektiven ADAMTS7-Inhibitoren.

4.2 Sequenz- und Strukturvergleich

Um sequenzielle und strukturelle Gemeinsamkeiten bzw. Unterschiede der 19 ADAMTS-Proteine zu identifizieren, wurden auf sequenzieller Ebene multiple Alignments mit PROMALS3D [65] (Substitutionsmatrix: BLOSUM62) und auf struktureller Ebene mit Schrödinger Maestro durchgeführt. Die Sequenzidentitäten zwischen den 19 Proteinen wurden im Gesamten sowie für die Metalloproteinase-Domänen und die Binde-taschen mit dem Programm SIAS [161] bestimmt.

Ein Vergleich über die gesamte Sequenzlänge der ADAMTS-Proteine gibt einen ersten Ansatz über den Grad der sequenziellen Konservierung innerhalb dieser Proteinfamilie. Auf die Aufführung des kompletten multiplen Alignments wird hier aus Gründen der Übersichtlichkeit verzichtet. In Tabelle 4.6 sind die paarweisen Sequenzidentitätswerte für die ADAMTS-Proteinfamilie aufgeführt. Die höchste Sequenzidentität besteht mit 60,7 % zwischen ADAMTS6 und -10, die niedrigste mit 17,6 % zwischen ADAMTS13 und -19. Die mittlere Sequenzidentität aller Vergleiche beträgt 29,4 % mit einer Standardabweichung von 7,7 %. Anhand dieser Werte wird deutlich, dass die Sequenzidentität unter den ADAMTS-Proteinen eine hohe Spannweite hat. ADAMTS7 weist über die gesamte Sequenz mit 44,5 % die höchste Sequenzidentität zu ADAMTS12 auf und mit 21,2 % die niedrigste zu ADAMTS13. Strukturvergleiche für die gesamten ADAMTS-Proteine waren nicht möglich, da keine vollständigen experimentellen 3D-Daten für Proteine dieser Familie vorliegen.

4.2.1 Vergleich der Metalloproteinase-Domäne

Bevor die Bindetaschen der ADAMTS-Proteine im Detail untersucht werden sollten, wurden die Metalloproteinase-Domänen hinsichtlich ihrer Sequenz und Struktur verglichen. Dies sollte eine differenziertere Auskunft über den genauen Grad der Konservierung dieses Bereichs ermöglichen.

Tabelle 4.6: Paarweiser Vergleich der Sequenzidentitäten der kompletten ADAMTS-Proteine in Prozent. Die Identitätswerte sind farbkodiert: 100 % identisch - orange, 0 % identisch - weiß.

Sequenzidentität	ADAMTS1	ADAMTS2	ADAMTS3	ADAMTS4	ADAMTS5	ADAMTS6	ADAMTS7	ADAMTS8	ADAMTS9	ADAMTS10	ADAMTS12	ADAMTS13	ADAMTS14	ADAMTS15	ADAMTS16	ADAMTS17	ADAMTS18	ADAMTS19	ADAMTS20
ADAMTS1	100	26,7	26,1	50,1	39,4	28,2	29,4	46,3	35,9	28,1	28,2	21,2	24,7	48,3	28,7	26,6	28,4	25,9	34,7
ADAMTS2	26,7	100	56,3	26,0	26,8	26,5	25,2	28,3	25,4	26,8	22,7	20,3	51,3	29,1	26,4	25,9	26,0	23,4	23,2
ADAMTS3	26,1	56,3	100	25,0	25,8	27,4	24,4	26,9	23,7	26,1	23,1	18,7	54,8	27,6	25,6	24,6	26,2	21,8	22,2
ADAMTS4	50,1	26,0	25,0	100	37,2	30,3	29,7	41,5	37,2	30,2	28,9	21,6	25,0	44,1	28,7	27,2	28,7	27,2	35,6
ADAMTS5	39,4	26,8	25,8	37,2	100	29,1	29,0	39,8	34,4	28,8	27,1	19,7	26,3	39,1	27,4	25,6	26,1	26,1	32,9
ADAMTS6	28,2	26,5	27,4	30,3	29,1	100	35,6	29,0	32,7	60,7	33,9	21,1	26,2	30,3	34,2	28,9	35,2	30,2	29,3
ADAMTS7	29,4	25,2	24,4	29,7	29,0	35,6	100	30,3	24,2	35,4	44,5	21,2	24,5	31,3	31,9	29,8	31,7	28,3	23,2
ADAMTS8	46,3	28,3	26,9	41,5	39,8	29,0	30,3	100	38,1	27,9	27,6	22,2	26,7	45,0	27,4	26,3	27,6	25,9	36,7
ADAMTS9	35,9	25,4	23,7	37,2	34,4	32,7	24,2	38,1	100	31,3	24,3	21,0	24,0	39,6	27,5	25,6	27,8	23,9	43,9
ADAMTS10	28,1	26,8	26,1	30,2	28,8	60,7	35,4	27,9	31,3	100	32,2	21,8	26,0	30,9	32,9	27,8	33,3	27,9	29,7
ADAMTS12	28,2	22,7	23,1	28,9	27,1	33,9	44,5	27,6	24,3	32,2	100	19,6	22,6	28,4	30,8	28,6	31,1	27,8	24,0
ADAMTS13	21,2	20,3	18,7	21,6	19,7	21,1	21,2	22,2	21,0	21,8	19,6	100	19,7	22,5	20,8	22,3	20,4	17,6	22,1
ADAMTS14	24,7	51,3	54,8	25,0	26,3	26,2	24,5	26,7	24,0	26,0	22,6	19,7	100	26,8	25,3	25,7	24,8	22,7	22,3
ADAMTS15	48,3	29,1	27,6	44,1	39,1	30,3	31,3	45,0	39,6	30,9	28,4	22,5	26,8	100	29,7	26,6	29,2	28,1	36,7
ADAMTS16	28,7	26,4	25,6	28,7	27,4	34,2	31,9	27,4	27,5	32,9	30,8	25,3	29,7	100	28,3	55,7	26,6	25,5	
ADAMTS17	26,6	25,9	24,6	27,2	25,6	28,9	29,8	26,3	25,6	27,8	28,6	22,3	25,7	26,6	28,3	100	29,0	54,6	25,8
ADAMTS18	28,4	26,0	26,2	28,7	26,1	35,2	31,7	27,6	27,8	33,3	31,1	20,4	24,8	29,2	55,7	29,0	100	26,7	26,4
ADAMTS19	25,9	23,4	21,8	27,2	26,1	30,2	28,3	25,9	23,9	27,9	27,8	17,6	22,7	28,1	26,6	54,6	26,7	100	24,0
ADAMTS20	34,7	23,2	22,2	35,6	32,9	29,3	23,2	36,7	43,9	29,7	24,0	22,1	22,3	36,7	25,5	25,8	26,4	24,0	100

Der Sequenzvergleich der Metalloproteinase-Domänen könnte außerdem einen ersten Ansatz für nicht konservierte Aminosäuren in der Bindetasche geben, welche mögliche Interaktionspartner für selektiven Liganden darstellen könnten. In Anhang A.2.1 in Abbildung A.5 befindet sich das multiple Sequenzalignment der 19 Metalloproteinase-Domänen der ADAMTS-Proteine. Anhand dieses Alignments und der daraus resultierenden paarweisen Sequenzidentitäten (siehe Tabelle 4.7) können Aussagen über die Gleichartigkeit dieser Domäne getätigt werden. Die paarweisen Sequenzidentitäten der ADAMTS-Metalloproteinase-Domänen liegen zwischen 28,0 % (ADAMTS5 und -13) und 86,8 % (ADAMTS2 und -3). Der durchschnittliche Sequenzidentität beträgt bei diesem Vergleich $41,8\% \pm 11,3\%$. Anhand des Alignments konnte somit ein hoher Konservierungsgrad der Metalloproteinase-Domänen aufgezeigt werden. Folglich ist die Metalloproteinase-Domäne der ADAMTS-Proteine im Vergleich zur gesamten Sequenz stärker konserviert. Die höchste Sequenzidentität zu ADAMTS7 weist erneut ADAMTS12 mit 76,8 % auf. Die niedrigste Sequenzidentität zeigt ADAMTS13 mit 35,7 % zu ADAMTS7.

Die Sequenzanalyse wurde durch einen Strukturvergleich der Metalloproteinase-Domänen ergänzt. Für den Vergleich der experimentellen sowie der modellierten Daten der ADAMTS-Metalloproteinase-Domänen wurden diese mit Schrödinger Maestro alignt und anhand des $RMSD_{C\alpha}$ verglichen. In Tabelle 4.8 sind die paarweisen Vergleichswerte für alle Mitglieder der ADAMTS-Familie aufgeführt. Es sind deutliche strukturelle Unterschiede in der Metalloproteinase-Domäne zu erkennen. Der durchschnittliche $RMSD_{C\alpha}$ liegt bei $8,3 \text{ \AA} \pm 4,1 \text{ \AA}$. Der größte Unterschied zeigt sich zwischen ADAMTS4 und -19 mit $16,7 \text{ \AA}$. Der niedrigste Unterschied konnte zwischen ADAMTS16 und -18 mit $1,1 \text{ \AA}$ beobachtet werden. ADAMTS7 hat eine durchschnittliche Abweichung von $7,3 \text{ \AA} \pm 3,1 \text{ \AA}$ zu den $C\alpha$ -Atomen der 18 ADAMTS-Proteine.

Tabelle 4.7: Paarweiser Vergleich der Sequenzidentitäten der ADAMTS-Metalloproteinase-Domänen in Prozent. Die Identitätswerte sind farbkodiert: 100 % identisch - orange, 0 % identisch - weiß.

Sequenzidentität	ADAMTS1	ADAMTS2	ADAMTS3	ADAMTS4	ADAMTS5	ADAMTS6	ADAMTS7	ADAMTS8	ADAMTS9	ADAMTS10	ADAMTS12	ADAMTS13	ADAMTS14	ADAMTS15	ADAMTS16	ADAMTS17	ADAMTS18	ADAMTS19	ADAMTS20
ADAMTS1	100	31,2	30,7	63,3	59,5	39,5	42,4	62,9	53,6	39,0	41,4	29,5	30,2	66,2	35,4	31,0	35,0	30,5	46,9
ADAMTS2	31,2	100	86,8	31,2	32,7	36,6	37,1	33,7	35,1	36,1	34,6	31,2	79,7	37,6	37,6	32,7	40,0	31,2	34,6
ADAMTS3	30,7	86,8	100	30,7	34,6	36,6	36,1	34,1	35,1	35,1	33,2	31,2	83,2	35,1	35,6	31,2	38,5	30,2	35,6
ADAMTS4	63,3	31,2	30,7	100	49,0	40,8	45,5	55,5	50,2	39,8	42,7	31,9	31,7	61,0	38,3	36,0	38,8	34,1	45,5
ADAMTS5	59,5	32,7	34,6	49,0	100	39,5	43,3	54,3	46,9	37,1	39,0	28,0	33,7	57,6	34,5	34,3	34,0	32,4	43,5
ADAMTS6	39,5	36,6	36,6	40,8	39,5	100	53,1	34,6	51,7	77,6	53,1	34,3	39,1	42,9	48,5	41,6	51,9	42,0	41,6
ADAMTS7	42,4	37,1	36,1	45,5	43,3	53,1	100	39,8	49,8	53,6	76,8	35,7	37,1	49,0	53,9	41,2	53,9	40,3	47,8
ADAMTS8	62,9	33,7	34,1	55,5	54,3	34,6	39,8	100	52,2	34,1	37,9	28,5	34,2	58,6	35,0	34,1	35,4	29,4	47,4
ADAMTS9	53,6	35,1	35,1	50,2	46,9	51,7	49,8	52,2	100	48,3	51,2	29,5	39,1	52,7	45,6	38,2	45,6	33,8	63,3
ADAMTS10	39,0	36,1	35,1	39,8	37,1	77,6	53,6	34,1	48,3	100	52,1	32,9	38,1	42,4	48,1	41,6	50,0	40,6	39,2
ADAMTS12	41,4	34,6	33,2	42,7	39,0	53,1	76,8	37,9	51,2	52,1	100	33,3	35,6	45,7	52,4	42,2	54,4	43,6	47,4
ADAMTS13	29,5	31,2	31,2	31,9	28,0	34,3	35,7	28,5	29,5	32,9	33,3	100	32,2	29,5	33,0	30,9	31,1	28,5	33,8
ADAMTS14	30,2	79,7	83,2	31,7	33,7	39,1	37,1	34,2	39,1	38,1	35,6	32,2	100	34,7	37,1	34,7	36,6	32,2	37,1
ADAMTS15	66,2	37,6	35,1	61,0	57,6	42,9	49,0	58,6	52,7	42,4	45,7	29,5	34,7	100	39,8	36,2	39,3	34,8	45,5
ADAMTS16	35,4	37,6	35,6	38,3	34,5	48,5	53,9	35,0	45,6	48,1	52,4	33,0	37,1	39,8	100	43,2	80,1	45,6	38,3
ADAMTS17	31,0	32,7	31,2	36,0	34,3	41,6	41,2	34,1	38,2	41,6	42,2	30,9	34,7	36,2	43,2	100	46,6	62,9	36,8
ADAMTS18	35,0	40,0	38,5	38,8	34,0	51,9	53,9	35,4	45,6	50,0	54,4	31,1	36,6	39,3	80,1	46,6	100	46,1	40,3
ADAMTS19	30,5	31,2	30,2	34,1	32,4	42,0	40,3	29,4	33,8	40,6	43,6	28,5	32,2	34,8	45,6	62,9	46,1	100	32,5
ADAMTS20	46,9	34,6	35,6	45,5	43,5	41,6	47,8	47,4	63,3	39,2	47,4	33,8	37,1	45,5	38,3	36,8	40,3	32,5	100

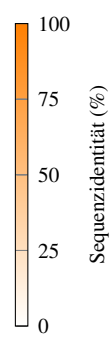
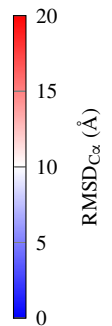


Tabelle 4.8: Strukturvergleich der Metalloproteinase-Domänen der ADAMTS-Proteine anhand des RMSD_{Cα} in Å. Die strukturellen Unterschiede zwischen den Proteinen sind mittels eines Farbverlaufs dargestellt: Von 0 Å Unterschied - blau bis 10 Å Unterschied - weiß bis 20 Å Unterschied - rot. * - Röntgenkristallstruktur.

RMSD _{Cα}	ADAMTS1*	ADAMTS2	ADAMTS3	ADAMTS4*	ADAMTS5*	ADAMTS6	ADAMTS7	ADAMTS8	ADAMTS9	ADAMTS10	ADAMTS12	ADAMTS13	ADAMTS14	ADAMTS15	ADAMTS16	ADAMTS17	ADAMTS18	ADAMTS19	ADAMTS20
ADAMTS1*	0,0	5,1	5,1	7,8	1,3	12,3	5,0	2,8	5,4	12,0	5,1	4,0	7,8	1,4	4,9	14,0	5,0	14,0	2,9
ADAMTS2	5,1	0,0	2,2	8,1	5,2	12,2	6,3	5,7	5,7	12,0	6,3	4,9	5,7	5,2	4,9	13,8	5,0	13,8	5,3
ADAMTS3	5,1	2,2	0,0	7,9	5,2	12,4	6,4	5,9	5,4	12,2	6,5	4,6	5,4	5,3	4,7	14,1	4,8	14,0	5,3
ADAMTS4*	7,8	8,1	7,9	0,0	7,9	15,2	10,1	8,5	4,8	15,1	10,2	7,6	7,5	7,8	7,1	16,3	7,2	16,7	8,1
ADAMTS5*	1,3	5,2	5,2	7,9	0,0	12,2	4,9	3,0	5,6	12,0	5,0	4,3	8,0	1,7	5,1	13,9	5,2	13,9	3,1
ADAMTS6	12,3	12,2	12,4	15,2	12,2	0,0	10,8	11,8	13,8	2,4	10,7	12,3	13,6	12,2	13,1	6,9	13,1	6,6	12,2
ADAMTS7	5,0	6,3	6,4	10,1	4,9	10,8	0,0	4,2	8,0	10,5	2,0	5,6	9,0	4,9	6,6	13,1	6,6	13,0	4,9
ADAMTS8	2,8	5,7	5,9	8,5	3,0	11,8	4,2	0,0	6,1	11,6	4,3	3,9	8,4	2,7	6,0	13,8	5,7	13,8	2,8
ADAMTS9	5,4	5,7	5,4	4,8	5,6	13,8	8,0	6,1	0,0	13,6	8,2	4,9	6,7	5,6	4,6	15,1	4,7	15,3	5,6
ADAMTS10	12,0	12,0	12,2	15,1	12,0	2,4	10,5	11,6	13,6	0,0	10,4	12,1	13,3	11,9	12,9	7,0	12,8	6,7	11,9
ADAMTS12	5,1	6,3	6,5	10,2	5,0	10,7	2,0	4,3	8,2	10,4	0,0	5,6	9,1	5,0	6,6	12,9	6,6	12,9	4,9
ADAMTS13	4,0	4,9	4,6	7,6	4,3	12,3	5,6	3,9	4,9	12,1	5,6	0,0	7,4	4,3	4,1	14,1	4,2	14,1	2,9
ADAMTS14	7,8	5,7	5,4	7,5	8,0	13,6	9,0	8,4	6,7	13,3	9,1	7,4	0,0	7,9	6,8	14,4	6,7	14,3	8,1
ADAMTS15	1,4	5,2	5,3	7,8	1,7	12,2	4,9	2,7	5,6	11,9	5,0	4,3	7,9	0,0	5,0	13,8	5,1	13,9	3,2
ADAMTS16	4,9	4,9	4,7	7,1	5,1	13,1	6,6	6,0	4,6	12,9	6,6	4,1	6,8	5,0	0,0	14,7	1,1	14,8	4,9
ADAMTS17	14,0	13,8	14,1	16,3	13,9	6,9	13,1	13,8	15,1	7,0	12,9	14,1	14,4	13,8	14,7	0,0	14,7	3,3	14,0
ADAMTS18	5,0	5,0	4,8	7,2	5,2	13,1	6,6	5,7	4,7	12,8	6,6	4,2	6,7	5,1	1,1	14,7	0,0	14,8	5,0
ADAMTS19	14,0	13,8	14,0	16,7	13,9	6,6	13,0	13,8	15,3	6,7	12,9	14,1	14,3	13,9	14,8	3,3	14,8	0,0	13,9
ADAMTS20	2,9	5,3	5,3	8,1	3,1	12,2	4,9	2,8	5,6	11,9	4,9	2,9	8,1	3,2	4,9	14,0	5,0	13,9	0,0



Die minimale Distanz zu ADAMTS7 zeigt ADAMTS12 ($\text{RMSD}_{\text{C}\alpha} = 2,0 \text{ \AA}$). Die maximale Differenz zu ADAMTS7 wurde für ADAMTS17 mit einem $\text{RMSD}_{\text{C}\alpha}$ von $13,1 \text{ \AA}$ berechnet. Die strukturellen Unterschiede in den Metalloproteinase-Domänen befinden sich dabei in den Loop-Regionen der Proteine. Die Bereiche mit ausgebildeten Sekundärstrukturen zeigen optisch nur geringe strukturelle Abweichungen (nicht abgebildet).

Der Sequenzvergleich der Metalloproteinase-Domänen bestätigt die hohe Konservierung dieser Domäne (durchschnittlich $41,8 \%$) in der ADAMTS-Familie. Im Gegensatz dazu sind große strukturelle Unterschiede (vergleiche Tabelle 4.8) in dieser Proteindomäne identifiziert worden. Es kann demzufolge nicht von der starken Konservierung auf eine strukturelle Ähnlichkeit geschlossen werden. Die strukturellen Unterschiede der Metalloproteinase-Domänen sind daher ein erster Ansatz, der auf Diversität in diesem Bereich trotz hoher sequenzieller Konservierung hindeutet.

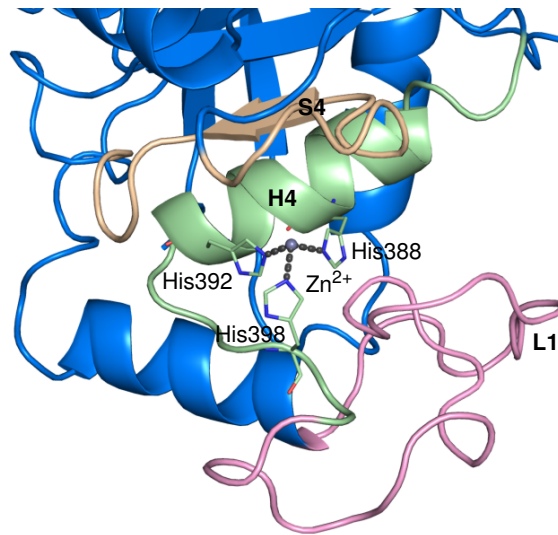
4.2.2 Vergleich der Bindetasche

Aufgrund des hohen Konservierungsgrades der Metalloproteinase-Domänen, stellt sich die Frage, ob die Bindetaschen der ADAMTS-Proteine ebenfalls hoch konserviert sind. Um das zu untersuchen, wurde die Bindetasche in drei Bereiche unterteilt (vergleiche Abschnitt 3.3.1) und diese einzeln betrachtet. Dabei handelt es sich um die Bereiche um das β -Faltblatt S4 (ADAMTS7: Met350-Met365), die α -Helix H4 (ADAMTS7: Leu380-Gly400) und die Loop-Region L1 (ADAMTS7: Ser401-Arg430). Ist im Folgenden von den Regionen bzw. Bereichen um S4, H4 und L1 die Rede, sind diese definierten Teilsequenzen bzw. -strukturen gemeint. In der Abbildung 4.1a sind die Bereiche, welche die Tasche beschreiben, im ADAMTS7-Homologiemodell dargestellt.

Zur Analyse der sequenziellen Konservierung wurde das multiple Alignment über diese Bereiche betrachtet, welches in Abbildung 4.1b dargestellt ist. Zur besseren Übersichtlichkeit sind an dieser Stelle in Tabelle 4.9 nur die Sequenzidentitäten der drei Bindetaschenregionen im Vergleich zu ADAMTS7 aufgeführt. Die vollständigen paarweisen Sequenzidentitäten der ADAMTS-Familie über die drei betrachteten Bereiche sind im Anhang A.2.1 in den Tabellen A.1-A.3 dargestellt.

Anhand des Alignments zeigt sich bereits visuell eine hohe sequenzielle Übereinstimmung in einem Teilbereich der Helix H4 (Ala383-Gly400), welcher auch das Zink-bindende Motiv umfasst. Der Vergleich der Sequenzidentitäten in den Bereichen verdeutlicht diese Feststellung. Die Bereiche um S4 und H4 sind über die ADAMTS-Familie mit einer durchschnittlichen Sequenzidentität von $46,5 \pm 17,9 \%$ bzw. $63,3 \pm 10,9 \%$ stark konserviert. Im Gegensatz dazu ist die Region des L1-Loops mit einer durchschnittlichen Sequenzidentität von $25,2 \pm 12,9 \%$ weniger stark konserviert. Die höchsten Sequenzidentitäten zu ADAMTS7 in den drei betrachteten Regionen weist erneut ADAMTS12 mit $82,4 \%$ (S4-Bereich), $95,2 \%$ (H4-Bereich) bzw. $66,7 \%$ (L1-Loop) auf. Die niedrigste Sequenzidentität zu ADAMTS7 weisen ADAMTS13 im S4-Bereich ($18,8 \%$), ADAMTS8 und ADAMTS14 im H4-Bereich ($52,4 \%$) und ADAMTS1 und ADAMTS4 mit je $16,7 \%$ im L1-Loop auf. Somit stellt die L1-Loop-Region aufgrund der geringsten sequenziellen Konservierung, im Gegensatz zu den Bereichen um H4 und S4, einen möglichen „Angriffspunkt“ für Interaktionsbereiche mit selektiven Inhibitoren dar. Hier bestätigt sich zudem, dass ADAMTS5 zu ADAMTS7 die höchsten Sequenzidentitäten, im Vergleich zu den ADAMTS-Proteinen mit bekannten Röntgenkristallstrukturen, aufweist.

Um einen Einblick in die Struktur der ADAMTS-Bindetaschen zu erhalten, wurde ein Vergleich dieser anhand der bekannten experimentellen ADAMTS-Daten sowie der erstellten ADAMTS-Homologiemodelle



(a) Einteilung der ADAMTS7-Bindetasche

Bereiche	Bereich um S4					Bereich um H4					Loop-Bereich L1					
	350	355	360	365		380	385	390	395	400	405	410	415	420	425	430
ADAMTS1	..QT	CDTLGMADV	CTV			IQA	AFTTAH	ELGH	VENMPH	DDAK	QCASL	NGVNO	.DSHM	MASMLS	N..LDH	SQPW
ADAMTS2	...	SGMQGYAP	VTGM			FSS	AFVVAH	ETGH	VLCMEH	DGQGNR	QGD..EVR	.LGSIM	APLVQA	..AFHR	FHWSR	
ADAMTS3	...	AGMQGYAP	VTGM			FSS	AFVVAH	ETGH	VLCMEH	DGQGNR	QGD..ETA	.MGSVM	APLVQA	..AFHR	YHWSR	
ADAMTS4	..ST	CDTLGMADV	CTV			LQS	AFTAAH	ELGH	VENMLH	DNSKPC	ISLNGP	LSTSRH	VMAVMAH	..VDPE	EPWSP	
ADAMTS5	..HS	CDTLGMADV	CTI			LHA	AFTVAH	ETGH	LLGLSH	DDSKF	CEETFG	STEDKRL	MSSILTS	..IDASK	PWSK	
ADAMTS6	KNKP	CGTLGLAS	VAGM			LGS	AFTVAH	ETGH	NFGMNH	DGIGNS	CGT.KGHE	.AAKL	MAHITA	..NTNP	FSWSA	
ADAMTS7	MNRP	CETLGLSH	VAGM			LPL	AFTVAH	ELGH	SFEGIO	HDSGND	CEP..VGK	.RPFIM	SPOLLY	..DAAP	LTWSR	
ADAMTS8	.EGL	CDTLGVAD	ICTI			LQA	AHTLAH	ELGH	VLSMPH	DDSKP	CTRLFG	PMPG.KHHV	MAPLFVH	..LNQT	LPWSP	
ADAMTS9	.HDK	CDTLGLAEL	CTI			LST	AFTVAH	ELGH	VENMPH	DDNNK	CKEE.GVKS	.POHV	MAPTLNF	..YTNP	WMWSK	
ADAMTS10	KNKP	CGTLGLAP	VGGM			LAT	AFTVAH	ELGH	TFGMNH	DGVNS	CGA.RGQD	.PAKL	MAHITM	..KTNP	FWWS	
ADAMTS12	FNRP	CETLGLSH	LSGM			LPL	AFTVAH	ELGH	SFEGIO	HDKEND	CEP..VGR	.HPYIM	SRQLQY	..DPTP	LTWSK	
ADAMTS13	.DGN	RQVRGVT	QLCGA			FDL	GVTTAH	ELGH	SFELHD	GAPGSG	CGP...SGHV	MASDGAA	..PRAGL	AWSP		
ADAMTS14SGYAP	VTGM			FSS	AFVVAH	ETGH	VLCMEH	DGQGN	QAD..ETS	.LGSVM	APLVQA	..AFHR	FHWSR	
ADAMTS15	..TT	CDTLGMADV	CTM			LPS	AFTVAH	ELGH	VENMPH	DDNVK	VCEEVFG	KLRR.ANHM	SPTLIQ	..IDRAN	PWSA	
ADAMTS16	KNEP	CDTLGFAP	ISGM			LGL	AFTVAH	ELGH	NFGMIH	DGEGNM	CKKS...EGNI	MSP	TLAG	..RNGV	FSWSP	
ADAMTS17	KDEP	CDTVGIAY	LSGM			LNL	AFTVAH	ELGH	NLGMNH	DDHSS	CAG...RSHI	M	SGE	VKGR	NSDLS	
ADAMTS18	KNEP	CDTLGFAP	ISGM			LGL	AFTVAH	ELGH	NFGMIH	DGEGNP	CRKA...EGNI	MSP	TLTG	..NNGV	FSWSS	
ADAMTS19	KDEP	CDTVGIAY	LSGM			LNL	AFTVAH	ELGH	NMGINH	DDNDH	PSCAD...GLHI	M	SGE	WIKG	QNLGD	
ADAMTS20	.KEK	KNMLGLSYL	CTI			LIS	AFTVAH	ELGH	TLGVQH	DDNPR	CKE..MKVT	.KYHV	MAPALSF	..HMS	PWSN	

(b) Alignment der Bindetaschenregion

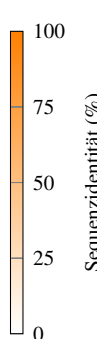
Abbildung 4.1: Zoom in die ADAMTS7-Bindetasche sowie Sequenzalignment der ADAMTS-Bindetaschen. (a) ADAMTS7-Bindetasche dargestellt anhand der drei definierten Regionen, welche farbig markiert sind: Bereich um β -Faltblatt S4 (Met350-Met365) beige, Region um α -Helix H4 (Leu380-Gly400) pastellgrün und Loop-Region L1 (Ser401-Arg430) rosé. Das Protein ist in Cartoon-Darstellung, die Zink-koordinierenden Histidine als Linien veranschaulicht. Das Zink-Ion ist als graue Kugel zu erkennen, Zink-koordinierende Bindungen als grau gestrichelte Linien. (b) Ausschnitt aus dem multiplen Sequenzalignment der Metalloproteinase-Domänen aller ADAMTS-Proteine zur Darstellung des Bereichs der Bindetasche. Konservierte Aminosäuren sind grau hinterlegt. Die Sequenzen der ADAMTS-Domänen stammen aus der UniProt-Datenbank [169]. Die Nummerierung des Alignments sowie die Einteilung der Bereiche über dem Alignment beziehen sich auf ADAMTS7. Das Gleiche gilt für die Markierung der Sekundärstrukturelemente in diesen Bereichen unter dem Alignment.

durchgeführt. Zu diesem Zweck wurden die ADAMTS-Bindetaschen align und anhand ihrer Koordinaten verglichen. Zur besseren Übersichtlichkeit sind die kompletten paarweisen Vergleiche der ADAMTS-Strukturen über die drei Bereiche im Anhang A.2.1 in den Tabellen A.4-A.6 aufgeführt. Die drei strukturellen Analysen der ADAMTS-Proteine im Vergleich zum Target ADAMTS7 sind in Tabelle 4.10 zusammengefasst.

Es zeigt sich, dass der Bereich um die α -Helix H4 ähnlich in allen ADAMTS-Proteinen ist. Der durchschnittliche $RMSD_{C\alpha}$ beträgt $1,3\text{ \AA} \pm 0,9\text{ \AA}$ für diesen Bereich. Der $RMSD_{C\alpha}$ reicht von $0,5\text{ \AA}$ bei AD-

Tabelle 4.9: Paarweiser Vergleich der Sequenzidentitäten der drei Bereiche, welche die Bindetasche definieren, im Vergleich zu ADAMTS7 (in Prozent). Zur Vereinfachung sind die betrachteten Bereiche der ADAMTS7-Bindetasche folgendermaßen abgekürzt: ADAMTS7-S4 (Met350-Met365), ADAMTS7-H4 (Leu380-Gly400) und ADAMTS7-L1 (Ser401-Arg430). Die Identitätswerte sind farbkodiert: 100 % identisch - orange, 0 % identisch - weiß.

	ADAMTS7-S4	ADAMTS7-H4	ADAMTS7-L1
ADAMTS1	40,0	61,9	16,7
ADAMTS2	38,5	57,1	30,0
ADAMTS3	38,5	57,1	26,7
ADAMTS4	35,7	61,9	16,7
ADAMTS5	40,0	61,9	26,7
ADAMTS6	70,6	66,7	23,3
ADAMTS7	100	100	100
ADAMTS8	31,3	52,4	20,0
ADAMTS9	37,5	61,9	26,7
ADAMTS10	64,7	66,7	23,3
ADAMTS12	82,4	95,2	66,7
ADAMTS13	18,8	61,9	17,9
ADAMTS14	50,0	52,4	26,7
ADAMTS15	46,7	66,7	23,3
ADAMTS16	52,9	71,4	35,7
ADAMTS17	35,3	66,7	27,6
ADAMTS18	52,9	71,4	35,7
ADAMTS19	41,2	66,7	24,1
ADAMTS20	37,5	66,7	23,3



A vertical color scale legend on the right side of the table indicates sequence identity percentages. The scale ranges from 0 (white) at the bottom to 100 (dark orange) at the top, with intermediate markers at 25, 50, and 75. The label 'Sequenzidentität (%)' is positioned vertically to the right of the scale.

AMTS8 und ADAMTS15 bis 1,7 Å bei ADAMTS5 und ADAMTS12 (siehe Tabelle A.5). Die hohe strukturelle Ähnlichkeit kann mit der konservierten Sekundärstruktur begründet sein. Hierbei bildet ADAMTS10 jedoch eine Ausnahme: Mit einem $\text{RMSD}_{\text{C}\alpha}$ bis zu 4,1 Å im Vergleich zu ADAMTS16 weicht der Bereich um die Helix H4 deutlich von den anderen ADAMTS-Proteinen ab. Das gleiche Ergebnis spiegelt sich auch wider, wenn nur die ADAMTS7-Vergleiche betrachtet werden: Die maximale strukturelle Abweichung zu ADAMTS7 weist ADAMTS10 mit 3,8 Å auf. Die durchschnittliche Abweichung des Bereiches um H4 zu ADAMTS7 beträgt $1,0 \text{ \AA} \pm 0,7 \text{ \AA}$, was somit mit der hohen sequenziellen Identität im Einklang ist.

Im Gegensatz dazu zeigen die beiden Regionen S4 und L1 innerhalb der Proteinfamilie großer strukturelle Unterschiede. Die durchschnittlichen $\text{RMSD}_{\text{C}\alpha}$ -Werte betragen $6,2 \text{ \AA} \pm 4,3 \text{ \AA}$ für die S4-Region (siehe Tabelle A.4) und $6,2 \text{ \AA} \pm 2,3 \text{ \AA}$ für die Region um den L1-Loop (siehe Tabelle A.6) für alle Vergleiche. Folglich sind die größten, gleich verteilten Unterschiede im L1-Loop-Bereich vorhanden. Werden nur die Vergleiche zu ADAMTS7 betrachtet, zeigen diese eine durchschnittliche strukturelle Abweichung von $5,0 \text{ \AA} \pm 4,0 \text{ \AA}$ für die Region um S4 und $6,1 \text{ \AA} \pm 1,7 \text{ \AA}$ für den L1-Loop. Die strukturellen Unterschiede zu ADAMTS7 bestehen in der Region um S4 vor allem zu ADAMTS2 ($\text{RMSD}_{\text{C}\alpha} = 10,2 \text{ \AA}$), -3 ($\text{RMSD}_{\text{C}\alpha} = 10,5 \text{ \AA}$) und -14 ($\text{RMSD}_{\text{C}\alpha} = 16,8 \text{ \AA}$). Die weiteren paarweisen Vergleiche betragen zwischen 1,3 Å und 6,1 Å. Im Bereich L1 liegen die Vergleichswerte zwischen 3,3 Å (ADAMTS3) und 11,4 Å (ADAMTS4). Somit sind die strukturellen Unterschiede im L1-Loop zu allen ADAMTS-Proteinen ähnlich bzw. ähnlicher als in der S4-Region. Der L1-Loop ist sequenziell weniger stark konserviert als die restliche Bindetasche, was ein Grund für die beobachteten strukturellen Unterschiede sein könnte. Die strukturelle Diversität in dieser Region sollte die Möglichkeit bieten selektive ADAMTS7-Inhibitoren zu entwickeln bzw. Eigenschaften für solche abzuleiten.

Tabelle 4.10: Strukturvergleich der drei Bereiche, welche die Bindetasche definieren, im Vergleich zu ADAMTS7 anhand des $\text{RMSD}_{\text{C}\alpha}$ in Å. Zur Vereinfachung sind die betrachteten Bereiche der ADAMTS7-Bindetasche folgendermaßen abgekürzt: ADAMTS7-S4 (Met350-Met365), ADAMTS7-H4 (Leu380-Gly400) und ADAMTS7-L1 (Ser401-Arg430). Die strukturellen Unterschiede zwischen den Proteinen sind mittels eines Farbverlaufs dargestellt: Von 0 Å Unterschied - blau zu 10 Å Unterschied - weiß bis 20 Å Unterschied - rot. * - Röntgenkristallstruktur.

$\text{RMSD}_{\text{C}\alpha}$	ADAMTS7-S4	ADAMTS7-H4	ADAMTS7-L1
ADAMTS1*	6,0	0,8	6,8
ADAMTS2	10,2	0,7	5,6
ADAMTS3	10,5	0,9	3,3
ADAMTS4*	6,1	1,0	11,4
ADAMTS5*	6,1	1,1	7,3
ADAMTS6	1,3	0,8	6,0
ADAMTS7	0,0	0,0	0,0
ADAMTS8	4,0	0,8	6,0
ADAMTS9	3,8	0,8	5,1
ADAMTS10	2,4	3,8	5,6
ADAMTS12	2,1	1,1	4,3
ADAMTS13	4,0	0,8	5,9
ADAMTS14	16,8	0,9	4,3
ADAMTS15	5,8	0,9	7,5
ADAMTS16	1,7	0,7	6,6
ADAMTS17	1,8	0,8	5,8
ADAMTS18	2,3	0,9	7,1
ADAMTS19	1,7	1,0	6,3
ADAMTS20	4,2	0,7	4,6

Sequenz- sowie Strukturvergleiche der Bindetaschen der ADAMTS-Proteine konnten aufdecken, dass der Bereich um die Helix H4 nicht nur sequenziell hoch konserviert ist, sondern außerdem strukturell geringe Abweichungen aufweist (im Vergleich zu ADAMTS7 maximal 3,8 Å zu ADAMTS10). Dieser Bereich bietet somit wahrscheinlich keinen Angriffspunkt für selektive Eigenschaften. Die Regionen um S4 und L1 zeigen stärkere Abweichungen. Vor allem in der Loop-Region L1 weisen alle ADAMTS-Proteine ähnlich große strukturelle Abweichungen ($\text{RMSD}_{\text{C}\alpha}$) zu ADAMTS7 auf, was auf eine vielversprechende Region im Hinblick auf selektive Interaktionen hindeutet.

4.3 Bindemuster von bekannten Inhibitoren in ADAMTS-Proteinen

Neben dem Wissen um strukturelle Unterschiede könnten bekannte Bindemuster in ADAMTS-Proteinen weitere Aufschlüsse über Eigenschaften von Inhibitoren in dieser Proteinfamilie geben. Daher wurden bekannte Protein-Ligand-Komplexe untersucht, um deren Interaktionen aufzuschlüsseln. Ziel war es, Aufschluss über wichtige nicht selektive Wechselwirkungen sowie die Bindung im Allgemeinen in ADAMTS-Proteinen zu erhalten.

In der ADAMTS Proteinfamilie sind kokristallisierte Strukturen von ADAMTS1, -4 und -5 veröffentlicht. In den Abbildungen 4.2a-4.2j sind die zehn ADAMTS-Röntgenkristallstrukturen mit den jeweiligen Inhibitoren und deren Wasserstoffbrücken dargestellt. In den bekannten Strukturen binden die ADAMTS-Inhibitoren in die Bindetasche der Metalloproteinase-Domäne, in welcher ein katalytisches Zink-Ion über drei Histidine gebunden ist. Außerdem wird deutlich, dass die ADAMTS-Bindetaschen einen „Graben“ bilden, in dem die Inhibitoren an das katalytische Zink-Ion koordinieren. Dies ist in den Abbildungen A.6a-A.6j im Anhang, welche eine Oberflächendarstellung der Bindetaschen beinhalten, veranschaulicht. Die

kokristallisierten Inhibitoren beinhalten entweder eine Carboxygruppe ($-\text{COOH}$) (ADAMTS4) oder die funktionelle Gruppe der Hydroxamsäuren ($-\text{CO}-\text{NHOH}$) als Zink-Bindemotiv.

Eine detaillierte Analyse der drei ADAMTS1-Bindemuster (siehe Abb. 4.2a-4.2c) zeigte, dass in allen eine Wasserstoffbrücke zwischen einem Sauerstoffatom des Inhibitors (2JIH: Carbonylgruppe ($\text{C}=\text{O}$); 3Q2G, 3Q2H: Sulfonylgruppe ($-\text{SO}_2-$)) und Leu370-N (ADAMTS7: Leu357) ausgebildet ist. Zusätzlich liegt bei 2JIH und 3Q2G eine weitere Wasserstoffbrücke zu Gly371-O bzw. Gly371-N (ADAMTS7: Gly358) vor. Leu370 und Gly371 befinden sich in der Faltblattstruktur analog zu ADAMTS7-S4. Bis auf hydrophobe Wechselwirkungen sind in den Strukturen 3Q2G und 3Q2H keine weiteren Interaktionen zwischen Protein und Inhibitor beobachtet worden (vergleiche Abb. A.6b und Abb. A.6c im Anhang). In der Röntgenkristallstruktur 2JIH hingegen liegen weitere Wechselwirkungen vor. Der Inhibitor 097 interagiert mit Ser432-O (ADAMTS7: Pro417) und Leu434-N (ADAMTS7: Leu419) über Wasserstoffbrücken. Ser432 und Leu434 liegen in dem Loop-Bereich, welcher analog zum ADAMTS7-L1-Loop ist. Außerdem bildet Glu402-O_e (ADAMTS7: Glu389) eine Wasserstoffbrücke zur Zink-bindenden Gruppe des Inhibitors 097 aus. Den Inhibitoren der Röntgenkristallstrukturen von 3Q2G und 3Q2H fehlen in diesem Bereich Protonendonoren bzw. -akzeptoren zur Ausbildung dieser Wechselwirkungen.

Der Inhibitor 886 in der ADAMTS4-Röntgenkristallstruktur (siehe Abb. 4.2d) bindet neben dem Zink-Ion über drei Wasserstoffbrücken an die Bindetasche. Die Wasserstoffbrücken wurden zwischen der Carboxygruppe bzw. der Sulfonylgruppe des Inhibitors und Gly331-N, Leu330-N und Gln362-N_e beobachtet. Die drei interagierenden Aminosäuren Gly331 (ADAMTS7: Gly358), Leu330 (ADAMTS7: Leu357) und Gln362 (ADAMTS7: Glu389) befinden sich im β -Faltblatt analog zu ADAMTS7-S4 bzw. der α -Helix analog zu ADAMTS7-H4. Zusätzlich stabilisieren aromatische Wechselwirkungen zwischen Phe357 und dem Inhibitor sowie hydrophobe Wechselwirkungen mit dem ADAMTS7-L1-Loop-Analogon die Bindung (siehe Abb. A.6d im Anhang). Der Inhibitor weist keine Protonendonoren bzw. -akzeptoren in diesem Bereich auf, sodass keine Wasserstoffbrücken in dieser Loop-Region ausgebildet werden.

Die Inhibitoren in den sechs ADAMTS5-Bindemustern (siehe Abb.4.2e-4.2j) zeigen neben der Zink-Koordinierung alle Wasserstoffbrücken zu Leu379-N (ADAMTS7: Leu357), Gly380-O (3B8Z zusätzlich Gly380-N) (ADAMTS7: Gly358) und Glu411-O_e (ADAMTS7: Glu389). Die Interaktionen mit den Aminosäuren Gly380 und Glu389 bilden sich über das Zinkbindemotiv der Inhibitoren aus. Die Wasserstoffbrücke zu Leu379 wurde entweder über eine Sulfonylgruppe oder eine Carbonylgruppe des jeweiligen Inhibitors ausgebildet. Außerdem liegen in allen Strukturen, außer 3B8Z, eine oder mehrere Wasserstoffbrücken zum ADAMTS7-L1-Loop-Analogon. Eine Wasserstoffbrücke zwischen dem Inhibitor und Ser440-O (ADAMTS7: Ser416) wurde in 3HY9 und 3HYG beobachtet. Diese Interaktion zwischen Inhibitor und Ser441-O (ADAMTS7: Pro417) lag in allen ADAMTS5, außer 3B8Z, vor und eine Wasserstoffbrücke zu Leu443-N (ADAMTS7: Leu419) ist in 2RJQ, 3HY7, 3HY9 und 3LJT vorhanden. Beim Inhibitor 294 von 3B8Z fehlt in dieser Loop-Region ein Donor bzw. Akzeptor zur Ausbildung einer Interaktion dieser Art. Die Inhibitoren von 2RJQ und 3HY7 wechselwirken zusätzlich über Asp377-O bzw. Asp377-O_o (ADAMTS7: Glu355) mit der Proteinbindetasche. Den weiteren vier ADAMTS5-Inhibitoren fehlt in diesem Bereich eine entsprechende funktionelle Gruppe zur Ausbildung einer Wasserstoffbrücke.

Insgesamt wurden folgende Gemeinsamkeiten in den ADAMTS-Bindemustern aufgedeckt: Neben der Zink-Koordinierung der Liganden sind im Bereich um das Zink-Ion Wasserstoffbrücken vorhanden, welche die Koordination stabilisieren könnten. Die in diesem Bereich wechselwirkenden Aminosäuren alignen in ADAMTS7 zu Leu357, Gly358 und Glu389. In insgesamt acht von zehn Strukturen der ADAMTS-

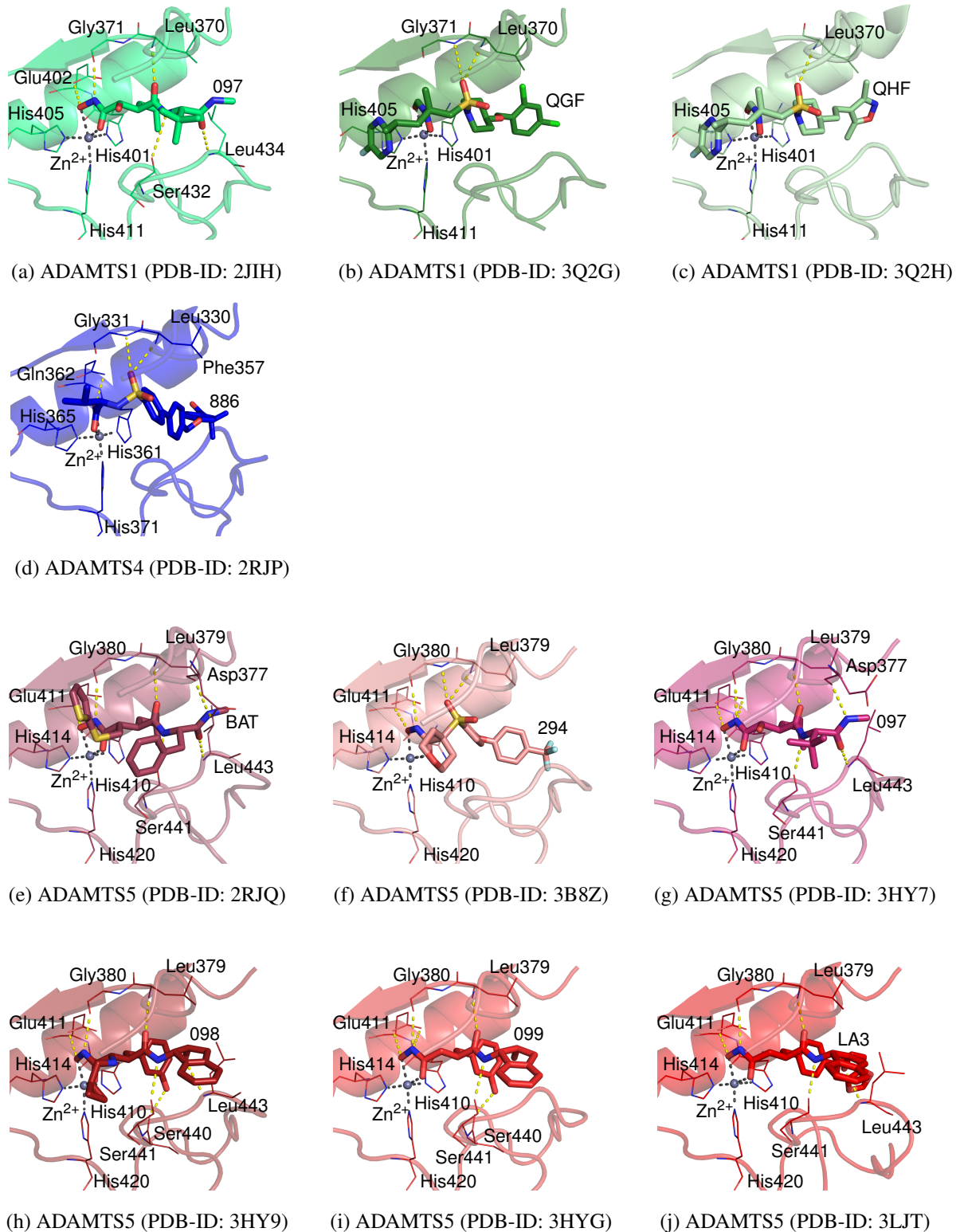


Abbildung 4.2: Bindemuster der bekannten mit Inhibitor kokristallisierten ADAMTS-Röntgenkristallstrukturen. Die Zink-koordinierenden Histidine sind als Linien veranschaulicht, genau wie die jeweils mit dem Inhibitor wechselwirkenden Aminosäuren des Proteins. Die Inhibitoren werden durch Stäbchen dargestellt und die Zink-Ionen sind als graue Kugeln zu erkennen. Zink-koordinierende Bindungen sind durch grau gestrichelte Linien dargestellt, Wasserstoffbrücken durch gelb gestrichelte Linien. Die zu der Bindetasche von ADAMTS7 analogen Bereiche (Regionen um H4, S4 und L1-Loop) sind in transparenter Cartoon-Darstellung visualisiert.

Komplexe findet sich eine Wasserstoffbrücke zur Seitenkette des Glu389-Analogons von ADAMTS7. Ein Blick auf das ADAMTS7-Homologiemodell zeigt, dass in diesem Modell über die Seitenkette der konservierten Aminosäure Glu389 das Zink-Ion koordiniert wird. Somit ist unwahrscheinlich, dass eine zusätzliche Wasserstoffbrücke von diesem Aminosäurerest zu Liganden in ADAMTS7 vorliegen wird. Eine Wasserstoffbrücke zeigen alle Bindemuster zwischen Inhibitor und dem Leu357-N-Analogon von ADAMTS7. Diese Aminosäure ist konserviert in 13 ADAMTS-Proteinen, außer in ADAMTS2, -3, -13, -14, -17 und 19. In sieben der zehn Strukturen, welche alle ADAMTS5-Röntgenkristallstrukturen sowie die ADAMTS1-Röntgenkristallstruktur 2JIH einschließt, liegt eine Wasserstoffbrücke zum Gly358-O-Analogon vor. Alternativ liegt in zwei Strukturen (3Q2G, 2RJP) eine Wasserstoffbrücke zum Gly358-N-Analogon vor. Diese Aminosäure ist über alle ADAMTS-Proteine hinweg sequenziell konserviert. Leu357 und insbesondere Gly358, liegen in geringer Distanz zum Zink-Ion im β -Faltblatt S4 von ADAMTS7. Eine Wasserstoffbrücke könnte eine wichtige Wechselwirkung für die Protein-Ligand-Interaktion neben der Zink-Koordinierung sein.

Die Interaktionen in der zum Zink-Ion abgewandten Seite der Bindetasche sind differenzieller. Hier wechselwirken die ADAMTS7-L1-Loop-Analoga mit einigen Inhibitoren über unterschiedliche Aminosäuren. Im L1-Bereich ist in zwei, fünf bzw. sechs Röntgenkristallstrukturen eine Wasserstoffbrücke zwischen dem ADAMTS7-Ser416-Analogon, dem ADAMTS7-Leu419-Analogon bzw. dem Serin-Rest, welcher zu Pro417 in ADAMTS7 aligns, zum Liganden vorhanden. Die drei Aminosäuren sind über die ADAMTS-Proteine nicht konserviert. Interaktionen in dieser Region könnten somit eine selektive Bindung in den ADAMTS-Proteinen ermöglichen, weil, wie oben bereits beschrieben wurde, dieser Bereich strukturell am wenigsten ähnlich zwischen den Bindetaschen ist. Vier weitere Liganden zeigen keine Interaktion in diesem Bereich über Wasserstoffbrücken, da ihnen entsprechende funktionelle Gruppen fehlen. Bei diesen stabilisieren hydrophobe Wechselwirkungen die Bindung in diesem Bereich.

Die Analyse von bekannten ADAMTS-Inhibitor-Komplexen zeigt, dass neben der Zink-Koordinierung Wasserstoffbrücken in der Umgebung des Zink-Ions scheinbar für eine Stabilisierung dieser Bindung an das Protein sorgen. Außerdem treten häufig Wasserstoffbrücken zwischen Ligand und Protein im Bereich der L1-Loop-Region auf. Dieser Bereich wurde im vorigen Abschnitt bereits als strukturell differenziell in dieser Proteinfamilie identifiziert. Da der L1-Loop-Bereich auch sequenziell weniger stark konserviert ist als die restliche Bindetasche, könnten über Wechselwirkungen mit nicht konservierten Aminosäuren in diesem Bereich die geforderte selektive Bindung von ADAMTS7-Inhibitoren erreicht werden.

4.4 Zusammenfassung

Sequenzielle und strukturelle Untersuchungen der ADAMTS-Familie brachten erste Aufschlüsse über Gemeinsamkeiten und Unterschiede in diesen Proteinen. Der ADAMTS-Sequenzvergleich zeigt über die gesamte Sequenz für 7 der 171 Vergleiche bereits Identitäten über 50 % (z. B. ADAMTS6 und ADAMTS10). In der detaillierten Analyse der Metalloproteinase-Domäne konnten für 33 der 171 Vergleich Sequenzidentitäten über 50 % ausgemacht werden. So hat z. B. die ADAMTS7-Metalloproteinase-Domäne eine Sequenzidentität von 76,8 % zu ADAMTS12. Die detaillierte Untersuchung der Sequenzidentität der ADAMTS-Bindetasche zeigt eine niedrigere Konservierung im Bereich des L1-Loops. Es wurde postuliert, dass dieser Bereich für die selektive Inhibition von ADAMTS7 eine wichtige Rolle spielen könnte.

Es gelang außerdem das Defizit an räumlichen ADAMTS-Strukturen durch Homology Modeling auszugleichen. Somit war es möglich, die 3D-Strukturen der ADAMTS-Proteinfamilie hinsichtlich ihrer Gemeinsamkeiten und Unterschiede zu untersuchen. Der Vergleich der Metalloproteinase-Domänen der ADAMTS-Proteine zeigte erste Unterschiede: Strukturelle Abweichungen von 1,1 Å bis 16,7 Å (Durchschnitt: 8,3 Å) wurden gemessen, wobei innerhalb der ausgebildeten und konservierten Sekundärstrukturen nur geringe Abweichungen vorliegen. Die detaillierte Untersuchung der ADAMTS-Bindetaschen ergab außerdem, dass sich aus dem L1-Loop vermutlich potenzielle Eigenschaften für selektive ADAMTS7-Inhibitoren ableiten lassen. Strukturelle Abweichungen zwischen den ADAMTS-Proteinen wurden hauptsächlich in dieser Region, welche keine definierte Sekundärstruktur aufweist, beobachtet.

Eine zusätzliche Analyse bekannter Protein-Inhibitor-Komplexe der ADAMTS-Familie lieferte neue Erkenntnisse über Eigenschaften von bzw. Wechselwirkungen mit Inhibitoren bzw. der ADAMTS7-Bindetasche. So kann postuliert werden, dass Wasserstoffbrücken in der Umgebung des Zink-Ions die Ligandenbindung stabilisieren könnten und es sich dabei um eine nicht selektive Interaktion, welche in der ADAMTS-Proteinfamilie essenziell sein könnte, handelt. Weiterhin konnten Wechselwirkungen mit nicht konservierten Aminosäuren im Bereich des L1-Loops identifiziert werden. Folglich bestätigte auch diese Analyse, dass sich anhand des L1-Loops vermutlich Eigenschaften für selektive ADAMTS7-Inhibitoren ableiten lassen. Die erhaltenen Ergebnisse ergaben wichtige Informationen und Anhaltspunkte für das im folgenden Kapitel thematisierte Virtual Screening.

5 Ergebnisse und Diskussion: ADAMTS7-Virtual Screening

In den vorangegangenen Kapiteln 3 und 4 wurde die Berechnung einer 3D-Struktur für ADAMTS7 sowie die Identifizierung von sequenziellen und strukturellen Gemeinsamkeiten und Unterschieden innerhalb der ADAMTS-Proteinfamilie dargestellt. Basierend auf diesen Ergebnissen sollte die Anreicherung von potenziellen ADAMTS7-Inhibitoren in einer Datenbank mithilfe eines Virtual Screenings für das Targetprotein untersucht werden. Hierfür wurde anhand des Homologiemodells die ADAMTS7-Bindetasche durch ein Pharmacophore charakterisiert (siehe Abschnitt 5.1). Ein Pharmacophore bietet die Möglichkeit, eine große Anzahl an Liganden hinsichtlich deren Eignung als potenzieller Bindungspartner eines Proteins abzuschätzen bzw. zu filtern. Hier galt es außerdem die Informationen über essenzielle und selektive Interaktionspunkte in der Bindetasche, welche bereits in Kapitel 4 erlangt wurden, einfließen zu lassen.

Des Weiteren sollte für das ADAMTS7-Virtual Screening mit einem Redocking die Güte von verschiedenen Docking-Programmen hinsichtlich der Vorhersage bekannter nativer ADAMTS-Protein-Inhibitor-Komplexe überprüft werden (siehe Abschnitt 5.2). Hierdurch sollte sichergestellt werden, dass das ausgewählte Programm die Bindung an ADAMTS7 prognostizieren kann. Im Anschluss wird die Durchführung des ADAMTS7-Virtual Screenings (Abschnitt 5.3) beschrieben. Hierbei wurden mittels des Pharmacophores und Docking-Studien Filter zur Identifizierung von potenziellen ADAMTS7-Inhibitoren angewandt.

5.1 ADAMTS7-Pharmacophore

Eine Methode zur Charakterisierung von Ligandeneigenschaften, die wichtig für die Interaktion bzw. Bindung in eine Proteinbindetasche sein können, ist die Darstellung mittels eines Wechselwirkungsmodells, welches hier anhand eines Pharmacophores (siehe Abschnitt 2.6) dargestellt wird. Das Pharmacophore wurde in dieser Arbeit anhand der Targetstruktur, demnach strukturbasiert, ermittelt.

Es gibt mehrere Möglichkeiten für die Erstellung eines strukturbasierten Pharmacophores. In dieser Arbeit wurde das ADAMTS7-Pharmacophore nach der *e-Pharmacophore*-Methode von Loving und Salam [133, 134] erstellt (siehe Abschnitt 2.6.3). Für dieses Verfahren werden Docking-Posen von Liganden im Zielprotein benötigt. Da keine publizierten Inhibitoren und damit keine bekannten Protein-Inhibitor-Konformationen existieren, sollte eine Gruppe an Verbindungen gedockt werden, um die potenziellen Interaktionsbereiche in der Bindetasche ausfindig zu machen. Hierfür wird von Schrödinger die Glide Fragmentdatenbank [135], die 667 Fragmente beinhaltet, zur Verfügung gestellt. Die zu dockenden Fragmente wurden mit Schrödinger LigPrep (Version 2.8) vorbereitet und anschließend mit Schrödinger Glide (Version 6.1) [136] (siehe Abschnitt 2.7.3) in die Bindetasche des ADAMTS7-Homologiemodells gedockt. Die Bindetasche wurde vorher anhand eines Grids definiert. Als Zentrum des Grids wurde das erste Histidin (His388) des Zink-Bindemotivs definiert. Die Kantenlänge der Grid-Box betrug 25 Å. Es wurde ein Docking mit *extra precision* (XP) und flexibel gehaltenen Fragmenten analog zum Vorgehen von Loving *et al.* [134]

durchgeführt. Es wurden maximal 100 Posen pro Ligand ausgegeben. Die Posen wurden anschließend einer Minimierung unterzogen und die daraus resultierenden Docking-Posen als Grundlage zur Erstellung des ADAMTS7-Pharmacophores genutzt.

Nach Loving *et al.* [134] wird das Pharmacophore durch die vier bestbewerteten Features sowie ausgenommenem Volumen (gelbe Kugeln) basierend auf der Proteinbindetasche definiert. In Tabelle 5.1 sind die vier Features für ADAMTS7 mit Scores und assoziierten Aminosäuren aufgeführt sowie in Abbildung 5.1a visuell dargestellt. Es handelt sich um die drei Wasserstoffbrückendonor-Features D1, D2 und D3 sowie das Wasserstoffbrückenakzeptor-Feature A1. Feature D3 liegt in räumlicher Nähe zum Zink-Ion in der Bindetasche, die anderen drei Features sind in der L1-Loop-Region lokalisiert.

Tabelle 5.1: Ergebnis der *e-Pharmacophore*-Methode nach Loving und Salam [133, 134] für ADAMTS7. Für jedes Feature sind Rang, Score, Quelle des Scores und die assoziierte Aminosäure angegeben.

Rang	Feature	Score	Quelle für Score	assoziierte Aminosäure
1	D1	-2,17	Wasserstoffbrücke	Ser416 (Proteinrückgrat)
2	A1	-1,69	Wasserstoffbrücke	Leu420 (Proteinrückgrat)
3	D2	-1,49	Wasserstoffbrücke	Asp422 (Seitenkette)
4	D3	-0,91	Wasserstoffbrücke	Glu389 (Seitenkette)

Mit der *e-Pharmacophore*-Methode wird das Volumen der Bindetasche über sogenannte ausgenommenen Volumen [134] im Pharmacophore charakterisiert. Diese dienen dazu, die räumliche Ausdehnung der Bindetasche zu definieren. Die ausgenommenen Volumen werden über Proteinatome lokalisiert, die in einer 20 Å großen Umgebung bzw. Tasche um die 15 räumlich am weitesten entfernten Fragmente bzw. Posen liegen. Der Radius dieser Volumen wird festgelegt als die Hälfte des van der Waals-Radius des jeweiligen Atoms. Da das Pharmacophore anhand des ADAMTS7-Homologiemodells erstellt wurde, fand bisher das Sampling des S1'-Loops (siehe Abschnitt 3.4) keine Berücksichtigung in der Charakterisierung der Bindetasche. Aus diesem Grund wurde das ADAMTS7-Pharmacophore manuell angepasst. Die vier weiteren Konformationen des S1'-Loops wurden über die ausgenommenen Volumen gelegt und in dem Fall, dass eine Konformation mehr Raum in der Bindetasche hat, wurden die entsprechenden Volumen entfernt. In der Abbildung 5.1b ist diese manuelle Anpassung zu erkennen. Diese Anpassung erlaubt die Berücksichtigung der Flexibilität der Bindetasche im ADAMTS7-Pharmacophore.

5.1.1 Evaluierung der ADAMTS7-Pharmacophore-Features

Zur Findung von selektiven Inhibitoren sollten die vier Features des ADAMTS7-Pharmacophores möglichst Wechselwirkungen bzw. Eigenschaften selektiver Liganden widerspiegeln. Aus diesem Grund wurde untersucht, ob die zu den Features korrespondierenden Aminosäuren in der ADAMTS7-Bindetasche über die ADAMTS-Proteine konserviert sind. Eine nicht konservierte Aminosäure wäre ein Hinweis auf eine potenziell selektive Interaktionsmöglichkeit. Da bereits in den vorigen Analysen eine hohe sequenzielle Konservierung beobachtet wurde, sollten die Features auch räumlich im Vergleich zu den ADAMTS-Proteinen betrachtet werden.

Der Wasserstoffbrückendonor D1 liegt in räumlicher Nähe zum Sauerstoffatom des Proteinrückgrats der Aminosäure Ser416. D2 befindet sich nahe der der Carboxygruppe der Asp422-Seitenkette. Das dritte Wasserstoffbrückendonor-Feature D3 liegt in räumlicher Nähe zur Carboxygruppe in der Glu389-Seitenkette. Das Wasserstoffbrückenakzeptor-Feature A1 ist nahe dem Stickstoffatom des Proteinrückgrats der

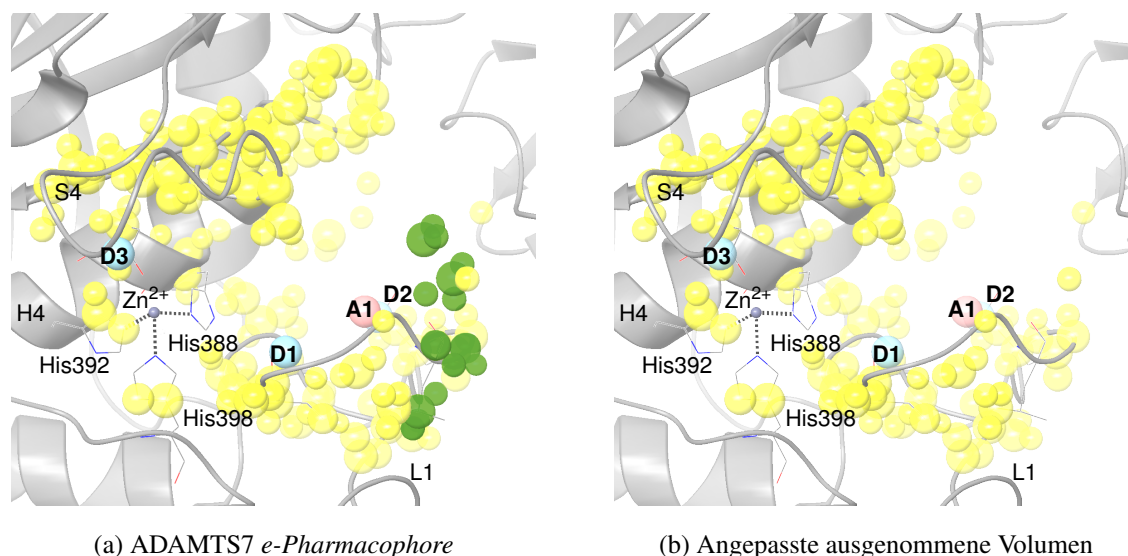


Abbildung 5.1: Ergebnis der *e-Pharmacophore*-Methode für ADAMTS7. (a) ADAMTS7-Bindetasche (grau, Cartoondarstellung) sowie die vier am besten bewerteten Features (fettgedruckt) der angewendeten Methode. Die Wasserstoffbrückendonator-Features sind als blaue Kugeln dargestellt, das Akzeptor-Feature als rote Kugel. Die gelben und grünen Kugeln stellen das ausgenommene Volumen, die Begrenzung der Tasche, dar. Die grünen Kugeln sind dabei die Volumenelemente, die mit Konformationen des S1'-Loops kollidieren. (b) Um die Flexibilität des S1'-Loops zu berücksichtigen, wurden die ausgenommenen Volumenelemente entsprechend angepasst und die grünen Kugeln entfernt.

Aminosäure Leu420 lokalisiert. Die Aminosäuren Leu420, Ser416 und Asp422, somit die assoziierten Aminosäuren der Features A1, D1 und D2, sind in der ADAMTS-Familie nicht konserviert (siehe Anhang A.2.1, Abb. A.5). D3 ist das einzige der vier Features, dessen assoziierte Aminosäure über die ADAMTS-Proteinfamilie sequenziell konserviert ist. Daraus folgt, dass die drei Features D1, D2 und A1 in der ADAMTS7-Bindetasche Bereiche als Interaktionspunkte definieren, welche eine Selektivität aufweisen könnten. Gleichzeitig liegen diese drei Features strukturell im Bereich des L1-Loops, welcher in den vorigen Analysen (vergleiche Abschnitt 4.2) strukturelle Unterschiede zwischen den ADAMTS-Proteinen zeigte. Glu389 hingegen liegt in der Helix H4, welche über die ADAMTS-Proteinfamilie nicht nur sequenziell konserviert sondern auch strukturell ähnlich ist.

Zwei der drei Features (D1 und A1), deren assoziierte Aminosäuren nicht konserviert sind, leiten sich aus Interaktionen mit dem Proteinrückgrat ab, welche somit unabhängig von der sequenziellen Konservierung sind. Daher wurde im Folgenden anhand der räumlichen Strukturen der ADAMTS-Bindetaschen untersucht, ob an der Position der vier Features in den anderen ADAMTS-Proteinen die gleichen Eigenschaften vorliegen. Hierfür wurden die Röntgenkristallstrukturen bzw. die ADAMTS-Homologiemodelle aus Abschnitt 4.1 genutzt. Die ADAMTS-Bindetaschen wurden zu diesem Zweck strukturell und visuell mit Schrödinger SiteMap (Version 2.9) [125, 126] charakterisiert. In Abbildung 5.2 ist die SiteMap beispielhaft für ADAMTS7 dargestellt.

Die erstellten SiteMaps der ADAMTS-Proteine sowie das ADAMTS7-Pharmacophore wurde basierend auf den Proteinstrukturen aligniert. So konnte verglichen werden, ob die vier ADAMTS7-Pharmacophore-Features in den anderen ADAMTS-Proteinen an den gleichen Stellen vorliegen. Dieser Vergleich zeigte für das Wasserstoffbrückendonator-Feature D2, dass nur noch in der ADAMTS14-Bindestelle in dieser Umgebung eine Donor-Region vorliegt. Das Wasserstoffbrückenakzeptor-Feature A1 kommt ebenfalls nur noch in einer anderen ADAMTS-Bindestelle vor, nämlich ADAMTS6. Im Gegensatz dazu ist das Feature D3 an

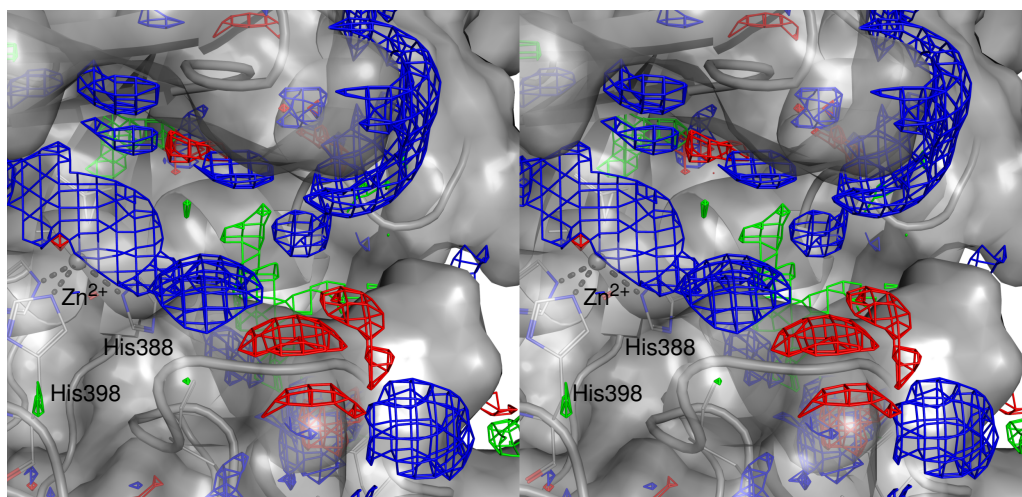


Abbildung 5.2: Beispielfähige *cross-eye*-Darstellung der SiteMap für ADAMTS7. Protein (grau) in Cartoon- mit Oberflächendarstellung. Zink-koordinierende Histidine sind als Linien, Zink-Ion als dunkelgraue Kugel und Zink-koordinierende Bindungen durch grau gestrichelte Linien dargestellt. Rote Fläche - Wasserstoffbrückenakzeptor-Region, blaue Fläche - Wasserstoffbrückendonor-Region, grüne Fläche - hydrophobe Region.

seiner Position in allen ADAMTS-Proteinen zu finden. Das Feature D1 liegt an dieser Stelle ebenfalls in den Bindetaschen von ADAMTS1, -2, -3, -5, -6, -8, -13, -15, -16 und -18 vor.

Zusammengefasst sind die assoziierten Aminosäuren der Features D1, D2 und A1 nicht konserviert und liegen in einer strukturell diversen Region. Im Gegensatz dazu ist die assoziierte Aminosäure von D3 konserviert und auch strukturell findet sich diese Eigenschaft in allen ADAMTS-Bindetaschen. Asp422, die mit D2 assoziierte Aminosäure, ist verglichen mit den alignierten Aminosäuren, die hydrophob oder größer sind, hydrophil. ADAMTS-Aminosäuren, die mit Leu420 (A1) aligniert sind, sind kleiner und/oder polar. Die mit D1 assoziierte Aminosäure ist nicht konserviert, jedoch ließ sich strukturell zeigen, dass eine Donorstelle in zehn ADAMTS-Proteinen in diesem Bereich vorhanden ist. Räumlich konnte zudem nachgewiesen werden, dass die Features D2 und A1 an ihrer Position in nur jeweils einem anderen ADAMTS-Protein und in der strukturell diversen L1-Loop-Region in der ADAMTS7-Bindetasche vorliegen. Diese Beobachtungen deuten darauf hin, dass die Features A1 und D2 in dem ADAMTS7-Pharmacophore Interaktionspunkte beschreiben, die von großer Bedeutung für die Selektivität von ADAMTS7-Liganden sein könnten.

5.1.2 Erweiterung des ADAMTS7-Pharmacophores

Ein Schwachpunkt der *e-Pharmacophore*-Methode bzw. von Schrödinger Phase ist, dass die Koordinierung von Metall-Ionen bzw. entsprechende Features nicht detektiert werden [131]. Daher sollten Eigenschaften für die Zink-Koordinierung in der ADAMTS7-Bindetasche manuell in das Pharmacophore hinzugefügt werden. Außerdem galt es, den Erkenntnisse aus der Analyse der bekannten ADAMTS-Bindemuster (vergleiche Abschnitt 4.3) Rechnung zu tragen.

Zur Charakterisierung der Wechselwirkungen zum Zink-Ion wurden in das ADAMTS7-Pharmacophore basierend auf dem Liganden des ADAMTS7-Templates (PDB-ID: 3HYG (ADAMTS5)) Features hinzugefügt. Hierfür ist die Röntgenkristallstruktur mit dem ADAMTS7-Modell strukturell aligniert worden und die Features der Zink-bindenden Gruppe des Liganden dem 4-Punkt-Pharmacophore hinzugefügt worden. Infolgedessen fand eine Erweiterung des ADAMTS7-Pharmacophores um zwei weitere Features statt. Dabei handelt es sich um die zwei Akzeptor-Features A2 und A3 (siehe Abb. 5.3a). Denn ein spezielles Metall-

Interaktions-Feature stellt das Programm Schrödinger Phase, welches für die Erstellung der Pharmacophore-Hypothesen genutzt wurde, nicht zur Verfügung.

Eine weitere Anpassung wurde zusätzlich am ADAMTS7-Pharmacophore vorgenommen: Wie bereits erwähnt (siehe Abschnitt 4.3) haben 70 % bzw. 20 % der bekannten Röntgenkristallstrukturen, welche mit einem Inhibitor kokristallisiert sind, eine Wasserstoffbrücke zwischen dem jeweiligen Inhibitor und dem Sauerstoff- bzw. Stickstoffatom des Aminosäurerests, welcher in ADAMTS7 zu Gly358 alignt. Im Alignment (siehe Anhang A.2.1, Abb. A.5) zeigt sich, dass Gly358 über alle ADAMTS-Proteine sequenziell konserviert ist. Aktuell gibt es keine wissenschaftliche Literatur, die darüber Aufschluss gibt, ob diese Interaktion notwendig für die Bindung an die ADAMTS-Bindestelle ist oder ob diese eine nicht spezifische Inhibition zur Folge haben könnte. Dies ist eine Fragestellung, die anhand weiterführender Experimente geklärt werden muss. Trotzdem besteht die Annahme, dass eine Wasserstoffbrücke zwischen Ligand und Gly358 essenziell für die Protein-Inhibitor-Interaktion in der ADAMTS-Proteinfamilie sein könnte. Aus diesem Grund wurde dem ADAMTS7-Pharmacophore ein weiteres Feature hinzugefügt. Dabei handelt es sich um das Wasserstoffbrückendonator-Feature D4. Das Feature D4 liegt in räumlicher Nähe zu Gly358-O in der β -Faltblattstruktur S4 (siehe Abb. 5.3b).

5.1.3 Beschreibung des ADAMTS7-Pharmacophore-Modells

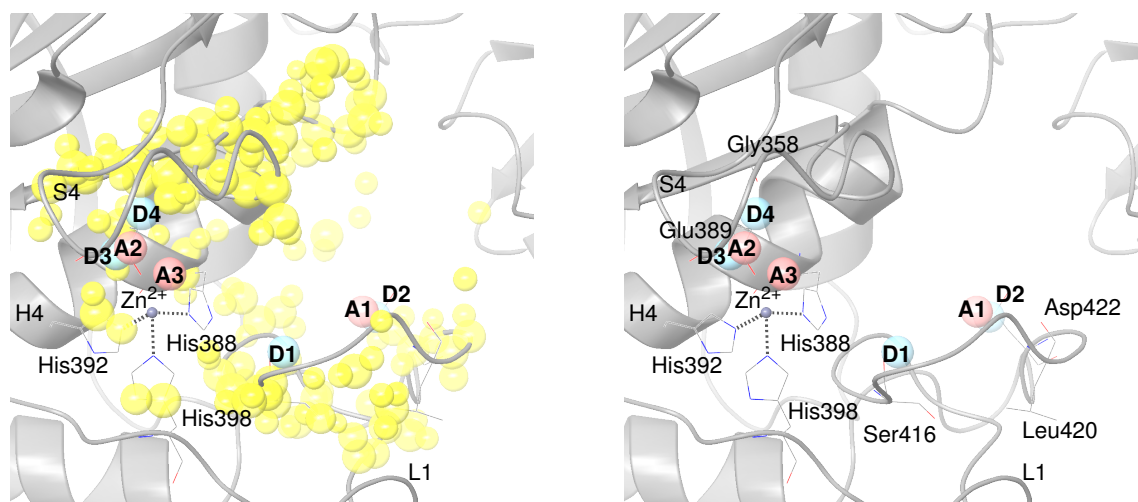
Final besteht das ADAMTS7-Pharmacophore aus sieben Features (siehe Abb. 5.3c). Diese stellen mögliche wichtige chemische Eigenschaften potenzieller ADAMTS7-Inhibitoren dar. Auf Basis der genutzten *e-Pharmacophore*-Methode wurden die besten vier Features selektiert sowie drei manuell erstellte Features hinzugefügt. Die Features A2, A3, D3 und D4 liegen in der Umgebung des Zink-Ions, die weiteren drei Features A1, D1 und D2 auf der Zink-abgewandten Seite im L1-Loop. Die Features A2 und A3 stehen für die Zink-Koordination, A1 und D2 möglicherweise für die Selektivität und D4 für die potenziell essenzielle Interaktion zur Aminosäure Gly358.

Die Radien der Pharmacophore-Features betragen 2 Å. Die Begrenzung der Bindetasche ist in dem Pharmacophore mittels ausgenommenen Volumen definiert worden. Die Radien dieser Volumen betragen die Hälfte der van der Waals-Radien der assoziierten Atome. Ausgewählte Distanzen zwischen den Pharmacophore-Features sind in Abbildung 5.3c dargestellt.

Yang [130] postuliert in seiner Arbeit, dass mehr als sieben Features in einem Pharmacophore ungeeignet wären bzw. ein Zuviel an Features ein Virtual Screening zu stark einschränken könnte. Somit bietet das hier erstellte ADAMTS7-Pharmacophore eine Ausgangsbasis zur Filterung von potenziellen Inhibitoren, welche in einem Virtual Screening angewendet werden sollte, um für die Inhibition von ADAMTS7 passende Liganden zu finden.

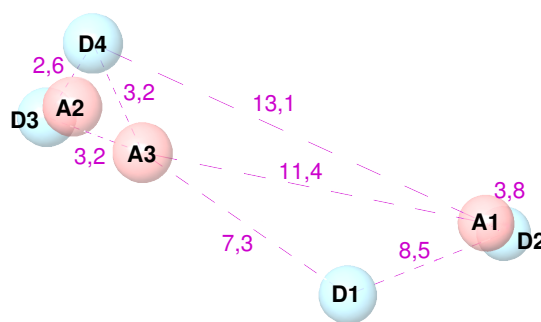
5.2 Redocking zur Evaluierung von Dockingmethoden

Im Abschnitt 2.7 wurde bereits erläutert, dass Redocking eine Methode zur Evaluierung verschiedener Algorithmen hinsichtlich deren Güte bezüglich des Auffindens der nativen Protein-Ligand-Konformation und deren Interaktionen ist. In dieser Arbeit sollte das Redocking dazu dienen die vier zur Verfügung stehenden Docking-Programme und deren Scoring-Funktionen im Hinblick auf ADAMTS-Proteine zu vergleichen. Es sollte jenes Programm gefunden werden, welches die bekannten ADAMTS-Protein-Inhibitor-



(a) ADAMTS7-Pharmacophore in der Bindetasche mit ausgenommenen Volumen

(b) ADAMTS7-Pharmacophore in der Bindetasche



(c) Finales ADAMTS7-Pharmacophore

Abbildung 5.3: Finales ADAMTS7-Pharmacophore. (a) Sieben Features (fettgedruckt) plus ausgenommene Volumen. Die Wasserstoffbrückendonoren-Features sind als blaue Kugeln dargestellt, das Akzeptor-Feature als rote Kugel. Die gelben Kugeln stellen das ausgenommene Volumen dar. (b) ADAMTS7-Pharmacophore ohne die ausgenommenen Volumen, aber mit den zu den jeweiligen Features (fettgedruckt) assoziierten Aminosäuren in Linien-Darstellung. (c) Die sieben ADAMTS7-Pharmacophore-Features sowie Darstellung ausgewählter Distanzen (in Å) zwischen den Features als violett gestrichelte Linien.

Konformationen korrekt vorhersagen kann und sich somit wahrscheinlich am Besten für die Abschätzung der Bindung in ADAMTS-Protein eignet.

Zum Zeitpunkt der Arbeit waren die folgenden zehn Protein-Inhibitor-Komplexe (vergleiche Tabelle 3.1) bekannt:

- die ADAMTS1-Strukturen 2JIH, 3Q2G und 3Q2H,
- die ADAMTS4-Struktur 2RJP,
- die ADAMTS5-Strukturen 2RJQ, 3B8Z, 3HY7, 3HY9, 3HYG und 3LJT.

Diese wurden für das Redocking genutzt. Gedockt wurde immer in die Proteinkette A. Die Röntgenkristallstrukturen wurden für die folgenden Analysen mit dem Schrödinger Protein Preparation Wizard (Schrödinger Suite 2013-2; Epik Version 2.5; Impact Version 6.0; Prime Version 3.5) [170] folgendermaßen vorbereitet: Wassermoleküle wurden entfernt, fehlende Seitenketten wurden hinzugefügt bzw. modelliert, die Bestimmung von Wasserstoffbrücken wurde optimiert sowie eine eingeschränkte (engl. *restrained*) Mini-

mierung [170] durchgeführt. Während einer eingeschränkten Minimierung werden die Wasserstoffatome frei minimiert, wobei die Bewegung von schweren Atomen soweit erlaubt ist, dass nicht native Bindungen, Winkel und Zusammenstöße vermieden werden. Dies dient dazu kleine Fehler in den Röntgenkristallstrukturen vor dem Redocking zu korrigieren.

Folgende Docking-Programme (siehe Abschnitt 2.7) standen zur Verfügung und wurden getestet:

- AutoDock4 [139],
- AutoDockVina [140],
- PLANTS [141] mit den Scoring-Funktionen $\text{PLANTS}_{\text{CHEMPLP}}$ sowie $\text{PLANTS}_{\text{PLP}}$ und
- Schrödinger Glide [136].

Das Zentrum der Bindetaschen der Proteine wurden in allen Programmen durch den jeweils ersten Histidin-Rest des Zink-Bindemotivs (ADAMTS1: His401, ADAMTS4: His361, ADAMTS5: His410) definiert. Parameter für die Grid-Erstellung und das Redocking mit AutoDock4 sowie AutoDockVina sind im Anhang A.3.1 aufgeführt, genau wie das genutzte Skript zum Docken mit PLANTS. Für Schrödinger Glide wurden die Standardparameter genutzt. Im Rahmen der Redocking-Prozedur wurden die jeweiligen Liganden in die zehn Kristallstrukturen der bekannten ADAMTS-Proteinstrukturen gedockt. Die Evaluierung der vier Programme erfolgte anhand des RMSD (siehe Abschnitt 2.5.4) zwischen der nativen Pose des Inhibitor in der Röntgenkristallstruktur im Vergleich zur basierend auf der jeweiligen Scoring-Funktion bestbewerteten Docking-Pose. Die Ergebnisse des Redockings sind in Abbildung 5.4 grafisch sowie im Anhang A.3.1 in Tabelle A.7 in numerischer Form aufgeführt.

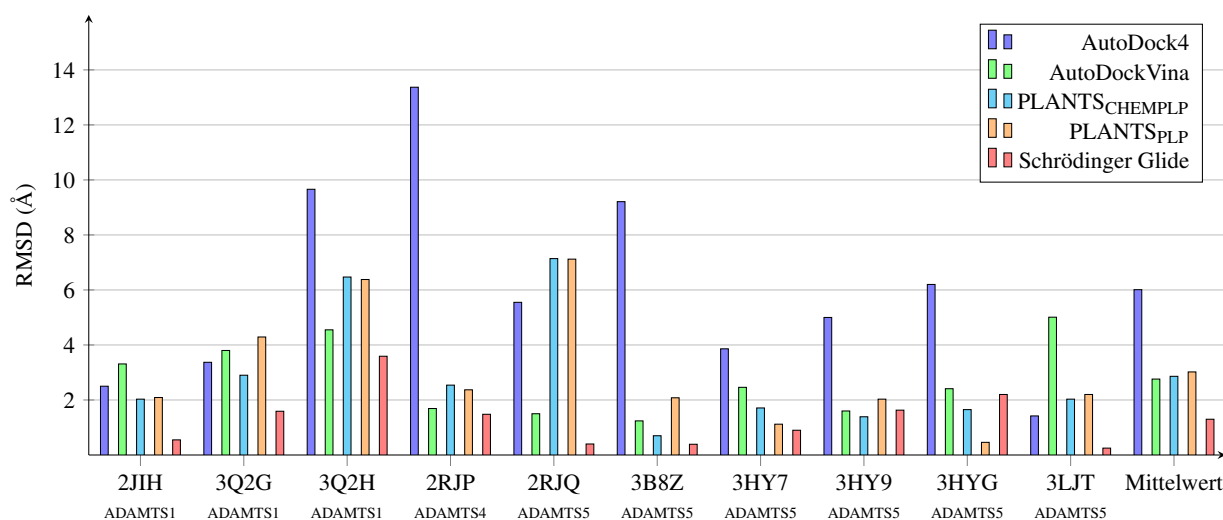


Abbildung 5.4: Redockingergebnisse für die jeweiligen ADAMTS-Protein-Inhibitor-Komplexe. Für jeden Komplex ist der RMSD zwischen nativer und gedockter Pose angegeben sowie die Mittelwerte über alle Komplexe für das jeweilige Programm.

Die Docking-Studie ergab die höchsten RMSD-Werte (zwischen 3,9 Å (3HY7) und 13,4 Å (2RJP)) der getesteten Programme bei sechs Strukturen (3Q2H, 2RJP, 3HYG, 3HY7, 3HY9, 3B8Z) für AutoDock4. AutoDockVina zeigte beim Redocking von zwei der zehn Röntgenkristallstrukturen (2JIH (RMSD = 3,3 Å), 3LJT (RMSD = 5,0 Å)) im Vergleich die schlechtesten Ergebnisse. PLANTS mit der Scoring-Funktion $\text{PLANTS}_{\text{CHEMPLP}}$ dockte den Inhibitor von 2RJQ (RMSD = 7,1 Å) am ungenausten und das Redocking von 3Q2G mit PLANTS und der Scoring-Funktion $\text{PLANTS}_{\text{PLP}}$ zeigte im Vergleich das schlechteste Ergebnis (RMSD = 4,3 Å). Glide hingegen zeigt bei keinem der zehn Redockings das schlechteste Ergebnis. Insgesamt wurde mit AutoDock4 das schlechteste Docking-Ergebnis ($\overline{\text{RMSD}} = 6,0 \text{ Å}$) beim Redocking für die

ADAMTS-Proteine von den genutzten Docking-Programmen beobachtet. Genauer offenbarte sich, dass der Docking-Algorithmus von AutoDock4 im Mittel etwa um das 5-fach schlechtere Ergebnisse als der Schrödinger Glide-Algorithmus, welcher durchschnittlich das beste Ergebnis zeigte ($\overline{\text{RMSD}} = 1,3 \text{ \AA}$), erreichte. Die Dockings mit PLANTS und den Scoring-Funktionen $\text{PLANTS}_{\text{CHEMPLP}}$ und $\text{PLANTS}_{\text{PLP}}$ sowie AutoDockVina ergaben im Durchschnitt Abweichungen von rund 3 \AA zur nativen Pose. Somit scheint AutoDock4 nicht die (annähernd) korrekte Pose von ADAMTS-Inhibitoren abschätzen zu können. Inhibitor-Posen sind mit diesem Programm unter anderem um 180° gedreht gedockt (2RJP). Im Gegensatz dazu dockt Schrödinger Glide acht (3Q2G, 3Q2H, 2JIH, 2RJP, 3LJT, 3HY7, 3B8Z, 2RJQ) der zehn getesteten ADAMTS-Inhibitoren mit dem niedrigsten RMSD-Wert (zwischen $0,3 \text{ \AA}$ (2LJT) und $3,6 \text{ \AA}$ (3Q2H)) im Vergleich zu den anderen Programmen. PLANTS erreichte die genaueste Vorhersage für den 3HYG- (RMSD = $0,5 \text{ \AA}$) und den 3HY9-Komplex (RMSD = $1,4 \text{ \AA}$), jeweils einmal mit der Scoring-Funktion $\text{PLANTS}_{\text{CHEMPLP}}$ und einmal mit der Scoring-Funktion $\text{PLANTS}_{\text{PLP}}$.

Die Redockingergebnisse zeigen innerhalb der bekannten, genutzten Komplexe keine Abhängigkeit von der Auflösung der Struktur. Zwar ist der maximale RMSD von $13,4 \text{ \AA}$ beim Redocking mit AutoDock4 für die Struktur 2RJP, welche die niedrigste Auflösung ($2,8 \text{ \AA}$) aller getesteten Strukturen aufweist, ermittelt worden. Jedoch betragen die Abweichungen zwischen nativer Struktur und Docking-Pose für die vier weiteren genutzten Programme im Durchschnitt $2,0 \text{ \AA}$ für diesen Protein-Inhibitor-Komplex. Dieser Wert ist vergleichbar mit dem Redocking von der Struktur mit der höchsten Auflösung (3HYG, Auflösung: $1,4 \text{ \AA}$), welches eine durchschnittliche Abweichung von $1,7 \text{ \AA}$ zeigt. Auch die Anzahl an rotierbaren Bindungen (11) oder die Größe des Liganden von 2RJP (vergleiche Abb. 4.2) scheint keine Erklärung für das schlechte AutoDock4-Docking dieser Röntgenkristallstruktur zu sein. Denn 2RJQ, 3Q2G und 3Q2H zeigen deutlich bessere Ergebnisse beim Docking mit AutoDock4 bei vergleichbaren strukturellen Parametern (Anzahl rotierbarer Bindungen: 12, 10, 10). Folglich hat die Auflösung der Röntgenkristallstrukturen sowie die Größe und Flexibilität der Liganden in diesen Versuchen keinen belegbaren Einfluss auf die Genauigkeit der Dockingergebnisse.

Aus dem ADAMTS-Redocking resultierte, dass der Schrödinger Glide-Docking-Algorithmus für die bekannten ADAMTS-Komplexe bei acht von zehn Vorhersagen die geringste Abweichung zur nativen Konformation zeigte. Die am besten bewerteten Docking-Posen von Schrödinger Glide sind nahezu gleich der nativen Protein-Inhibitor-Konformationen. Daher kann davon ausgegangen werden, dass der Glide-Algorithmus für ähnliche Strukturen gute bis sehr gute Ergebnisse liefern wird und die Bindung zwischen ADAMTS-Proteinen und Inhibitoren potenziell am besten vorhersagen kann. Aus diesem Grund wurde im Folgenden für das ADAMTS7-Virtual Screening bzw. die weiteren Docking-Studien im Rahmen dieser Arbeit Schrödinger Glide genutzt.

5.3 Virtual Screening

Nach der Erstellung des Pharmacophores und der Beurteilung der Docking-Programme bestand das nächste Ziel in der Suche und Identifizierung von potenziellen ADAMTS7-Inhibitoren aus einer großen Datenbank. Die gesamte Virtual Screening-Pipeline ist in der Abbildung 5.5 dargestellt. Im Folgenden werden die einzelnen Schritte genauer erläutert.

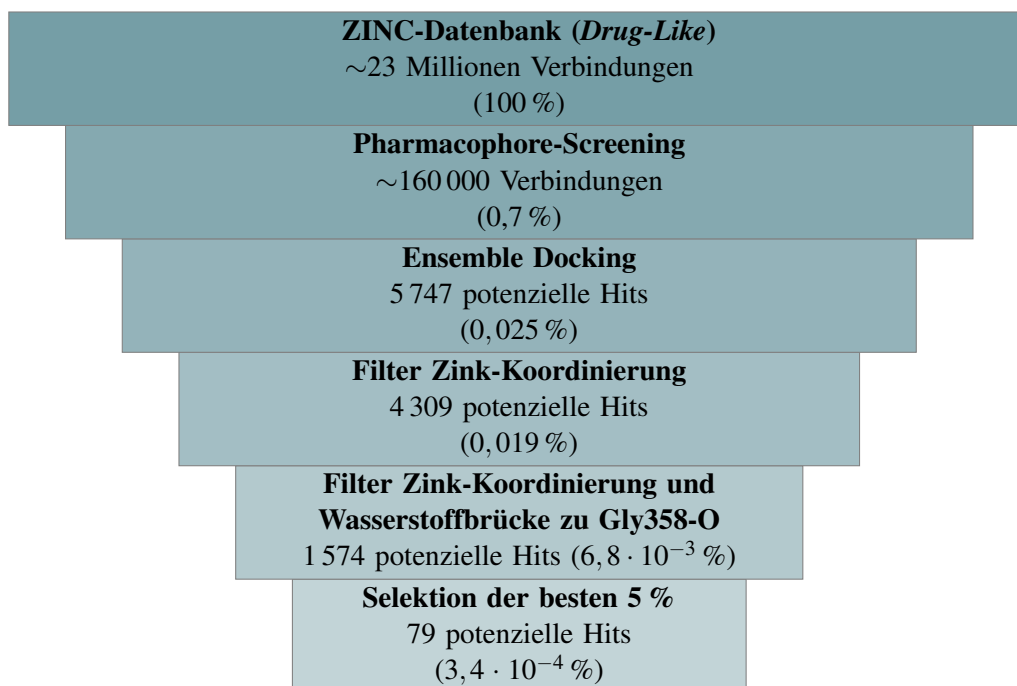


Abbildung 5.5: Virtual Screening-Route für ADAMTS7. In jedem Schritt ist die verbleibende Anzahl an Verbindungen und in Klammern der prozentuale Anteil an Verbindungen im Vergleich zur Ausgangsmenge aufgeführt.

5.3.1 Datenbank

Die Grundlage bzw. Ausgangsmenge für die Suche nach ADAMTS7-Inhibitoren sollte eine möglichst große Anzahl an unterschiedlichen Molekülen sein, die hinsichtlich ihrer Eigenschaften und der potenziellen Bindung an ADAMTS7 gescreent werden sollte. Es gibt einige öffentliche Datenbanken, in denen eine Vielzahl an Verbindungen gespeichert sind, die so z. B. für *in silico*-Analysen zugänglich sind. Die ZINC-Datenbank [153] ist eine solche Datenbank, die als frei verfügbare Ressource aufgebaut wurde und kommerziell erhältliche Stoffe für Virtual Screenings enthält. In dieser Datenbank sind Verbindungen aufgrund ihrer Eigenschaften in verschiedene Untergruppen, z.B. *Drug-Like*, *Lead-Like* und *Fragment-Like*, aufgeteilt. Die Untergruppen der ZINC-Datenbank werden außerdem nach ihrer Verfügbarkeit oder anderen Eigenschaften weiter unterteilt in *Standard*, *Clean*, *In Stock* und *Boutique*. *Standard* enthält die größte Variabilität an Verbindungen. Aus dem *Clean*-Datensatz wurden Verbindungen herausgefiltert, die eine potenziell problematische Funktionalität aufweisen könnten. *In Stock* umfasst Verbindungen, die sofort lieferbar sind und *Boutique* beinhaltet teure Verbindungen (> 100\$ pro Probe), die in keiner der drei anderen Gruppen vorhanden sind [153].

Die ZINC-Datenbank wurde in dieser Arbeit als Grundlage für die Suche nach ADAMTS7-Inhibitoren genutzt. Dabei wurde sich auf eine Untergruppe an Verbindungen beschränkt: die *Drug-Like* Verbindungen. Bei der Suche nach ADAMTS7-Inhibitoren wird nicht nur nach Molekülen gesucht, welche die Proteinfunktion inhibieren. Diese Stoffe müssen ebenfalls bestimmte Eigenschaften erfüllen, um vom Organismus aufgenommen werden zu können. Die Gruppe der *Drug-Like* Verbindungen erfüllen alle die Lipinski *Rule of 5* (siehe Abschnitt 2.8.1). Da diese Untergruppe der ZINC-Datenbank somit wirkstoffartige Moleküle enthält, fiel die Wahl auf diese. In dieser Arbeit wurde der *Clean*- und der *Boutique*-Teil der *Drug-Like*-ZINC-Datenbank für das Virtual Screening genutzt (im dreidimensionalen mol2-Format). Dabei handelte es sich zum damaligen Zeitpunkt um etwa 23 Millionen Moleküle.

Für das Virtual Screening wurden die zu nutzenden Daten der Moleküle in einer eigenen Datenbank gespeichert. Hierfür wurde die Schrödinger Software genutzt. Die Daten aus der ZINC-Datenbank sind bereits mit Schrödinger LigPrep vorbereitet. Mit Schrödinger Phase (Version 3.7) [131] wurde eine weitere Vorbereitung der ZINC-Verbindungen durchgeführt. Dabei wurden den potenziellen Inhibitoren metallbindende Zustände (engl. *metal binding states*) hinzugefügt, welche die potenzielle Interaktion zum Zink-Ion ermöglichen sollten sowie Konformere der Verbindungen (maximal 100) erzeugt.

Die rund 23 Millionen Verbindungen bilden somit die Basis für das Virtual Screening bei der Suche nach Inhibitoren von ADAMTS7.

5.3.2 Pharmacophore-Screening

Ein Docking mit einer Datenbank von mehreren Millionen Verbindungen in die Proteinbindetasche wäre sehr rechen- und zeitaufwendig. Somit erfolgte im zweiten Schritt die Filterung der Datenbank basierend auf dem ADAMTS7-Pharmacophore (vergleiche Abschnitt 5.1). Anhand dieses ersten Filterschritts sollten jene Moleküle selektiert werden, welche die Eigenschaften des Pharmacophores erfüllen und somit an ADAMTS7 binden könnten. In diesem Schritt sollten ebenso Verbindungen herausgefiltert werden, die aufgrund ihrer Größe nicht in die ADAMTS7-Bindetasche passen würden und somit als ADAMTS7-Inhibitoren ausgeschlossen werden können.

Für das Pharmacophore-Screening (Schrödinger Phase, Version 3.7) wurden die rund 23 Millionen Verbindungen aus der erstellten Datenbank mit dem Pharmacophore verglichen. Dabei gilt die Annahme, dass Liganden, die Eigenschaften des Pharmacophores aufweisen, potenziell auch an die Bindetasche binden. Um keine zu starre Filterung vorzunehmen, wurden nicht alle sieben Features des Pharmacophores (vergleiche Abb. 5.3c) als Pflichtbedingung für die Moleküle gesetzt. Dies könnte die chemische Variabilität des Ergebnisses einschränken, was im schlechtesten Fall auch zu keinem Resultat führen könnte. Daher wurde zwischen „Muss“- und „Kann“-Bedingungen unterschieden. Verbindungen wurde aus dem Datensatz aussortiert, wenn keine Zink-bindende Gruppe enthalten, also die Features A2 und A3 nicht erfüllt wurden. Außerdem mussten potenzielle Inhibitoren das Feature D4 für die Wasserstoffbrücke zu Gly358-O vorweisen. Von den übrigen vier Features (A1, D1, D2, D3) musste mindestens ein weiteres Feature, neben A2, A3 und D4, zutreffen. Hier wurde nicht strenger gefiltert, um nicht zu viele mögliche Kandidaten zu verlieren. Insgesamt musste jeder Ligand mindestens vier Features erfüllen, um in diesem Schritt nicht aussortiert zu werden.

Durch das Pharmacophore-Screening wurde die Datenbank an chemischen Verbindungen, welche insgesamt etwa 23 Millionen Verbindungen ausmachten, auf etwa 160 000 Verbindungen reduziert. Dies entspricht einer Verkleinerung des Datensatzes der möglichen Hits um etwa 99,3 %. Der so reduzierte Datensatz sollte im Anschluss anhand von Docking-Studien auf seine Bindeaffinität zu ADAMTS7 getestet werden.

5.3.3 Ensemble Docking

Im nächsten Schritt wurden die verbleibenden Verbindungen mit einem Docking in die Bindetasche eingepasst und mittels einer Scoring-Funktion bewertet. Somit konnte die Affinität der potenziellen Hits zur ADAMTS7-Bindetasche abgeschätzt werden. Folglich wurden die übrigen 160 000 Verbindungen, welche im Pharmacophore-Screening nicht aussortiert wurden, in das ADAMTS7-Homologiemodell sowie in die vier Konformationen der ADAMTS7-Bindetasche gedockt. Dieser Schritt diente dazu, die Liste an potenziellen Liganden von ADAMTS7 weiter zu reduzieren, da trotz Erfüllung der Pharmacophore-Eigenschaften

nicht alle Liganden eine gute Bindeaffinität zum Target aufweisen würden. Des Weiteren gilt die Annahme, dass unter den bestbewerteten Verbindungen und ihren Docking-Posen eine Anreicherung von ADAMTS7-Inhibitoren vorliegt.

Für das Docking wurden Grids, welche die Bindetasche definieren, analog zu Abschnitt 5.2 erstellt. Das Gridzentrum stellte die Aminosäure His388, erstes Histidin im Zink-Bindemotiv, dar und die Box hatte eine Kantenlänge von 25 Å. Weitere Parameter wurden für die Erstellung des Grids nicht verändert. Mit diesen Grids wurde das Docking anhand des Virtual Screening Workflows mit Schrödinger Glide (Version 6.2) durchgeführt. Der Schrödinger Virtual Screening Workflow ermöglicht das Screenen von einer Vielzahl an Verbindungen in mehreren Proteinstrukturen, wie es hier der Fall war, bei einer Vergleichbarkeit der Ergebnisse. In diesem Arbeitsschritt wurde folgendermaßen vorgegangen: Als Docking-Kraftfeld wurde OPLS2005 gewählt, wie bereits bei der ADAMTS7-MD-Simulation. Zuerst wurden die Liganden mit einem HTVS-Docking in die Bindetasche gedockt. Die Liganden wurden flexibel gehalten und die Docking-Posen anschließend minimiert. Dabei wurde für jede Verbindung nur die beste Pose, ausgehend vom Docking Score, analysiert. Nach diesem Schritt wurde von den rund 160 000 Verbindungen nur das beste 1 % der Posen, gemessen am Docking-Score, weiter betrachtet. Dieses beste Prozent wurde erneut in die Bindetasche mit einem XP-Docking eingepasst. Der zweite Docking-Schritt (XP) stellt eine Verfeinerung der ersten Abschätzung der Bindung dar. Die Liganden wurden wie im vorigen Docking-Schritt flexibel gehalten. Es wurden maximal zehn Posen pro Verbindung generiert, die anschließend energieminiert wurden. Für die jeweiligen Dockings wurden keine Bedingungen gesetzt. Diese Docking-Schritte wurden für jedes Conformational Ensemble innerhalb des Workflows separat durchgeführt. Das Verfahren, bei dem ein Datensatz in mehrere Proteinstrukturen gedockt wird, wird auch *Ensemble Docking* genannt. Diese Methode und Vorgehensweise trägt der Flexibilität der Bindetasche Rechnung, die im Docking allgemein als starr betrachtet wird.

Insgesamt lagen nach dem Ensemble Docking 8 093 Posen von 5 747 Verbindungen vor. Dabei handelt es sich jeweils um die bestbewerteten Posen in den jeweiligen Konformationen der ADAMTS7-Bindetasche. Da jede Verbindung in jede Konformation gedockt wurde, liegen mehr Posen als Verbindungen vor. In den folgenden Untersuchungen wurde nur die bestbewertete Pose jeder Verbindung unabhängig von der ADAMTS7-Konformation untersucht, welche anhand des Docking-Scores bestimmt wurde. Im weiteren Verlauf der Arbeit wurden die erhaltenen Docking-Posen visuell analysiert und gefiltert.

5.3.4 Screening und Filterung der Docking-Posen

Nach der Durchführung des Ensemble Dockings wurden die Docking-Posen der chemischen Verbindungen in den Bindetaschen untersucht. Nicht nur der Docking-Score spielt eine wichtige Rolle bei der Abschätzung der Bindeaffinität der Liganden. Ebenso sind bestimmte Interaktionsmuster, wie die Zink-Koordinierung, wahrscheinlich essenziell für die Bindung von ADAMTS-Inhibitoren (vergleiche Abschnitt 4.3). Daher wurde eine Filterung der erhaltenen Docking-Posen nach weiteren Kriterien vorgenommen.

Filter Zink-Koordinierung

Bereits im Vorhinein wurde auf die Wichtigkeit der Zink-Koordinierung von ADAMTS-Inhibitoren hingewiesen. Mithilfe des Pharmacophores wurden Strukturen identifiziert, welche ein Zink-Bindemotiv enthalten. Folglich musste im nächsten Filterschritt überprüft werden, ob die Liganden in ihren Docking-Posen das Zink-Ion koordinieren. Als Grenzwert für die Koordinierung des Zink-Ions wurde ein Abstand von bis

zu 3 Å zwischen Zink-Bindemotiv und Zink-Ion festgelegt. Nach Harding [171], der die Geometrie von Metall-Ligand-Interaktionen in experimentell bestimmten Strukturen untersuchte, beträgt der Abstand zwischen Zink-Ionen und interagierenden Ligand-Atomen zwischen 1,9 Å und 2,5 Å. Hu und Shelver [172] zeigten ferner für MMPs, dass die Distanz zwischen Zink-Ion und Zink-koordinierenden Atomen der Liganden zwischen 1,8 Å und 3,0 Å liegt. In den bekannten, kokristallisierten ADAMTS-Strukturen beträgt der maximale Abstand zwischen Bindemotiv des Liganden und Zink-Ion etwa 2,6 Å. Um keine potenziellen Hits zu verwerfen, wurde daher ein maximaler Abstand von 3 Å zwischen Bindemotiv und Zink-Ion gewählt. Insgesamt erfüllten 4 309 Docking-Posen bzw. Moleküle dieses Kriterium über alle ADAMTS7-Strukturen. Die Startdatenbank konnte bis zu diesem Schritt bereits um 99,081 % reduziert werden.

Filter Zink-Koordinierung und Wasserstoffbrücke zu Gly358-O

Eine wichtige Rolle bei der Bindung von potenziellen ADAMTS(7)-Inhibitoren spielt vermutlich die Aminosäure Gly358, welche über das Sauerstoffatom in der Lage ist eine Wasserstoffbrücke zum Liganden auszubilden. Vorangegangene Untersuchungen (siehe Abschnitte 4.3 und 5.1.2) deuten darauf hin, dass diese Wechselwirkung essenziell für die Bindung von ADAMTS(7)-Inhibitoren sein könnte. Aus diesem Grund wurden in einem weiteren Filterschritt die verbleibenden Docking-Posen auf das Vorhandensein einer Wasserstoffbrücke zwischen dem Ligand und Gly358-O untersucht. Von den 4 309 Zink-koordinierenden Liganden zeigen 1 574 jene Wasserstoffbrücke zu Gly358-O und wurden somit als potenzielle ADAMTS7-Inhibitoren selektiert.

In Abbildung 5.6 sind die Docking-Scores dieser 1 574 potentiellen Hits, aufsteigend sortiert dargestellt. Der niedrigste Docking-Score liegt bei -10,67 (ZINC48847273) und deutet somit auf die höchste

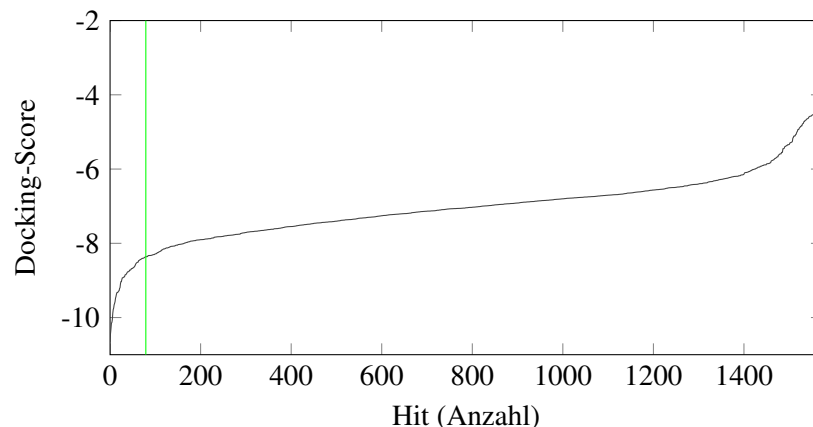


Abbildung 5.6: Docking-Scores der Verbindungen, welche die Zink-Koordinierung und die Wasserstoffbrücke zu Gly358-O in ADAMTS7 aufweisen, aufsteigend sortiert. Die Grenze der besten 79 Verbindungen ist mit einer grünen, vertikalen Linie markiert.

potenzielle Affinität der Verbindung ZINC48847273 zu ADAMTS7 hin. Nach einem sprunghaften Anstieg der Docking-Scores kommt es bei -8 zur Stagnation bzw. einem nur noch leichten Anstieg. Den höchsten Docking-Score (-2,49) und damit die potenziell schlechteste Affinität zu ADAMTS7 weist ZINC03995494 auf. Es wird davon ausgegangen, dass sich die potenziellen Inhibitoren unter den bestbewerteten Verbindungen mit den niedrigsten Docking-Scores befinden. Da keine bekannten ADAMTS7-Inhibitoren vorliegen, konnte diese Annahme hier nicht überprüft werden. Aufgrund der Stabilisierung der Docking-Scores und unter Berücksichtigung des Arbeitsumfangs des Konsortiums CVgenes@target wurden die besten 5 % der

potenziellen 1 574 Hits, basierend auf dem Docking-Score, zur experimentellen Testung vorgeschlagen. Die besten 5 % sind 79 chemische Verbindungen, welche einen Docking-Score von -8,366 oder niedriger aufweisen. Dies lässt auf eine gute Bindeaffinität zwischen ADAMTS7 und Verbindung schließen. Die 79 Verbindungen sind im Anhang A.3.2 mit Strukturformeln aufgelistet. Zusätzlich sollte diese Auswahl weiteren Analysen unterzogen werden, um deren potenziell selektive Wirkung für ADAMTS7 weiter zu untersuchen (siehe Kapitel 6).

Zusammengefasst ergab das Filtern der Posen aus dem Ensemble Docking etwa 1 600 Verbindungen, die in der ADAMTS7-Bindetasche das Zink-Ion koordinieren und eine Wasserstoffbrücke zu Gly358 ausbilden. Sortiert nach Docking-Score wurde ein Grenzwert bei 5 % der Liganden festgelegt und diese zur experimentellen Testung vorgeschlagen. Somit konnte im Virtual Screening eine Datenbank von rund 23 Millionen Strukturen auf 79 Verbindungen reduziert werden, die potenzielle Inhibitoren von ADAMTS7 darstellen (siehe Abb. 5.5).

5.4 Zusammenfassung

Auf Grundlage der sequenziellen und strukturellen Analysen der vorigen beiden Kapitel gelang es, die Bindetasche von ADAMTS7 zu charakterisieren und ein Pharmacophore-Modell für ADAMTS7 abzuleiten, welches insgesamt sieben Features enthält und die potenziellen Eigenschaften von ADAMTS7-Inhibitoren repräsentiert. Diese Ergebnisse konnten bereits in [173] publiziert werden. Mit einem Redocking galt es zudem aus vier Programmen, das beste Dockingprogramm für ADAMTS-Proteine zu identifizieren. Mit dem Docking-Programm Glide von Schrödinger konnten im Redocking die ähnlichsten Ergebnisse im Vergleich zu den nativen Konformationen bekannter ADAMTS-Protein-Inhibitor-Komplexe erzielt werden. Daher wurde davon ausgegangen, dass Glide sich am besten dafür eignet, die räumliche Lage von ADAMTS7-Inhibitoren vorherzusagen und deren Bindeaffinität abzuschätzen.

Anschließend konnte ein Virtual Screening zur Identifizierung von ADAMTS7-Inhibitoren durchgeführt werden. In diesem gelang es, die Ausgangsdatenbank ZINC (*Drug-Like*) [153] mit etwa 23 Millionen Wirkstoffkandidaten auf eine Auswahl von 79 potenziellen ADAMTS7-Inhibitoren zu reduzieren. Die Filterschritte umfassten das Screening mit dem ADAMTS7-Pharmacophore sowie Dockings. Die erhaltenen Docking-Posen wurden anhand des Docking-Scores sowie zwei weiterer Kriterien (Zink-Koordinierung und Wasserstoffbrücke zu Gly358-O) auf ihre Bindungseigenschaften zu ADAMTS7 getestet. Da keine ADAMTS7-Inhibitoren oder Liganden, welche an diese Bindetasche binden, bekannt sind, konnte die Güte der ADAMTS7-Virtual Screening-Pipeline nicht validiert werden. Die verbliebenen 79 Verbindungen wurden zur weiteren experimentellen Testung vorgeschlagen. Die Ergebnisse standen bis zur Fertigstellung der Arbeit noch aus.

6 Ergebnisse und Diskussion: Weiterführende Untersuchungen der potenziellen ADAMTS7-Inhibitoren

Mithilfe des Virtual Screenings konnten 79 potenzielle ADAMTS7-Inhibitoren identifiziert werden (Kapitel 5). Die Docking-Posen dieser 79 Hits wurden hinsichtlich ihrer Interaktionen in ADAMTS7 untersucht. So sollte ein Wechselwirkungsmuster ausgehen von den 79 potenziellen Inhibitoren abgeleitet werden, welches die vorhergesagte Bindeaffinität der potenziellen Inhibitoren zu ADAMTS7 erklärt (siehe Abschnitt 6.1).

Wie bereits in Kapitel 4 festgestellt wurde, sind sich die Proteine der ADAMTS-Familie sequenziell sowie strukturell ähnlich. Es ist daher nicht auszuschließen, dass potenzielle ADAMTS7-Inhibitoren auch an die anderen ADAMTS-Proteine binden. Dies könnte im Hinblick auf das Ziel selektive Inhibitoren für ADAMTS7 zu finden, zu Kreuzreaktionen und unerwünschten Nebenwirkungen führen. Aus diesem Grund sollte mit einem inversen Screening (Abschnitt 6.2) die Affinität der 79 potenziellen ADAMTS7-Inhibitoren zu den 18 weiteren ADAMTS-Proteinen abgeschätzt und die Selektivität der ADAMTS7-Hits bewertet werden (siehe Abschnitte 6.2.1 und 6.2.2).

Im Anschluss galt es zu ermitteln, ob ein Zusammenhang zwischen Länge, Volumen und der Selektivität von potenziellen ADAMTS7-Inhibitoren besteht (siehe Abschnitt 6.3.1). Außerdem wurden die Wechselwirkungen der identifizierten selektiven und nicht selektiven Inhibitoren mit dem ADAMTS7-Protein im Detail untersucht (siehe Abschnitte 6.3.2 und 6.3.3). Abschließend sollte die Stabilität der Bindung der potenziell selektiven Hits anhand ihrer Docking-Posen mittels MD-Simulationen validiert werden (siehe Abschnitt 6.3.4).

6.1 Wechselwirkungen zwischen ADAMTS7-Bindetasche und potenziellen Inhibitoren

Nach der Identifizierung von potenziellen ADAMTS7-Inhibitoren wird hier eine genauere Analyse der Docking-Posen dieser Verbindungen im Virtual Screening vorgenommen. Die Untersuchungen konzentrieren sich auf die 79 potenziellen ADAMTS7-Inhibitoren und deren Wechselwirkungen in der ADAMTS7-Bindetasche.

Interessant bei der Auswertung der Docking-Posen ist, ob die Wirkstoffkandidaten bestimmte Interaktionen (gehäuft) bei der Bindung an die Bindetasche eingehen. Dies hat den Zweck, zu analysieren, ob bestimmte Wechselwirkungen für die guten Docking-Scores in den 79 Posen und somit für eine wahrscheinlich hohe Bindeaffinität zu ADAMTS7 sorgen. Hierzu wurden die Wasserstoffbrückenbindungen, hydrophobe sowie aromatische Interaktionen in den Docking-Posen zwischen Bindetasche und jeweiligem potenziellen Inhibitor untersucht. Die Analysen erfolgten mit dem jeweiligen Hit in der Bindetasche, in der

das jeweilige Docking-Ergebnis gemessen wurde. Demnach wurden die Untersuchungen im ADAMTS7-Homologiemodell oder einem der vier Conformational Ensembles durchgeführt.

Zuerst erfolgt die Auswertung der Wasserstoffbrückenbindungen in den Docking-Posen. In Tabelle 6.1 sind diejenigen Aminosäuren aufgeführt, welche über Wasserstoffbrücken mit den Liganden interagieren sowie deren Häufigkeit in den 79 Docking-Posen. 22 verschiedene Interaktionen dieser Art wurden beobachtet. Die Wasserstoffbrücke zu Gly358-O (Pharmacophore-Feature D4) ist aufgrund des Filterkriteriums in allen Docking-Posen zu finden. Acht Wasserstoffbrücken sind nur bei jeweils einem Liganden zu finden (Häufigkeit 1,3 %) und daher ligandspezifisch einzuschätzen und nicht kennzeichnend für

Tabelle 6.1: Wasserstoffbrückenbindungen zwischen den 79 potenziellen ADAMTS7-Inhibitoren und der ADAMTS7-Bindetasche sowie deren Häufigkeit. Grau markiert sind jene Bindungen zwischen Liganden und Aminosäureseitenketten.

Aminosäure	Atom	Häufigkeit
Glu355	O	10,1 %
Thr356	O	1,3 %
Thr356	O _{γ1}	3,8 %
Leu357	N	1,3 %
Gly358	N	1,3 %
	O	100 %
Ser360	O	1,3 %
Thr385	O _{γ1}	8,8 %
His388	N _{ε2}	1,3 %
Glu389	O _{ε1}	6,3 %
	O _{ε2}	2,5 %
Gly409	N	1,3 %
Phe413	O	5,1 %
Ile414	O	8,9 %
Ser416	O	19,0 %
	O _γ	1,3 %
Pro417	O	6,3 %
Gln418	N	3,8 %
	O	2,5 %
Leu420	O	1,3 %
Tyr421	OH	2,5 %
Asp422	O _{δ2}	2,5 %

spezifische ADAMTS7-Inhibitoren. Vergleicht man die Häufigkeit der Wasserstoffbrücken mit dem erstellten ADAMTS7-Pharmacophore (vergleiche Abschnitt 5.1) sind die Features D1 (Ser416-O) und D3 (Glu389-O_{ε1}, Glu389-O_{ε2}) bzw. die Wasserstoffbrücken mit den assoziierten Aminosäuren mit jeweils 19,0 % bzw. 8,9 % am Häufigsten zu finden. Die Pharmacophore-Features A1 (Leu420-N) und D2 (Asp422-O_{δ2}) bzw. Interaktionen mit den Aminosäuren wurden nicht bzw. bei zwei Posen beobachtet. Basierend auf den Pharmacophore-Untersuchungen wurde postuliert, dass die Features A1 und D2 für die Selektivität von ADAMTS7-Inhibitoren stehen könnten (vergleiche Abschnitt 5.1). In den 79 Docking-Posen sind die Merkmale bzw. Interaktionen nicht bzw. selten identifiziert worden. Jedoch liegen die Hälfte der Wasserstoffbrücken zwischen den potenziellen Inhibitoren und Aminosäuren des L1-Loops vor, sodass in diesem Fall wahrscheinlich generell Wasserstoffbrückenbindungen mit dem L1-Loop für eine Bindeaffinität und Selektivität stehen könnten.

Des Weiteren ist bei der Analyse der Wasserstoffbrücken vor allem interessant, ob Wasserstoffbrücken mit nicht konservierten Seitenketten ausgebildet werden. Dies könnte auf Interaktionen hindeuten, die auf Selektivität schließen lassen, da sie in den anderen ADAMTS-Proteinen nicht ausgebildet werden können. Es konnten innerhalb der 79 Docking-Posen Wasserstoffbrücken zu sieben Seitenketten ausgemacht werden, die in Tabelle 6.1 grau markiert sind. Bei den interagierenden Aminosäuren handelt es sich um Thr356, Thr385, His388, Glu389, Ser416, Tyr421 und Asp422. Nicht konserviert über die ADAMTS-Proteine sind dabei die drei Aminosäuren Ser416, Tyr421 und Asp422, welche im L1-Loop lokalisiert sind. Da jedoch die Wasserstoffbrücken zwischen den Seitenketten und potenziellen Inhibitoren in nur 1-2 Verbindungen vorkommen, sind sie kein Hinweis auf potenziell spezifische ADAMTS7-Interaktionen. Auf die mögliche Wichtigkeit der Aminosäuren Ser416 und Asp422 wurde schon im Rahmen des ADAMTS7-Pharmacophores hingewiesen bzw. auf die Wichtigkeit des L1-Loops im vorigen Absatz.

Ein weiterer Aspekt in den Docking-Posen sind hydrophobe Wechselwirkungen zwischen Protein und potenziellen ADAMTS7-Inhibitoren. Hydrophobe Wechselwirkungen zählen im Gegensatz zu Wasserstoffbrücken zu den schwachen Wechselwirkungen. Zu zwölf Aminosäuren in der ADAMTS7-Bindetasche wurde diese Art der Interaktion festgestellt. Diese sind in Tabelle 6.2 aufgelistet. Bei etwa 46 % bzw. 41 % aller potenziellen ADAMTS7-Inhibitoren wurden hydrophobe Wechselwirkungen zu Leu357 bzw. Phe384 beobachtet. Rund 37 % aller Liganden zeigen Wechselwirkungen zu Pro417 und 29 % zu Leu419. Zwischen 15 % und 12,7% der Liganden zeigen hydrophobe Wechselwirkungen mit Pro381, Tyr421 und Leu359. Mit den Aminosäuren Ala348, Val362, Val408, Ile414 und Leu420 zeigen maximal 5 % der 79 Verbindungen hydrophobe Wechselwirkungen. Folglich sind sieben der zwölf Aminosäuren häufig bei Interaktionen dieser Art beteiligt. Von diesen sieben Aminosäuren sind Leu359, Pro417 und Leu419 nicht konserviert über die ADAMTS-Proteine. Von den zwölf Aminosäuren, welche hydrophobe Wechselwirkungen eingehen, sind insgesamt zehn nicht konserviert (Leu359, Val362, Pro381, Val408, Ile414, Pro417, Leu419, Leu420, Tyr421). Die hydrophoben Interaktionen könnten ein wichtiger Aspekt für eine gute (selektive) Bindeaffinität an ADAMTS7 darstellen.

Tabelle 6.2: Hydrophobe Wechselwirkungen zwischen den 79 potenziellen ADAMTS7-Inhibitoren und der ADAMTS7-Bindetasche sowie deren Häufigkeit.

Aminosäure	Häufigkeit
Ala348	2,5 %
Leu357	45,6 %
Leu359	12,7 %
Val362	5,1 %
Pro381	15,2 %
Phe384	40,5 %
Val408	2,6 %
Ile414	1,3 %
Pro417	36,7 %
Leu419	29,1 %
Leu420	1,3 %
Tyr421	12,7 %

Die letzten Wechselwirkungen, welche genauer betrachtet wurden, sind aromatische Wechselwirkungen. In den 79 Docking-Posen wurden wenige aromatische Interaktionen beobachtet. Dies liegt vor allem an der Tatsache, dass nur eine geringe Anzahl an aromatischen Aminosäuren in der ADAMTS7-Bindetasche vorliegt. In Tabelle 6.3 sind die Aminosäuren, welche diese Art der Wechselwirkung mit den 79 Verbindungen aufweisen, mit der Häufigkeit des Auftretens in den Docking-Posen aufgelistet. Alle aromatischen Interaktionen haben eine Häufigkeit von unter 11 %. Damit sind aromatische Wechselwirkungen wahrscheinlich nicht ausschlaggebend für die Docking-Scores bzw. die Bindeaffinität dieser Verbindungen zu ADAMTS7. Außerdem sind die vier Aminosäuren Phe384, His388, His392 und His398 über die ADAMTS-Proteine konserviert. His388, His392 und His398 sind zudem jene Histidin-Reste, welche das Zink-Ion in den ADAMTS-Proteinen koordinieren. Folglich sind diese Aminosäuren und aromatische Interaktionen wahrscheinlich kein Kriterium, welches für die Bindeaffinität bei den ADAMTS7-Inhibitoren sorgt.

Tabelle 6.3: Aromatische Wechselwirkungen zwischen den 79 potenziellen ADAMTS7-Inhibitoren und Aminosäureseitenketten der ADAMTS7-Bindetasche sowie deren Häufigkeit.

Aminosäure	Häufigkeit
Phe384	1,3 %
His388	10,1 %
His392	1,3 %
His398	6,3 %

Die Wechselwirkungen zwischen ADAMTS7 und den 79 Verbindungen zeichnen in den Docking-Posen ein vielfältiges Bild an Wasserstoffbrückenbindungen, hydrophoben sowie aromatischen Interaktionen. In diesem Fall konnte basierend auf diesen Ergebnissen keine eindeutige Schlussfolgerung über ein einheit-

liches Interaktionsmuster zwischen Protein und Liganden für diese 79 Verbindungen ausgemacht werden, welches eine gute Bindeaffinität an ADAMTS7 garantiert.

6.2 Inverses Screening

Im nächsten Schritt sollte untersucht werden, ob es sich bei den ausgewählten, bestbewerteten 79 Verbindungen möglicherweise um potenziell selektive oder nicht selektive ADAMTS7-Inhibitoren handelt. Es stellte sich die Frage, ob die potenziellen ADAMTS7-Inhibitoren zu den 18 anderen Proteinen der Familie eine ähnliche Bindeaffinität haben. Diese Information ist ein Hinweis bezüglich der ADAMTS7-Selektivität der Verbindungen. Mithilfe eines inversen Screenings (siehe Abschnitt 2.7.4) wurde dieser Fragestellung auf den Grund gegangen. Dafür wurden die 79 Verbindungen aus dem ADAMTS7-Virtual Screening in die Metalloproteinase-Domäne der 18 weiteren ADAMTS-Proteine gedockt, um die Bindeaffinität und die Wechselwirkungen zu vergleichen.

Für diese Untersuchungen wurden die bekannten räumliche Strukturen von ADAMTS-Proteinen sowie die bereits erstellten Homologiemodelle (siehe Abschnitt 4.1) jener ADAMTS-Proteine ohne bekannte 3D-Struktur genutzt. Für drei der 18 ADAMTS-Proteine (ADAMTS1, -4 und -5) lag zum Zeitpunkt des inversen Screenings mehr als eine experimentell bestimmte Struktur, welche die Metalloproteinase-Domäne abdeckt, vor. Aus diesem Grund wurde anhand der folgenden Kriterien, basierend auf der Studie (inverses Screening von Kinase-Inhibitoren) von Santiago *et al.* [174], jeweils eine beispielhafte Struktur ausgewählt:

1. Strukturen mit kokristallisiertem Liganden wurden apo Strukturen vorgezogen.
2. Strukturen mit der höchsten Auflösung wurden gewählt.
3. Besteht eine Struktur aus mehreren Proteinketten, wurde diejenige mit den besten strukturellen Parametern gewählt (anhand des B-Faktors und der Vollständigkeit).
4. (Möglichst) keine fehlenden Aminosäuren in der Proteinbindetasche.

In Tabelle 6.4 sind die ausgewählten, im inversen Screening das jeweilige Protein repräsentierenden, Röntgenkristallstrukturen aufgelistet. Für ADAMTS1 ist die gewählte Röntgenkristallstruktur 2JIH. Bei den

Tabelle 6.4: ADAMTS-Röntgenkristallstrukturen, welche für das inverse Screening ausgewählt wurden. Für jedes Protein ist die PDB-ID der gewählten Struktur sowie Auflösung, Proteinkette und Abdeckung der Proteinsequenz angegeben. Die Auswahl fand im Oktober 2014 anhand der damals bekannten Strukturen statt.

Protein	Struktur (PDB-ID)	Auflösung (Å)	Kette	Abdeckung
ADAMTS1	2JIH [49]	2,10	A	Ser256-Lys548
ADAMTS4	2RJP [47]	2,80	C	Ala214-His508
ADAMTS5	3HYG [50]	1,40	A	Ser264-Ile480

ADAMTS4- (2RJP) und ADAMTS-5-Strukturen (3HYG) handelt es sich um jene, die bereits bei den strukturellen Untersuchungen der ADAMTS-Bindetaschen als repräsentative Strukturen sowie als Template beim Homology Modeling genutzt wurden.

Wie bereits in Abschnitt 4.1 erwähnt, ist eine Aminosäure in der Bindetasche der ausgewählten Röntgenkristallstruktur 2RJP von ADAMTS4 mutiert (Glu362Gln). Diese Aminosäure wurde in die Wildtyp-Aminosäure umgewandelt (Maestro, Version 9.5) und mit dem Kraftfeld OPS2005 energieminiert (MacroModel, Version 10.3). Alle Proteinstrukturen, die für das inverse Screening ausgesucht wurden, wurden einer Vorbereitung mit dem Schrödinger Protein Preparation Wizard (Schrödinger Suite 2014-1) [170] unter-

zogen, analog zum Vorgehen beim Redocking (siehe Abschnitt 5.2). Die 15 ADAMTS-Homologiemodelle, welche alle die Metalloproteinase-Domäne umfassen, standen aus vorausgegangenen Analysen (siehe Abschnitt 4.1) zur Verfügung und wurden wie vorliegend verwendet. Sie wurden bereits einem Loop-Refinement sowie einer Energieminimierung unterzogen.

Es muss bei der Nutzung von Modellen in den folgenden Analysen beachtet werden, dass die ADAMTS-Homologiemodelle eine Vorhersage der räumlichen Struktur sind und keine 100 %ige Aussage zur Genauigkeit getroffen werden kann, da diese Fehler enthalten können. Im gleichen Atemzug können Röntgenkristallstrukturen je nach Auflösung ebenso Ungenauigkeiten und somit Fehler aufweisen. Durch die Anwendung bestimmter Methoden, wie z. B. der Energieminimierung, wurde versucht in den Modellen und den Röntgenkristallstrukturen jene Fehler zu minimieren.

Nach der Auswahl der räumlichen Proteinstrukturen wurden die Bindetaschen der 18 Strukturen mit einem Grid mit Schrödinger Glide (Version 6.2) charakterisiert. Die Grids haben jeweils die zum His388 von ADAMTS7 korrespondierenden Aminosäure als Zentrum, um die Vergleichbarkeit zu den ADAMTS7-Docking-Studien zu gewährleisten. Außerdem wurden die Grids mit einer Kantenlänge von 25 Å definiert und erlauben die Rotation von Hydroxyl- und Thiolgruppen auf Proteinseite. Das inverse Screening wurde analog zum letzten Docking-Schritt beim ADAMTS7-Virtual Screening (siehe Abschnitt 5.3.3) durchgeführt. Beim inversen Screening wurde somit ausschließlich ein präzises Docking mit *extra precision* (XP) mit den 79 Verbindungen in die 18 ADAMTS-Grids durchgeführt sowie deren Bindeaffinität mittels Docking-Scores vorhergesagt.

Wie im Abschnitt 2.7.4 angedeutet, sind Docking-Scores einer Verbindung in unterschiedlichen Proteinstrukturen nicht notwendigerweise vergleichbar. In dem hier betrachteten Fall liegen keine experimentellen Daten zur Bindung der Liganden vor, sodass keine Validierung vorgenommen werden konnte. Daher wurden die Docking-Scores aus dem inversen Screening sowie die Scores aus dem ADAMTS7-Virtual Screening für die 79 Verbindungen nach der MACS-Methode (siehe Abschnitt 2.7.4 bzw. [151]) normalisiert. Die MACS-Methode ermöglicht die Vergleichbarkeit der Docking-Scores innerhalb des inversen Screenings. Für den folgenden Vergleich der Bindeaffinitäten wurden die angepassten Scores (Berechnung: siehe Formel (2.13) in Abschnitt 2.7.4) herangezogen. Im Anhang in Tabelle A.9 sind die normalisierten Docking-Scores aufgelistet.

Um herauszufinden, ob die Verbindungen in den Strukturen der ADAMTS-Familie ähnlich wie in ADAMTS7 binden, wurden zur Auswertung des inversen Screenings die Docking-Posen anhand der gleichen bzw. entsprechend angepassten Kriterien des ADAMTS7-Virtual Screening bewertet:

1. Ist eine Zink-Koordinierung vorhanden?
2. Wird eine (möglicherweise essenzielle) Wasserstoffbrücke zwischen Inhibitor und der zu Gly358-O (ADAMTS7) korrespondierenden Aminosäure in dem jeweiligen Protein ausgebildet?
3. Zusätzlich wurden die Docking-Posen anhand des Docking-Scores klassifiziert. Hierbei wird für jeden Liganden der normalisierte Docking-Score des ADAMTS7-Virtual Screenings als Grenzwert gesetzt. Im Folgenden wird von einem „guten“ Docking-Score geschrieben, wenn der Docking-Score des Liganden in der jeweiligen Bindetasche j im inversen Screening gleich oder niedriger als der Grenzwert ist (d. h. $\text{Docking-Score}_{\text{Ligand},j} \leq \text{Docking-Score}_{\text{Ligand,ADAMTS7}}$). Andernfalls wird der Docking-Score als „schlecht“ bezeichnet (d. h. $\text{Docking-Score}_{\text{Ligand},j} > \text{Docking-Score}_{\text{Ligand,ADAMTS7}}$). Dies ist in der Erwartung begründet, dass Docking-Posen, die gleich oder besser bewertet werden als die ADAMTS7-Pose eine „gute“ Bindeaffinität, ähnlich wie zu ADAMTS7, aufweisen und ansonsten im

Vergleich zu ADAMTS7 „schlechter“ an das jeweilige Protein binden. Letzteres lässt darauf schließen, dass die Bindeaffinität zu ADAMTS7 höher sein könnte.

Es wurde für jeden Liganden nur die beste Docking-Pose je Protein, basierend auf dem Docking-Score, analysiert. In Abbildung 6.1 ist das Ergebnis des inversen Screenings für die 18 ADAMTS-Proteine zusammengefasst. Deutlich wird, dass in den ADAMTS-Proteinen zwischen 3,8 % (ADAMTS3) und 36,7 % (ADAMTS1) der untersuchten Verbindungen alle drei oben genannten Kriterien für eine potenziell gute Bindung bzw. Bindeaffinität erfüllen. Im Durchschnitt liegt der Anteil bei 16,6 %. Außerdem beträgt der Anteil der Liganden, die das Zink-Ion koordinieren und dabei einen guten Docking-Score unter dem Grenzwert des ADAMTS7-Virtual Screenings aufweisen zwischen 0 % (ADAMTS19) und 49,9 % (ADAMTS1) (Durchschnitt: 15,8 %). Diese Ergebnisse deuten darauf hin, dass in dem Datensatz potenzieller ADAMTS7-Inhibitoren nicht nur selektive Verbindungen zu finden sind. Gleichzeitig zeigt sich, dass zwischen 1,3 % (ADAMTS1) und 49,4 % (ADAMTS12 und ADAMTS17) der Liganden das Zink-Ion koordinieren, eine Wasserstoffbrücke zum Gly358-O-Analog ausbilden, aber eine geringere Bindeaffinität als zu ADAMTS7 aufweisen (Durchschnitt: 24,9 %). Durchschnittlich 35,8 % (zwischen 4,5 % (ADAMTS1) und 64,7 % (ADAMTS14) der getesteten Verbindungen weisen im inversen Screening lediglich eine Zink-Koordinierung zum ADAMTS-Protein auf, jedoch keine Wasserstoffbrücke zum Gly358-O-Analog sowie eine schlechte Bindeaffinität zum jeweiligen Protein. Außerdem zeigt sich, dass nur in ADAMTS5 2,5 % der getesteten Liganden im inversen Screening unter Fehlen der Zink-Koordinierung und der Wasserstoffbrücke eine gute Bindeaffinität zu den ADAMTS-Proteinen aufweisen. In allen anderen ADAMTS-Proteinen erfüllt kein Ligand dieses Kriterium.

Es stellten sich die folgenden Fragen: Können aus den Ergebnissen des inversen Screenings und den Sequenz- und Strukturvergleichen der ADAMTS-Bindetaschen aus Abschnitt 4.2.2 Rückschlüsse auf die Bindeaffinität der 79 Verbindungen gezogen werden? Lässt sich ein Unterschied in der Bindeaffinität der Liganden zu Proteinen erkennen, die zu ADAMTS7 weniger ähnlich bzw. strukturell diverser sind? Den geringsten Anteil an Verbindungen, die alle drei Kriterien erfüllen, zeigt ADAMTS3 mit 3,8 %. Das Protein ADAMTS3 ist im Vergleich zu ADAMTS7 in der Bindetasche weniger stark konserviert als die anderen ADAMTS-Proteine (vergleiche Tabelle 4.9): Die Sequenzidentität zur S4-Region beträgt 38,5 % (ADAMTS-Durchschnitt: 46,5 %), zur H4-Region 57,1 % (ADAMTS-Durchschnitt: 63,3 %) und zum L1-Loop 26,7 % (ADAMTS-Durchschnitt: 25,2 %). Strukturell zeigt das S4-Faltblatt-Analogon in ADAMTS3 eine Differenz von 10,5 Å zu ADAMTS7 und im L1-Loop unterscheidet sich ADAMTS3 um 3,3 Å zu ADAMTS7 (siehe Tabelle 4.10). Diese sequenziellen und strukturellen Unterschiede zwischen der ADAMTS3-Bindetasche zu ADAMTS7 könnten für den vergleichsweise geringen Anteil an Docking-Posen, welche die drei Kriterien erfüllen, ursächlich sein. Den höchsten Anteil an Verbindungen, die die drei Kriterien erfüllen, zeigte ADAMTS1 (36,7 %). Die ADAMTS1-Bindetasche weist eine durchschnittliche Sequenzidentität zu ADAMTS7 (außer im L1-Loop) im Vergleich zu den weiteren ADAMTS-Proteinen auf (vergleiche Tabelle 4.9): Die Sequenzidentität zur S4-Region beträgt 40,0 % (ADAMTS-Durchschnitt: 46,5 %), zur H4-Region 61,9 % (ADAMTS-Durchschnitt: 63,3 %) und zum L1-Loop 16,7 % (ADAMTS-Durchschnitt: 25,2 %). Eine hohe Konservierung ist daher nicht verantwortlich für dieses Ergebnis. Der $\text{RMSD}_{\text{C}\alpha}$ des L1-Loops beträgt 6,8 Å zwischen ADAMTS1 und ADAMTS7, was etwa der durchschnittlichen strukturellen Abweichung dieser Region zu ADAMTS7 (6,1 Å) entspricht (siehe Tabelle 4.10). Trotz dieser durchschnittlichen sequenziellen und strukturellen Unterschiede weisen etwa zwei Fünftel der getesteten Verbindungen eine ähnliche Bindeaffinität zu ADAMTS1 wie zu ADAMTS7 auf. Die Docking-Posen der poten-

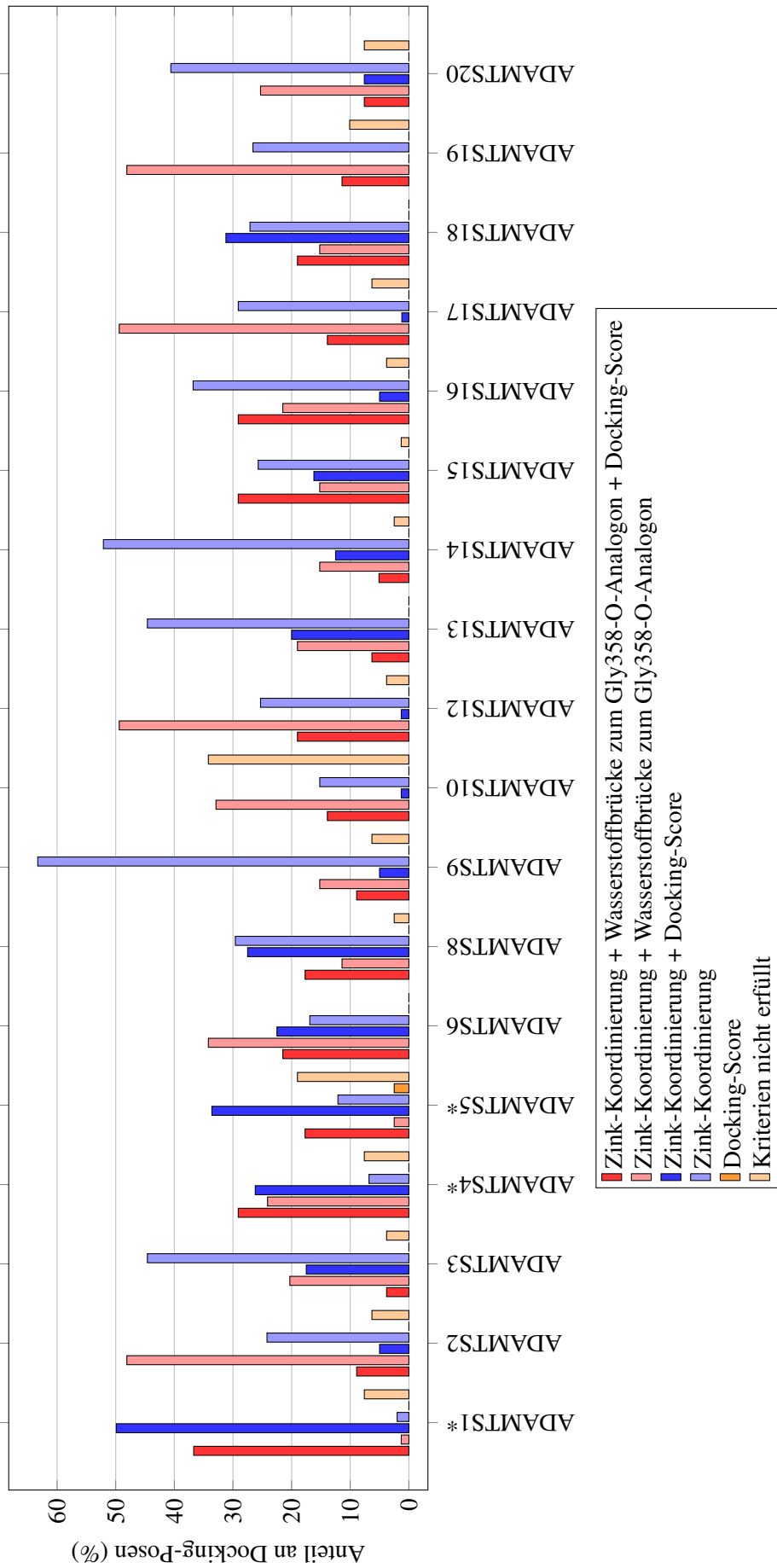


Abbildung 6.1: Auswertung der Docking-Posen des inversen Screenings für die ausgewählten 79 Verbindungen aus dem ADAMTS7-Virtual Screening. Die Klassifizierung fand entsprechend je nach Erfüllung der folgenden Kriterien statt: Zink-Koordinierung, Vorhandensein der Wasserstoffbrücke zum Gly358-O-Analogon und Docking-Score kleiner gleich Grenzwert (ADAMTS7). * - Röntgenkristallstruktur.

ziellen ADAMTS7-Inhibitoren in ADAMTS12, das zu ADAMTS7 sequenziell und strukturell ähnlichste Protein (vergleiche Abschnitt 4.2.2), erfüllen nur zu 19,0 % die drei Kriterien. Es wurde jedoch beobachtet, dass etwa die Hälfte der 79 Verbindungen in ADAMTS12 Zink-koordiniert sind und eine Wasserstoffbrücke zum Gly358-O-Analogon ausbilden. Folglich sind zwei Kriterien erfüllt, wobei die Bindeaffinität im Vergleich zu ADAMTS7 schlechter ist.

Die Auswertung des inversen Screenings zeigte, dass unter den Verbindungen, welche als potenzielle ADAMTS7-Inhibitoren identifiziert wurden, einige Verbindungen ebenfalls an die anderen 18 getesteten ADAMTS-Proteine unter den gewählten Filterkriterien mit einer ähnlichen Affinität binden. Folglich muss die Selektivität der Verbindungen genauer untersucht werden. Es konnte nicht gezeigt werden, dass von einer sequenziellen und strukturellen Ähnlichkeit bzw. Unähnlichkeit der ADAMTS-Bindetaschen (vergleiche Abschnitt 4.2) auf eine ähnlich affine bzw. eine weniger affine Bindung der 79 potenziellen ADAMTS7-Inhibitoren geschlossen werden kann. Eine Frage stellt sich besonders nach der vorausgegangenen Analyse: Gibt es Verbindungen, die im inversen Screening alle bzw. keines der Kriterien in allen getesteten ADAMTS-Proteinen erfüllen?

6.2.1 Potenziell nicht selektive Hits

Zur Beantwortung der gestellten Frage wurde zuerst geprüft, ob es ADAMTS7-Hits gibt, die im inversen Screening gut, d. h. unter Erfüllung der drei oben genannten Kriterien, an die 18 getesteten Strukturen binden. Wenn Verbindungen in mindestens einer anderen Bindetaschen an das Zink-Ion koordiniert sind, eine Wasserstoffbrücke zum Gly358-O-Analogon ausbilden und einen normalisierten Docking-Score kleiner gleich dem Score des ADAMTS7-Virtual Screenings aufweisen, kann davon ausgegangen werden, dass die Bindung ähnlich affin wie zu ADAMTS7 ist. Dies wäre ein Hinweis darauf, dass jene Verbindungen womöglich keine spezifischen Inhibitoren für ADAMTS7 sein könnten und auch auf andere Proteine dieser Familie eine Wirkung haben könnten. Daher sind diese ADAMTS7-Hits wahrscheinlich nicht selektiv.

Gibt es Strukturen, die in allen ADAMTS-Proteinen die drei Kriterien erfüllen? Auf keine der 79 gedockten Verbindungen trifft dies zu (vergleiche Tabelle A.9 im Anhang). Folglich zeigt keiner der 79 potenziellen ADAMTS7-Inhibitoren eine Affinität zur gesamten Proteinfamilie. Es zeigte sich, dass die untersuchten Verbindungen in maximal 10 weiteren ADAMTS-Proteinen die oben genannten Kriterien erfüllen: Fünf Verbindungen (ZINC00895294, ZINC05131990, ZINC08551175, ZINC45076864, ZINC71186797) zeigen (im Vergleich zu ADAMTS7) zu zehn ADAMTS-Proteinen eine gute Bindeaffinität, zwei Verbindungen (ZINC03869786, ZINC18141652) zu neun ADAMTS-Proteinen. Im Gegensatz dazu konnte bei elf Verbindungen (ZINC37374946, ZINC37414218, ZINC40536338, ZINC42430101, ZINC48858643, ZINC49571689, ZINC49974870, ZINC52011629, ZINC62424078, ZINC62424080, ZINC62468257) eine ähnliche Bindeaffinität wie zu ADAMTS7 nur bei einem weiteren ADAMTS-Protein beobachtet werden. Es wurde davon ausgegangen, dass Liganden, deren beste Docking-Pose in mindestens 50 % der getesteten Strukturen (entspricht neun ADAMTS-Proteinen) die drei Kriterien erfüllen, nicht zu den selektiven ADAMTS7-Inhibitoren gezählt werden können. Auf sieben der 79 Verbindungen trifft dies zu. Die sieben Verbindungen stammen aus dem *Clean*-Datensatz der *Drug-Like* ZINC-Datenbank. In Abbildung 6.2 sind die Strukturformeln dieser Verbindungen dargestellt. Die Auflistung der Strukturformeln zeigt, dass sich unter den sieben Verbindungen zwei Enantiomerenpaare befinden (ZINC00895294 (siehe Abb. 6.2a) und ZINC08551175 (siehe Abb. 6.2d); ZINC03869786 (siehe Abb. 6.2b) und ZINC05131990 (siehe Abb. 6.2c)). Bis auf ZINC71186797 zeigen die Verbindungen einen sehr ähnlichen strukturellen Aufbau: Sechs Verbin-

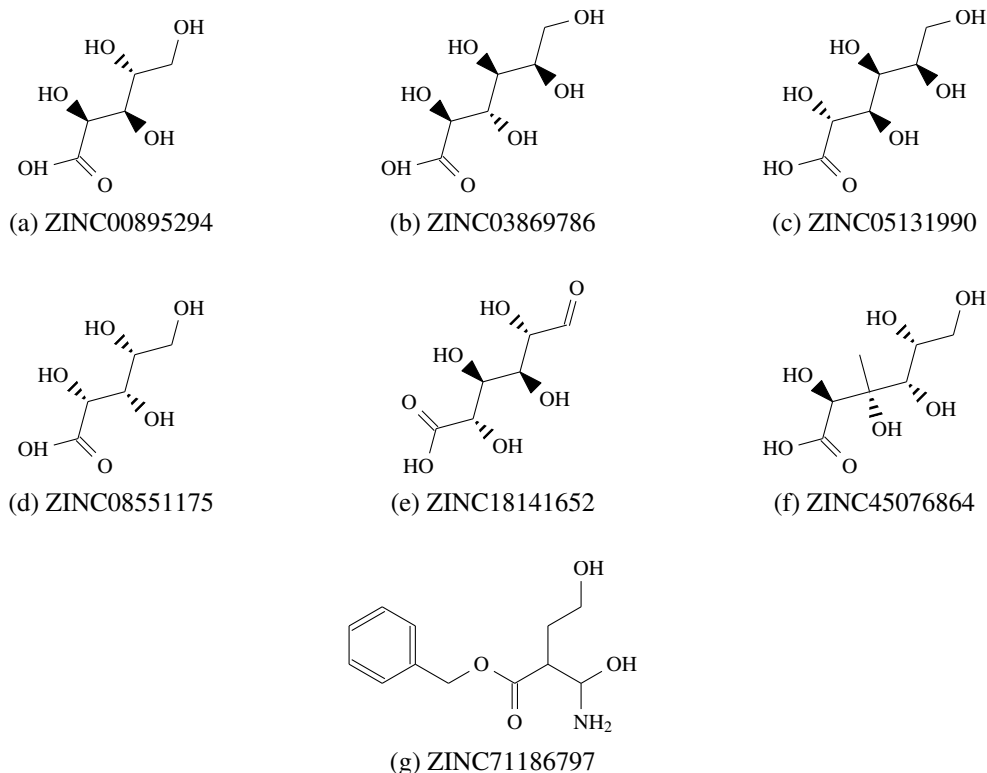


Abbildung 6.2: Strukturformeln der sieben potenziell für ADAMTS7 nicht selektiven Inhibitoren.

dungen (außer ZINC71186797 (siehe Abb. 6.2g)) setzen sich aus kurzen, linearen Kohlenstoffketten zusammen, welche jeweils eine $-COOH$ -Gruppe sowie vier bis fünf $-OH$ -Gruppen aufweisen. ZINC71186797 hingegen ist eine Verbindung mit zwei $-OH$ -Gruppen sowie einer Phenylgruppe. Dass sechs der sieben Verbindungen (sehr) ähnlich zueinander sind, zeigt auch der Vergleich der Strukturen mittels Fingerprints (Schrödinger Canvas, Version 1.9, Tanimoto similarity, MOLPRINT2D Fingerprints) in Tabelle 6.5. Durch ihren Aufbau können sich diese Moleküle alle in die länglichen Vertiefungen der ADAMTS-Bindetaschen ein- bzw. anpassen. Daher ist zu erwarten, dass sie wahrscheinlich in die meisten ADAMTS-Bindetaschen unter Erfüllung der gesetzten Kriterien binden. Auf das genaue Bindemuster der sieben Substanzen im ADAMTS7-Protein wird später in Abschnitt 6.3.2 eingegangen.

Tabelle 6.5: Ähnlichkeitsmatrix (Tanimoto-Koeffizient) der sieben potenziell nicht selektiven ADAMTS7-Hits anhand von Fingerprints. Die Ähnlichkeitswerte sind farbkodiert.

Ähnlichkeit	ZINC00895294	ZINC03869786	ZINC05131990	ZINC08551175	ZINC18141652	ZINC45076864	ZINC71186797
ZINC00895294	1,0	0,9	0,9	1,0	0,5	0,5	0,1
ZINC03869786	0,9	1,0	1,0	0,9	0,5	0,5	0,9
ZINC05131990	0,9	1,0	1,0	0,9	0,5	0,5	0,1
ZINC08551175	1,0	0,9	0,9	1,0	0,8	0,5	0,1
ZINC18141652	0,5	0,5	0,5	0,5	1,0	0,3	0,0
ZINC45076864	0,5	0,5	0,5	0,5	0,3	1,0	0,1
ZINC71186797	0,1	0,1	0,1	0,1	0,0	0,1	1,0

Somit konnte postuliert werden, dass es sich bei Verbindungen mit kurzen, linearen Kohlenstoffketten und mehreren –OH-Gruppen vermutlich nicht um selektive ADAMTS7-Inhibitoren handelt. Die sieben untersuchten, nicht selektiven Verbindungen sind wenig komplex und können wahrscheinlich auf Basis dieser Ergebnisse in die meisten ADAMTS-Bindetaschen bei gleichzeitig hoher Affinität binden.

6.2.2 Potenziell selektive Hits

Die interessantere Frage, welche durch das inverse Screening beantwortet werden sollte, ist, ob es unter den 79 Verbindungen Moleküle gibt, die bezogen auf die 18 getesteten ADAMTS-Strukturen nie alle drei Filterkriterien (Zink-Koordinierung, Wasserstoffbrücke zur Gly358-O analogen Aminosäure und normalisierter Docking-Score \leq normalisierter ADAMTS7-Docking-Score) auf einmal erfüllen. Insgesamt weisen 18 der 79 Strukturen des ADAMTS7-Virtual Screenings im inversen Screening keine Bindeaffinität zu den getesteten ADAMTS-Strukturen auf (vergleiche Tabelle A.9 im Anhang). Somit kann davon ausgegangen werden, dass es sich bei den 18 Verbindungen um potenziell selektive ADAMTS7-Inhibitoren handelt. 15 der 18 Verbindungen stammen aus dem *Boutique*-Teil der *Drug-Like* ZINC-Datenbank, die übrigen drei (ZINC86853423, ZINC94569517 und ZINC94569992) aus der *Clean*-Untergruppe.

In Abbildung 6.3 sind die 2D-Strukturen dieser Verbindungen dargestellt. Im Vergleich zu den sieben potenziell nicht selektiven Verbindungen (siehe Abschnitt 6.2.1), enthalten die Strukturen der 18 potenziell selektiven Hits nicht ausschließlich –OH-Gruppen. Sie weisen weitere funktionelle Gruppen auf und sind damit größer und komplexer: Bei den 18 Verbindungen ist jeweils mindestens ein (aromatischer) Carbocyclus oder Heterocyclus vorhanden. Des Weiteren finden sich Aminogruppen und Gruppen mit Schwefel-Atomen in den Molekülen. 13 der 18 Verbindungen beinhalten eine Peptidbindung. Daher wurde postuliert, dass diese komplexeren Liganden sich nicht an jede der ADAMTS-Bindetaschen einpassen können und daher vermutlich eine Selektivität für das ADAMTS7-Protein postuliert werden kann. Auf die Bindemuster im ADAMTS7-Protein wird in Abschnitt 6.3.3 eingegangen.

Ein Vergleich der 18 Strukturen mittel sogenannter Fingerprints (Schrödinger Canvas, Version 1.9, Tanimoto similarity, MOLPRINT2D Fingerprints) sollte Aufschluss darüber geben, wie ähnlich sich die potenziellen Hits sind (vergleiche Abschnitt 2.8.2). In Tabelle 6.6 sind die Ähnlichkeitswerte für die untersuchten Verbindungen dargestellt. Die Analyse zeigte, dass die 18 Verbindungen nicht sehr ähnlich zueinander sind: Die höchste Ähnlichkeit zueinander weisen die Verbindungen ZINC52012042 und ZINC52012291 mit 0,7 auf. Die weiteren Ähnlichkeitswerte liegen zwischen 0,0 und 0,5, wobei der Durchschnitt 0,1 beträgt. Die Unterschiedlichkeit der Strukturen ermöglicht die Abdeckung eines großen chemischen Raumes bei der Suche nach ADAMTS7-Inhibitoren.

Im inversen Screening wurden 18 Verbindungen ausgemacht, welche zu den untersuchten ADAMTS-Proteinen eine schlechtere Bindeaffinität als zu ADAMTS7 aufzeigten. Diese 18 Verbindungen sind daher unter den 79 ausgewählten Verbindungen aus dem ADAMTS7-Virtual Screening potenziell selektive Verbindungen und ermöglichen einen ersten Ansatz für die Charakterisierung der Selektivität von Inhibitoren für ADAMTS7.

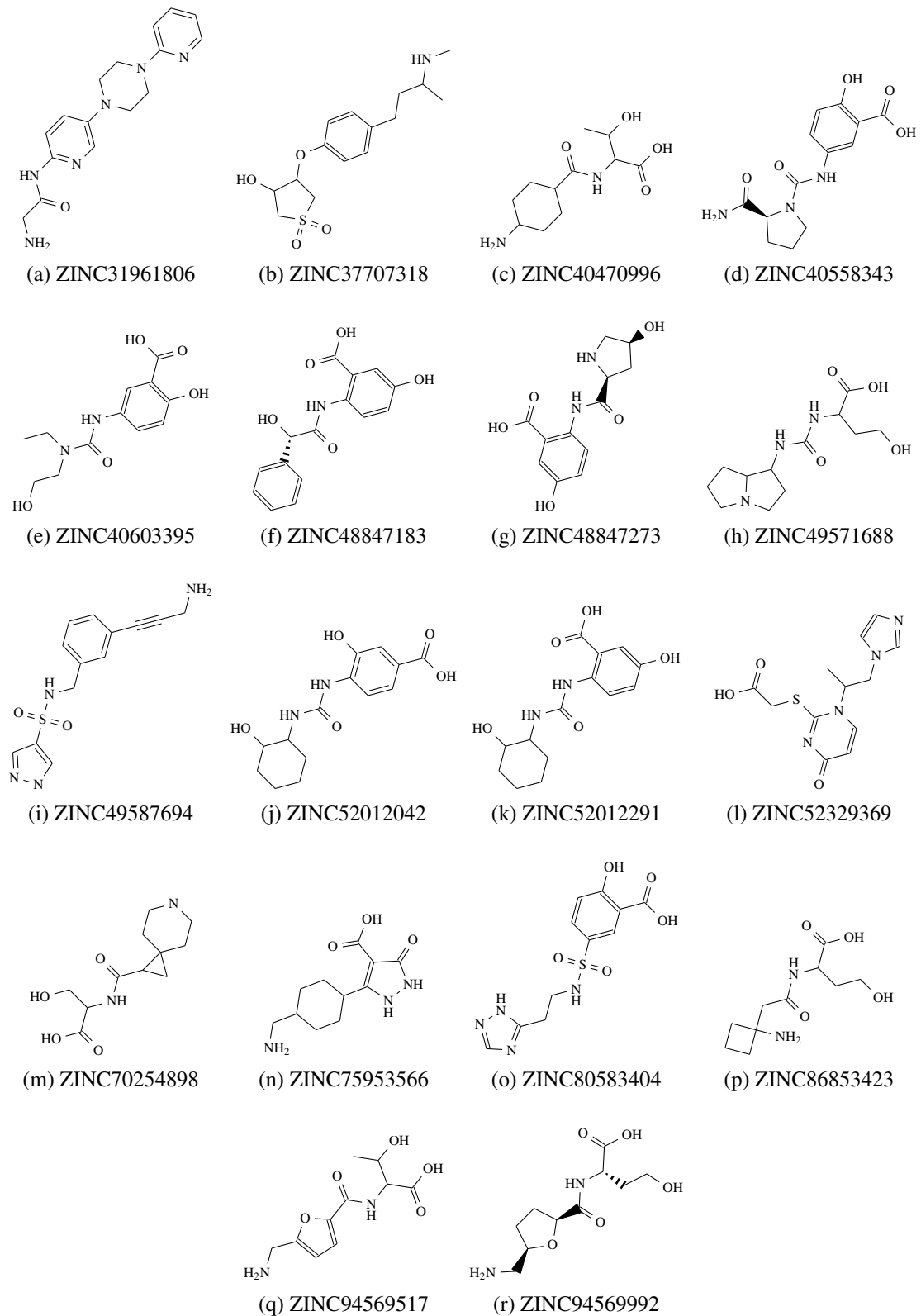


Abbildung 6.3: Strukturformeln der 18 für ADAMTS7 potenziell selektiven Inhibitoren.

Tabelle 6.6: Ähnlichkeitsmatrix (Tanimoto-Koeffizient) der 18 potenziellen ADAMTS7-Hits anhand von Fingerprints. Die Ähnlichkeitswerte sind farbkodiert.

Ähnlichkeit	ZINC31961806	ZINC37707318	ZINC40470996	ZINC40558343	ZINC40603395	ZINC48847183	ZINC48847273	ZINC49571688	ZINC49587694	ZINC52012042	ZINC52012291	ZINC52329369	ZINC70254898	ZINC75953566	ZINC80583404	ZINC86853423	ZINC94569517	ZINC94569992
ZINC31961806	1,0	0,0	0,0	0,1	0,0	0,1	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
ZINC37707318	0,0	1,0	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
ZINC40470996	0,0	0,1	1,0	0,0	0,0	0,0	0,1	0,1	0,0	0,0	0,0	0,1	0,3	0,0	0,0	0,2	0,3	0,2
ZINC40558343	0,1	0,0	0,0	1,0	0,5	0,2	0,2	0,0	0,0	0,2	0,2	0,0	0,0	0,1	0,2	0,0	0,0	0,0
ZINC40603395	0,0	0,0	0,0	0,5	1,0	0,2	0,2	0,1	0,0	0,2	0,3	0,0	0,0	0,1	0,2	0,0	0,0	0,0
ZINC48847183	0,1	0,0	0,0	0,2	0,2	1,0	0,5	0,0	0,1	0,2	0,4	0,0	0,0	0,1	0,2	0,0	0,0	0,0
ZINC48847273	0,1	0,0	0,1	0,2	0,2	0,5	1,0	0,0	0,0	0,2	0,4	0,0	0,0	0,1	0,1	0,0	0,0	0,0
ZINC49571688	0,0	0,0	0,1	0,0	0,1	0,0	0,0	1,0	0,0	0,1	0,1	0,1	0,1	0,0	0,0	0,3	0,1	0,3
ZINC49587694	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,0	1,0	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,0	0,0
ZINC52012042	0,0	0,0	0,0	0,2	0,2	0,2	0,2	0,1	0,0	1,0	0,7	0,0	0,0	0,0	0,1	0,1	0,0	0,0
ZINC52012291	0,0	0,0	0,0	0,2	0,3	0,4	0,4	0,1	0,0	0,7	1,0	0,0	0,0	0,1	0,1	0,0	0,0	0,0
ZINC52329369	0,0	0,0	0,1	0,0	0,0	0,0	0,0	0,1	0,0	0,0	0,0	1,0	0,1	0,0	0,0	0,1	0,1	0,1
ZINC70254898	0,0	0,0	0,3	0,0	0,0	0,0	0,0	0,1	0,0	0,0	0,0	0,1	1,0	0,0	0,0	0,3	0,1	0,2
ZINC75953566	0,0	0,0	0,0	0,1	0,1	0,1	0,1	0,0	0,0	0,1	0,1	0,0	0,0	1,0	0,1	0,0	0,0	0,0
ZINC80583404	0,0	0,0	0,0	0,2	0,2	0,2	0,1	0,0	0,1	0,1	0,1	0,0	0,0	0,1	1,0	0,0	0,0	0,0
ZINC86853423	0,0	0,0	0,2	0,0	0,0	0,0	0,0	0,3	0,0	0,0	0,0	0,1	0,3	0,0	0,0	1,0	0,1	0,4
ZINC94569517	0,0	0,0	0,3	0,0	0,0	0,0	0,0	0,1	0,0	0,0	0,0	0,1	0,1	0,0	0,0	0,1	1,0	0,1
ZINC94569992	0,0	0,0	0,2	0,0	0,0	0,0	0,0	0,3	0,0	0,0	0,0	0,1	0,2	0,0	0,0	0,4	0,1	1,0

Ähnlichkeit

6.3 Wechselwirkungen zwischen ADAMTS7-Bindetasche und potenziell (nicht) selektiven Inhibitoren

Bevor die Docking-Posen der potenziell nicht selektiven bzw. der potenziell selektiven ADAMTS7-Inhibitoren untersucht wurden, um zu analysieren welche Bereiche bzw. Interaktionen zum ADAMTS7-Protein wahrscheinlich für eine gute Bindeaffinität bzw. Selektivität sorgen, sollten die Verbindungen hinsichtlich ihrer Länge und ihres Volumens verglichen werden.

6.3.1 Analyse des Zusammenhangs zwischen Länge, Volumen und Selektivität von potenziellen ADAMTS7-Inhibitoren

Die strukturellen Daten der 7 nicht selektiven bzw. 18 selektiven Hits (siehe Abschnitte 6.2.1 und 6.3.3) ließen bereits vermuten, dass die potenziell selektiven Inhibitoren größer sind als die nicht selektiven. Diese Feststellung sollte validiert werden. Hierfür wurde die maximale Länge der Liganden in den ADAMTS7-Docking-Posen gemessen sowie das Volumen, welches die Verbindungen in den Docking-Posen einnehmen, berechnet.

In Abbildung 6.4 sind die Länge-Volumen-Daten für die 25 Verbindungen aufgetragen und je nach Gruppenzugehörigkeit farblich markiert. Hierbei ist zu erkennen, dass die nicht selektiven Hits kleiner sind ($\leq 7,8 \text{ \AA}$) und dementsprechend auch ein kleineres Volumen ($\leq 191 \text{ \AA}^3$) einnehmen als die potenziell selektiven ADAMTS7-Inhibitoren, deren Länge mindestens $7,7 \text{ \AA}$ und Volumen mindestens 178 \AA^3 beträgt. Basierend auf den Datenpunkten wurde davon ausgegangen, dass die beiden Gruppen linear trennbar sind.

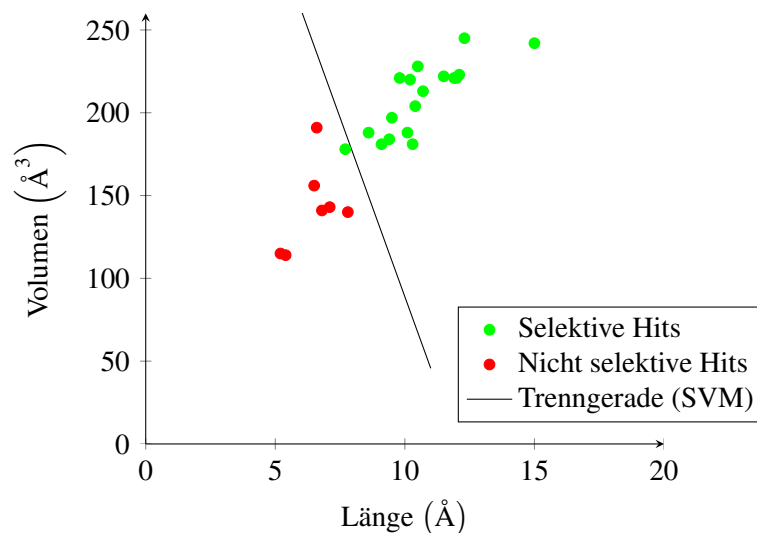


Abbildung 6.4: Längen- und Volumenvergleich der potenziell nicht selektiven und selektiven ADAMTS7-Hits.

Aus diesem Grund wurde mittels einer *Support Vector Machine* (SVM) [175] eine Gerade bestimmt, welche die beiden Gruppen möglichst gut trennen kann. Hierfür wurde mit R [176] (Skript: siehe Anhang A.4.2) eine Klassifikation vorgenommen. Das Resultat ist die folgende Klassifikationsgerade: $1,038 \cdot x + 0,024 \cdot y - 12,514 = 0$, wobei x die Länge (Å) und y das Volumen (Å³) des Liganden bezeichnen. Die Gerade ist in Abbildung 6.4 dargestellt. Nur ein potenziell selektiver Inhibitor wird anhand dieser Trenngeraden falsch klassifiziert und als nicht selektiv eingeordnet. Für die weiteren 24 Verbindungen wird eine korrekte Einordnung vorgenommen. Folglich kann ein Zusammenhang zwischen Länge, Volumen und potenzieller Selektivität für die ADAMTS7-Bindetasche postuliert werden.

Es ist somit wahrscheinlich, dass längere Liganden mit einem größeren Volumen eher selektiv an ADAMTS7 binden als kleinere Liganden. Dies bestätigt die These, dass sich kleinere Liganden besser in unterschiedliche große ADAMTS-Bindetaschen einpassen können und somit wahrscheinlich keine selektiven Bindemuster aufweisen. Somit ist die Größe bzw. das Volumen von potenziellen ADAMTS7-Liganden wahrscheinlich ein wichtiges Merkmal. Dabei muss ebenso beachtet werden, dass die potenziell selektiven Inhibitoren neben der Größe vielfältigere funktionelle Gruppen aufweisen als die nicht selektiven.

6.3.2 Wechselwirkungen der potenziell nicht selektiven Hits

Im Folgenden werden die Wechselwirkungen der sieben potenziell nicht selektiven ADAMTS7-Hits anhand der Docking-Posen des Virtual Screenings ausgewertet. Ziel ist es, Gemeinsamkeiten und Unterschiede in den Bindemustern auszumachen, um eventuell essenzielle Eigenschaften aufzudecken. Zu diesem Zweck sind in den Abbildungen 6.5a-6.5g die Docking-Posen der potenziellen, nicht selektiven Inhibitoren mit ADAMTS7 dargestellt. Diese zeigen die Wechselwirkungen über Wasserstoffbrücken. Im Anhang in den Abbildungen A.7a-A.7g sind die Inhibitor-Posen mit den Bindetaschen in Oberflächendarstellung visualisiert. Dies ermöglicht einen räumlichen Eindruck der Docking-Posen sowie die Darstellung der Eigenschaften der Bindetasche.

Ein Blick auf die sieben Docking-Posen zeigt, dass in allen Posen der Inhibitor an das Zink-Ion koordiniert sowie die Wasserstoffbrücke zwischen Gly358-O und Inhibitor vorhanden ist. Alle Verbindungen koordinieren das Zink-Ion über ihre $-\text{COOH}$ -Gruppe, bis auf den Liganden ZINC71186797. Dieser

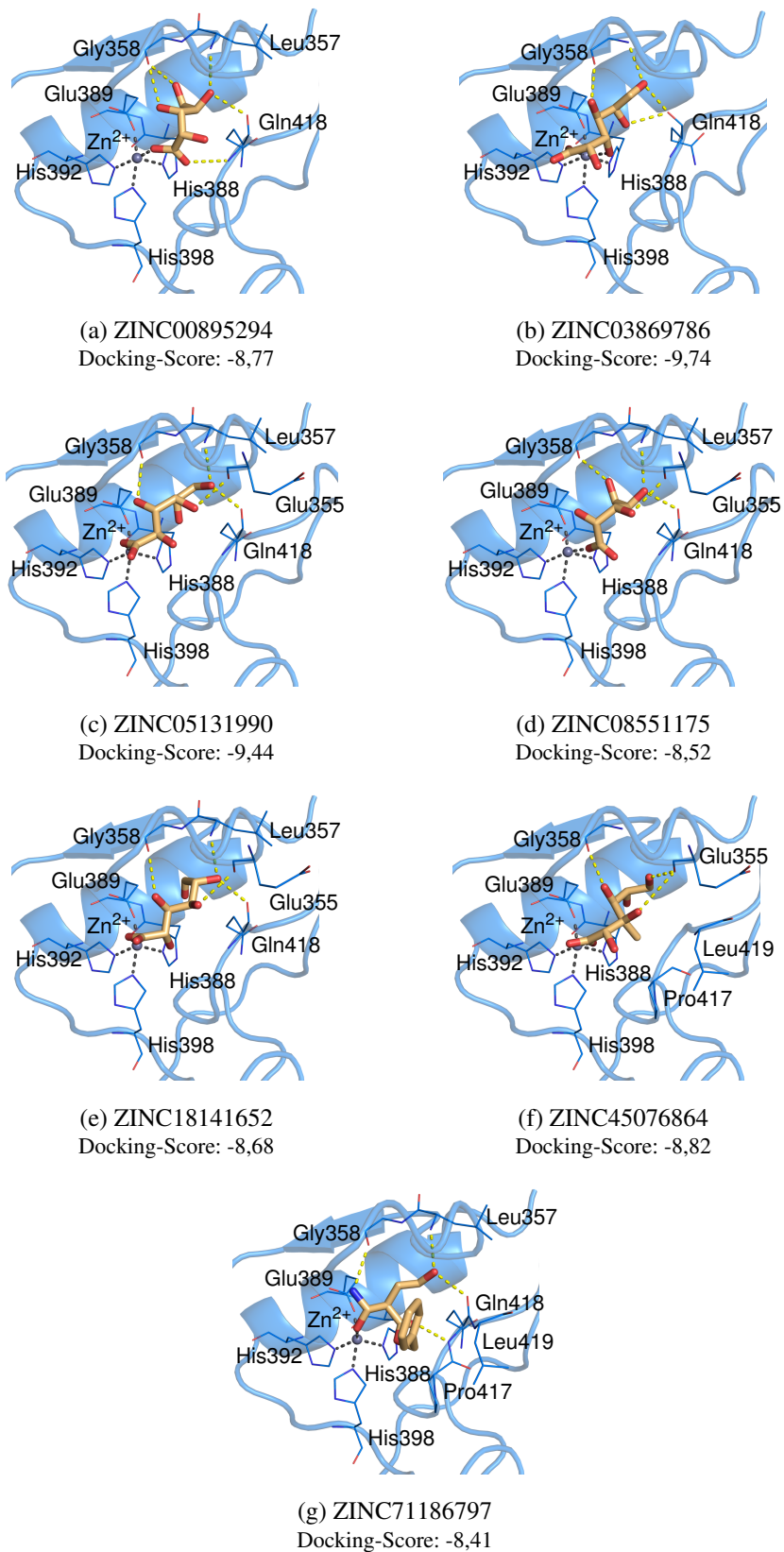


Abbildung 6.5: Docking-Posen der sieben potenziell nicht selektiven ADAMTS7-Inhibitoren. Die Zink-koordinierenden Histidine sind als Linien dargestellt, genau wie die jeweils mit dem Liganden wechselwirkenden Aminosäuren. Die Liganden werden durch orange Stäbchen dargestellt und die Zink-Ionen sind als graue Kugeln zu erkennen. Zink-koordinierende Bindungen sind durch grau gestrichelte Linien dargestellt, Wasserstoffbrücken durch gelb gestrichelte Linien. Die ADAMTS7-Bindetasche (Bereiche um S4, H4 und L1-Loop) ist in Cartoon-Darstellung visualisiert.

koordiniert das Zink-Ion über eine –OH-Gruppe. Alle Wasserstoffbrücken, bis auf jene zu Gly358-O in ZINC71186797, werden über die –OH-Gruppen der Liganden ausgebildet. Die Wasserstoffbrücke zwischen ZINC71186797 und Gly358-O wird über N2 in der –NH₂-Gruppe ausgebildet.

Die Docking-Posen der vier Verbindungen ZINC00895294 (siehe Abb. 6.5a), ZINC05131990 (siehe Abb. 6.5c), ZINC08551175 (siehe Abb. 6.5d) und ZINC18141652 (siehe Abb. 6.5e) mit dem ADAMTS7-Protein weisen ein ähnliches Bindemuster auf. In diesen vier Posen konnte neben der Gly358-O-Wasserstoffbrücke zum Inhibitor (ZINC00895294: O-3 und O-5, ZINC05131990: O-3, ZINC08551175: O-5, ZINC18141652: O-5) jeweils eine Bindung dieser Art zwischen Leu357-N und Inhibitor (ZINC00895294: O-4, ZINC05131990: O-6, ZINC08551175: O-4, ZINC18141652: O-7) beobachtet werden. Außerdem interagieren diese vier Verbindungen (ZINC00895294: O-4, ZINC05131990: O-6, ZINC08551175: O-4 sowie ZINC18141652: O-7) über Gln418-O im L1-Loop mit der Bindetasche. Zusätzlich findet sich in drei der vier Docking-Posen (ZINC05131990: O-4, ZINC08551175: O-3, ZINC18141652: O-6) jeweils eine Wasserstoffbrücke mit Glu355-O. ZINC00895294 bildet mit der zu Glu355 nächstgelegenen –OH-Gruppe bereits zwei Wasserstoffbrücken zu Leu357 und Gln418, sodass keine Interaktion mit Glu355 beobachtet wurde. Des Weiteren treten in den Docking-Posen von ZINC05131990 und ZINC18141652 neben den Wasserstoffbrücken zu Leu357 auch hydrophobe Interaktionen mit dieser Aminosäure auf.

Ein ähnliches Interaktionsmuster wie die vier oben erläuterten Posen zeigt der komplexeste, der potenziell nicht selektiven Inhibitoren, ZINC71186797 (siehe Abb. 6.5g) mit ADAMTS7. Es sind Wasserstoffbrücken zwischen N-1 und Gly358-O, O-3 und Leu357-N, O-3 und Gln418-O sowie O-2 und Gln418-N beobachtet worden. Zusätzlich liegen in dieser Pose hydrophobe Wechselwirkungen zwischen den Aminosäuren Pro417 und Leu419 der Phenylgruppe von ZINC71186797 vor. Diese funktionelle Gruppe findet sich in keiner weiteren Struktur dieser sieben potenziell nicht selektiven Hits. Somit konnte diese Interaktion nur in der Docking-Pose von ZINC71186797 beobachtet werden.

Die Docking-Pose von ZINC03869786 (siehe Abb. 6.5b), das Enantiomer zu ZINC05131990, zeigt, neben der Wasserstoffbrücken zwischen O-4 und Gly358-O, eine weitere Interaktion dieser Art (zwischen O-6 und Gly358-N). Letztere tritt in dieser Pose anstelle der Wasserstoffbrücke zu Leu357 in ZINC05131990 auf. Im L1-Loop wurden zwei Wasserstoffbrücken zwischen zwei –OH-Gruppen (O-5 bzw. O-6) und Gln418-O beobachtet. Somit bilden ZINC03869786 und ZINC05131990 jeweils vier Wasserstoffbrücken über ihre –OH-Gruppen in der ADAMTS7-Pose aus, jedoch in zwei Fällen zu unterschiedlichen Aminosäuren.

In der Docking-Pose von ZINC45076864 (siehe Abb. 6.5f) liegen neben der Wasserstoffbrücke zwischen O-3 und Gly358-O zwei Wasserstoffbrücken zwischen Inhibitor (O-1 bzw. O-7) und Glu355-O vor. ZINC45076864 interagiert zum L1-Loop der ADAMTS7-Bindetasche über hydrophobe Wechselwirkungen zu Pro417 und Leu419. Alle Docking-Posen, außer jene von ZINC45076864 (siehe Abb. 6.5f), zeigen somit Wechselwirkungen zum L1-Loop über eine Wasserstoffbrücke zwischen Ligand und Rückgrat von Gln418.

Insgesamt wird deutlich, dass die nicht selektiven Hits nur mit Aminosäuren in räumlicher Nähe zum Zink-Ion interagieren. Dies sind vor allem die über die ADAMTS-Proteine konservierten Aminosäuren Gly358 und Leu357 sowie die nicht konservierten Aminosäuren Glu355 und Gln418. Die Wechselwirkungen mit Glu355 und Gln418 werden über deren Proteinrückgrate eingegangen. Die drei Aminosäuren Glu355, Leu357 und Gly358 liegen zudem im sequenziell stark konservierten Bereich um das Faltblatt S4. Die in sechs der sieben Docking-Posen vorliegende Wasserstoffbrücke zu Gln418 im L1-Loop scheint in diesem Zusammenhang ebenso kein Merkmal für eine selektive Bindung zu sein. Somit sind diese vier In-

teraktionen unspezifisch und erklären, warum diese Liganden in mindestens 9 der 18 ADAMTS-Proteine im inversen Screening eine gute Bindeaffinität aufweisen.

Folglich scheint die Größe der Liganden in Kombination mit den daraus resultierenden Interaktionen im Bereich um das Zink-Ion eine universelle Bindung an die meisten ADAMTS-Proteine zu ermöglichen.

6.3.3 Wechselwirkungen der potenziell selektiven Hits

Im Folgenden wurden die Wechselwirkungen der Docking-Posen der 18 potenziell selektiven ADAMTS7-Hits untersucht werden. Gibt es gemeinsame Eigenschaften bzw. Unterschiede zwischen den 18 Verbindungen, die eine mögliche Selektivität begründen würden? Lässt sich daraus ein Wechselwirkungsmodell für potenziell selektive ADAMTS7-Inhibitoren ableiten?

In den folgenden Abbildungen 6.6-6.10 sind die Docking-Posen der 18 Moleküle aus dem ADAMTS7-Virtual Screening dargestellt. Zusätzlich sind im Anhang in den Abbildungen A.8a-A.8r die Posen mit den Bindetaschen in Oberflächendarstellung dargestellt. Dies ermöglicht einen räumlichen Eindruck der Bindung sowie die Darstellung der Eigenschaften der Bindetasche.

Untersucht man die Bindemuster der 18 Liganden mit ADAMTS7 zeigen sich verschiedene Interaktionsmuster, neben der Zink-Koordination und der Wasserstoffbrücke zu Gly358-O, die in allen Docking-Posen vorliegen. Im Folgenden werden die Wechselwirkungen der 18 potenziellen ADAMTS7-Inhibitoren in Gruppen mit ähnlichen Bindemustern erläutert. Zudem werden Gemeinsamkeiten und Unterschiede zu den Bindemustern der jeweiligen Liganden im inversen Screening dargestellt.

Die vier Liganden ZINC31961806, ZINC37707318, ZINC52329369 und ZINC75953566 zeichnen sich durch zwei bzw. drei ausgebildete Ringe in der Struktur aus und haben ein ähnliches Bindemuster im Target ADAMTS7 (siehe Abb. 6.6). Die Posen der zwei Liganden ZINC37707318 (siehe Abb. 6.6a) und ZINC52329369 (siehe Abb. 6.6b) zeigen neben der Wasserstoffbrücke zu Gly358-O (ZINC37707318: O-4, ZINC52329369: O-1) keine weiteren Wasserstoffbrücken zum Protein. ZINC75953566 (siehe Abb. 6.6c) weist neben der Wasserstoffbrücke zwischen O-1 und Gly358-O eine weitere Bindung dieser Art zwischen N2 und Gly358-N auf. Außerdem wurden hydrophobe Wechselwirkungen zu Phe384 (ZINC37707318 und ZINC52329369) und zu Leu420 (ZINC37707318) bzw. Leu357 (ZINC52329369) beobachtet. Im inversen Screening zeigen die Docking-Posen von ZINC37707318, dass der Ligand nicht in der Nähe des Zink-Ions bindet. In den Docking-Posen konnte daher keine Zink-Koordination und/oder keine Wasserstoffbrücke zum Gly358-O-Analogen beobachtet werden. Dies stellt wahrscheinlich eine schlechtere Bindung dar und was somit höhere Docking-Scores nach sich zieht. Für ZINC52329369 konnte in allen Docking-Posen im inversen Screening keine Interaktion zur Helix analog zu H4 (ADAMTS7: Phe384) beobachtet werden. Die Wechselwirkungen erfolgen über Wasserstoffbrücken zum L1-Loop-Analogen, was mit weniger guten Docking-Scores bewertet wird und somit auf eine schlechtere Bindeaffinität hinweist. Die Docking-Posen von ZINC75953566 in den anderen ADAMTS-Proteinen zeichnen sich alle durch eine Zink-Koordinierung aus. Eine Wasserstoffbrücke zum Gly358-O-Analogen liegt in 72 % der Fälle vor und es werden meist weitere Wasserstoffbrücken in der Bindetasche ausgebildet. Trotzdem wird im Vergleich der ADAMTS-Proteine die Bindung zu ADAMTS7 am Besten bewertet. Der Ligand ZINC31961806 (siehe Abb. 6.6d) hat genau wie die vorigen drei Hits wenige Wechselwirkungen mit der ADAMTS7-Bindetasche: Zusätzlich zur Zink-Koordinierung und zweier Wasserstoffbrücken zwischen N-5 bzw. N-6 und Gly358-O liegt eine Wasserstoffbrücke zwischen N-6 und Glu389-O_{ε1} vor. Der Inhibitor wird demnach nur in der Umgebung des Zink-Ions an das Protein gebunden. Zusätzlich liegen hydrophobe Wechselwirkungen mit Leu359 und

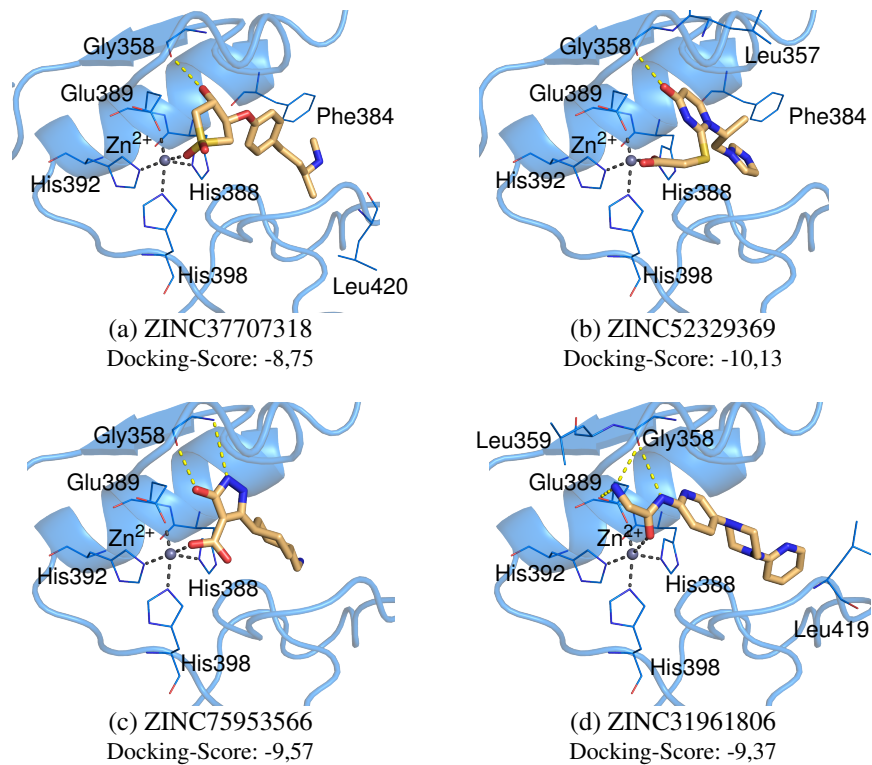


Abbildung 6.6: Docking-Posen der potenziell selektiven Hits ZINC37707318, ZINC52329369, ZINC75953566 und ZINC31961806 in ADAMTS7. Die Zink-koordinierenden Histidine sind als Linien dargestellt, genau wie die jeweils mit dem Liganden wechselwirkenden Aminosäuren. Die Liganden werden durch orange Stäbchen dargestellt und die Zink-Ionen sind als graue Kugeln zu erkennen. Zink-koordinierende Bindungen sind durch grau gestrichelte Linien dargestellt, Wasserstoffbrücken durch gelb gestrichelte Linien. Die Bindetasche (Bereiche um S4, H4 und L1-Loop) ist in Cartoon-Darstellung visualisiert.

Leu419 vor. Im inversen Screening ist der Ligand in einer anderen Position und nicht in der Nähe des Zink-Ions gedockt, sodass keine Zink-Koordination und/oder keine Wasserstoffbrücke zum Gly358-O-Analogon vorliegt.

Die vier ADAMTS7-Docking-Posen in Abbildung 6.6 zeigen somit keine Wasserstoffbrücken mit dem L1-Loop. In jeweils zwei Posen wurden hydrophobe Wechselwirkungen mit der Helix H4 bzw. dem L1-Loop beobachtet. Folglich konnten bei diesen vier Moleküle wenige bis keine weiteren Interaktionen festgestellt werden, auf welchen die Selektivität zu ADAMTS7 begründet sein könnte.

Die beiden potenziellen ADAMTS7-Inhibitoren ZINC48847183 (siehe Abb. 6.7a) und ZINC52012291 (siehe Abb. 6.7b) sind ähnliche Strukturen (Tanimoto-Koeffizient: 0,4) und zeigen ein ähnliches Interaktionsmuster in der ADAMTS7-Bindetasche: Beide Liganden bilden Wasserstoffbrücken zu Gly358-O (ZINC48847183: O-4, ZINC52012291: O-1) sowie Ser416-O (ZINC48847183: O-5, ZINC52012291: N-2) aus. Zusätzlich sind in der ZINC52012291-Pose hydrophobe Wechselwirkungen zur konservierten Aminosäure Leu357 und der nicht konservierten Aminosäure Leu419 vorhanden und in der ZINC48847183-Pose zu Leu357 und Phe384 (konserviert) sowie Pro381 und Leu419 (nicht konserviert). ZINC40603395 (siehe Abb. 6.7c) interagiert in der Bindetasche genau wie die beiden gerade erwähnten potenziellen Inhibitoren über Wasserstoffbrücken zwischen O-5 und Gly358-O sowie O-1 und Ser416-O. Außerdem ist in dieser Docking-Pose eine weitere Wasserstoffbrücke zwischen einer –OH-Gruppe (O-2) und der konservierten Aminosäure Thr385 (Thr385-O_γ1) vorhanden. Ferner liegen zu Leu357, Pro381 und Phe384 hydrophobe Wechselwirkungen vor.

Die Verbindung ZINC48847273 zeigt in ihrer ADAMTS7-Docking-Pose (siehe Abb. 6.7d) ähnliche hydrophoben Wechselwirkungen (Leu357, Phe384) wie ZINC48847173 (vergleiche Abb. 6.7a) sowie eine Wasserstoffbrücke zwischen N-2 und Ser416-O. Die Verbindungen ZINC48847273 und ZINC48847173 weisen mit 0,5 den zweithöchsten Ähnlichkeitswert (vergleiche Tabelle 6.6) der 18 potenziell selektiven ADAMTS7-Inhibitoren auf. ZINC48847273 (O-5, N-2) interagiert mit der ADAMTS7-Bindetasche zusätzlich mit den nicht konservierten Aminosäuren Ile414-O und Gln418-O des L1-Loops über Wasserstoffbrücken. Die Verbindungen ZINC52012291 und ZINC52012042 haben von den 18 betrachteten Verbindungen mit 0,7 den höchsten Ähnlichkeitswert (vergleiche Tabelle 6.6). In ihrem ADAMTS7-Bindemuster unterscheiden sich beide jedoch grundlegend. Während ZINC52012291 über eine Wasserstoffbrücke mit dem L1-Loop (Ser416-O) interagiert, ist zwischen dem potenziellen Hit ZINC52012042 (siehe Abb. 6.7e) und ADAMTS7 neben der Wasserstoffbrücke zwischen O-3 und Gly358-O eine weitere zwischen N-1 und 355-O im S4-Faltblattbereich der Bindetasche beobachtet worden. Überdies liegen hydrophobe Wechselwirkungen zu Leu357, Pro381 und Phe384 im S4- bzw. H4-Bereich sowie zu Pro417 und Tyr421 im L1-Loop vor.

Die Auswertung des inversen Screenings der fünf in Abbildung 6.7 betrachteten Moleküle zeigte Folgendes: In den 18 anderen ADAMTS-Bindetaschen liegt ZINC48847173 räumlich mehr im Bereich des L1-Loop-Analogons und bindet oftmals nicht Zink-koordiniert, sodass die Interaktion mit dem Zink-Ion und/oder dem Gly358-O-Analogon nicht möglich ist (in 14 von 18 Posen). Gleiches gilt für das inverse Screening von ZINC52012291. Die ADAMTS-Docking-Posen der beiden Liganden ZINC48847273 und ZINC40603395 sind in 17 von 18 Posen Zink-koordiniert, jedoch fehlt in 11 bzw. 13 der 18 Bindemuster die Wasserstoffbrücke zum Gly358-O-Analogon. Zusätzlich wird die Bindung im Vergleich zu ADAMTS7 schlechter bewertet. Die Posen von ZINC52012042 im inversen Screening sind alle Zink-koordiniert, jedoch ist nur in dreien die Wasserstoffbrücke zum Gly358-O-Analogon ausgebildet. Der Ligand liegt in

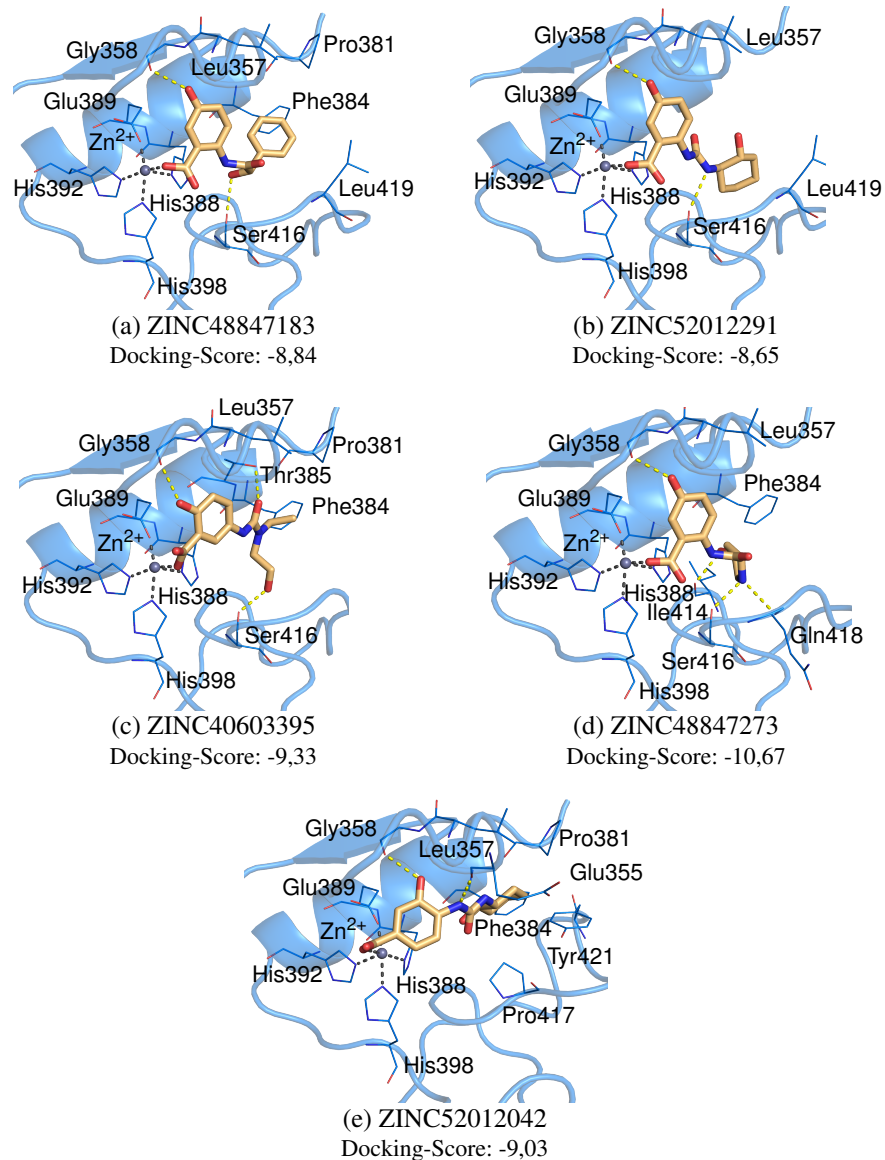


Abbildung 6.7: Docking-Posen der potenziell selektiven Hits ZINC48847183, ZINC52012291, ZINC40603395, ZINC48847273 und ZINC52012042 in ADAMTS7. Die Zink-koordinierenden Histidine sind als Linien dargestellt, genau wie die jeweils mit dem Liganden wechselwirkenden Aminosäuren. Die Liganden werden durch orange Stäbchen dargestellt und die Zink-Ionen sind als graue Kugeln zu erkennen. Zink-koordinierende Bindungen sind durch grau gestrichelte Linien dargestellt, Wasserstoffbrücken durch gelb gestrichelte Linien. Die Bindetasche (Bereiche um S4, H4 und L1-Loop) ist in Cartoon-Darstellung visualisiert.

den meisten Bindetaschen räumlich näher im Bereich des L1-Loop-Analogons, sodass die Wechselwirkung mit der zu Gly358 korrespondierenden Aminosäure nicht ausgebildet werden kann. Zudem weisen 3 der 18 Posen von ZINC52012042, in denen die Wasserstoffbrücken-Bedingung nicht erfüllt wird, eine höhere Bindeaffinität im Vergleich zu ADAMTS7 auf.

Die fünf Bindemuster in Abbildung 6.7 zeigen somit alle hydrophobe Wechselwirkungen zum S4-Bereich (Glu355, Leu357) in der ADAMTS7-Bindetasche. In vier der fünf Posen liegen Wechselwirkungen (hydrophobe, Wasserstoffbrücken) zur Helix H4 (Pro 381, Phe384) vor. Außerdem wurden in vier von fünf Posen Wasserstoffbrücken zum L1-Loop (Ile414, Ser416, Gln418) beobachtet, in drei der Posen hydrophobe Wechselwirkungen (Pro417, Leu419, Tyr412) zu diesem Bereich. Diese fünf potenziell selektiven Hits zeigen somit eine Vielzahl an Wechselwirkungen in der Bindetasche in allen drei Bereichen dieser.

Die Verbindungen ZINC40470996 und ZINC49571688 zeigen in der ADAMTS7-Bindetasche ein identisches Bindemuster: In der Pose von ZINC40470996 (siehe Abb. 6.8a) wurden neben der Wasserstoffbrücke zwischen O-4 und Gly358-O zwei Wasserstoffbrücken zwischen N-2 und Ile414-O sowie Ser416-O im L1-Loop beobachtet. Die gleichen Wechselwirkungen wurden in der Docking-Pose von ZINC49571688 (siehe Abb. 6.8b) identifiziert: Zwischen O-2 und Gly358-O sowie zwischen N-1 und Ile414-O bzw. N-2 und Ser416-O liegen Bindungen dieser Art vor. Zusätzlich wurden in beiden Docking-Posen hydrophobe Wechselwirkungen mit der konservierten Aminosäure Phe384 in der Helix H4 beobachtet.

Im inversen Screening fehlten in den Docking-Posen mit dem potenziellen Inhibitor ZINC40470996 die Wasserstoffbrücke zum Gly358-O-Analogon oder die Liganden interagierten nicht mit dem Loop analog zu L1, was zu einer schlechteren Bindeaffinität führte. Das inverse Screening von ZINC49571688 zeigte in 17 der 18 Posen eine Zink-Koordination, in 13 der 18 Protein-Ligand-Konformationen fehlt die Interaktion mit dem Gly358-O-Analogon. Der Ligand bindet dabei räumlich näher im Bereich des L1-Loop-Analogons.

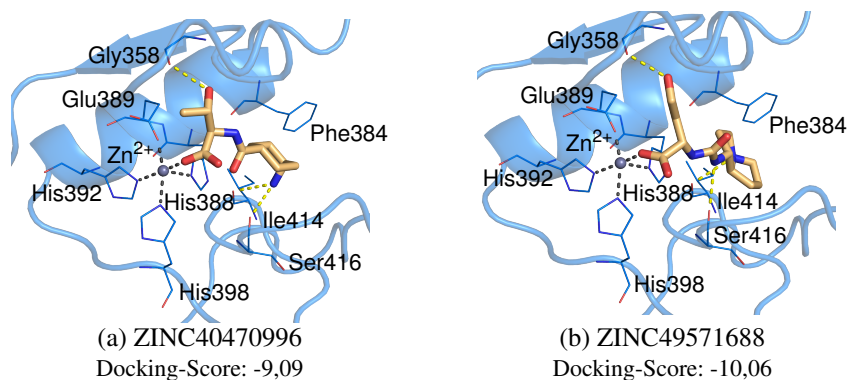


Abbildung 6.8: Docking-Posen der potenziell selektiven Hits ZINC40470996 und ZINC49571688 in ADAMTS7. Die Zink-koordinierenden Histidine sind als Linien dargestellt, genau wie die jeweils mit dem Liganden wechselwirkenden Aminosäuren. Die Liganden werden durch orange Stäbchen dargestellt und die Zink-Ionen sind als graue Kugeln zu erkennen. Zink-koordinierende Bindungen sind durch grau gestrichelte Linien dargestellt, Wasserstoffbrücken durch gelb gestrichelte Linien. Die Bindetasche (Bereiche um S4, H4 und L1-Loop) ist in Cartoon-Darstellung visualisiert.

Folglich konnten für die potenziell selektiven Hits ZINC40470996 und ZINC49571688 auch Wechselwirkungen in allen drei Bereichen (S4-Faltblatt, H4-Helix und L1-Loop) der Bindetasche beobachtet werden.

Die vier Verbindungen ZINC80583404, ZINC40558343, ZINC94569517 und ZINC9456992 zeigen mitunter Ähnlichkeiten im Bindemuster mit ADAMTS7: In der Docking-Pose von ZINC80583404 (siehe Abb. 6.9a) ist neben der Wasserstoffbrücke zwischen O-5 und Gly358-O eine Wasserstoffbrücke zwischen Gln418-O und N-2 vorhanden. Diese Interaktion wurde bereits gehäuft in den nicht selektiven Hits beob-

achtet (siehe Abschnitt 6.3.2). Zusätzlich kommt es zu hydrophoben Wechselwirkungen mit dem Leu357. Die gleiche Kombination der Wechselwirkungen wurde bei den potenziell nicht selektiven Hits (vergleiche Abb. 6.5) nicht beobachtet. Im inversen Screening wechselwirkt der Ligand häufiger mit anderen Aminosäuren des L1-Loop-Analogons und bindet daher räumlich näher in diesem Bereich, sodass keine Wasserstoffbrücke zwischen Gly358-O-Analogon und Ligand ausgebildet werden kann. Daher ist die Wasserstoffbrücke zu Gly358-O in diesem Liganden wahrscheinlich für die gute Bindeaffinität bei ADAMTS7 verantwortlich.

Neben einer Wasserstoffbrücke zu Gln418-O findet sich in den Docking-Posen von ZINC40558343 (siehe Abb. 6.9b) und ZINC94569517 (siehe Abb. 6.9c) zur ADAMTS7-Bindetasche noch eine Bindung dieser Art zwischen einem Stickstoff-Atom des Inhibitors (N-3 bzw. N-2) und Phe413-O bzw. Ser416-O und Phe413-O, demnach mit dem Rückgrat des L1-Loops. Außerdem ist in der ZINC40558343-Pose eine Wasserstoffbrücke zwischen einer -OH-Gruppe (O-1) und der konservierten Aminosäure Thr385-O_γ ausgebildet. Hydrophobe Wechselwirkungen wurden in beiden Posen zu Phe384 beobachtet und zu Leu357 nur in ZINC94569517. Im inversen Screening wurde bei ZINC40558343 und ZINC94569517 das gleiche Verhalten des Liganden wie bei ZINC80583404 beobachtet: Das Molekül liegt räumlich näher zum L1-Loop-Analogon und damit bindet eher auf der Zink-abgewandten Seite der Bindetasche. Daher kann keine Wasserstoffbrückenbindung zum Gly358-Analogon ausgebildet werden. Scheinbar ist die Bindung in der Nähe des Zink-Ions wichtig für eine gute Bindeaffinität. Die ADAMTS7-Docking-Pose von ZINC94569992 offenbart drei Wasserstoffbrücken zwischen Inhibitor und L1-Loop von ADAMTS7. Diese liegen zwischen N-1 und Phe413-O, Ile414-O sowie Ser416-O vor. Zusätzlich sind in dieser Pose hydrophobe Wechselwirkungen mit Phe384 vorhanden. Im inversen Screening zeigte sich für diesen Liganden, dass 17 der 18 Posen Zink-koordiniert sind und nur in vier Posen die Wasserstoffbrücke zum Gly358-O-Analogon vorlag. Die Bindung erfolgt auch hier zumeist eher in Richtung des L1-Loops, sodass die Interaktion zum Glycin nicht eingegangen wird. Gleichzeitig wird die Bindeaffinität von 17 der 18 Posen schlechter bewertet als bei ADAMTS7.

Die vier in Abbildung 6.9 dargestellten Bindemuster lassen vermuten, dass die Kombination aus Wechselwirkungen mit der Helix H4 (Phe384) und dem L1-Loop (Phe413, Ile414, Ser416, Gln418) zu einer guten Bindeaffinität und damit zu einer Selektivität für ADAMTS7 verantwortlich sein könnten. Dabei ist kein eindeutiges identisches Bindemuster abzuleiten, was natürlich auch durch die verschiedenen Strukturen der Inhibitoren begründet sein kann.

Ein vielfältiges Bindemuster zum L1-Loop weisen auch die drei Bindemuster von ZINC49587694, ZINC70254898 und ZINC86853423 in ADAMTS7 auf. In der ADAMTS7-Docking-Pose des potenziellen Inhibitors ZINC49587694 (siehe Abb. 6.10a) wurden neben der Wasserstoffbrücke zwischen N-3 und Gly358 zwei Wasserstoffbrücken zwischen Inhibitor (N-2, N-2) und L1-Loop (Pro417-O, Leu420-O) beobachtet. Außerdem liegen in dieser Pose aromatische sowie hydrophobe Wechselwirkungen mit dem Aminosäurerest Phe348 vor. Außerdem sind hydrophobe Wechselwirkungen mit Leu419 beobachtet worden. Dies scheint bei der Bindung zu ADAMTS7 eine gute Affinität hervorzurufen. Im inversen Screening liegt der Ligand in den weiteren ADAMTS-Proteinen mehr in räumlicher Nähe zum L1-Loop-Analogon hin. Daher fehlt in 4 der 18 Posen die Zink-Koordinierung und in 10 der 18 Posen dazu noch die Wasserstoffbrücke zum Gly358-O-Analogon. Die fehlenden Wechselwirkungen im Bereich des Zink-Ions resultieren scheinbar in einer schlechteren Bindeaffinität zu den ADAMTS-Proteinen. Die ADAMTS7-Docking-Pose mit ZINC70254898 (siehe Abb. 6.10b) zeigt eine Wasserstoffbrücke zwischen N-1 und Tyr421-OH und ZINC70254898 interagiert außerdem über hydrophobe Wechselwirkungen mit Pro381, Pro417 und Tyr421

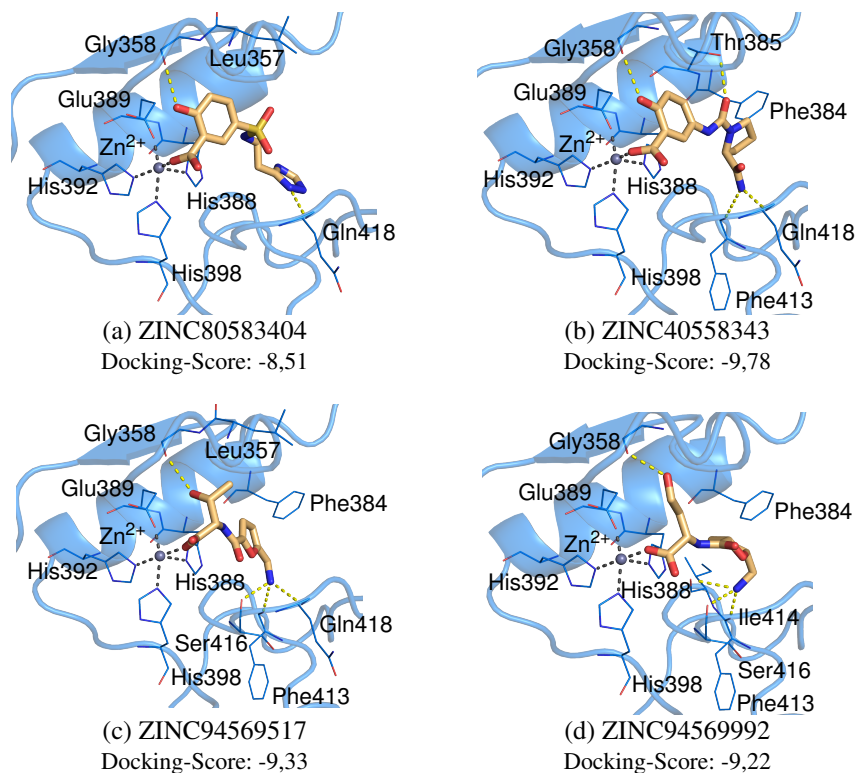


Abbildung 6.9: Docking-Posen der potenziell selektiven Hits ZINC80583404, ZINC40558343, ZINC94569517 und ZINC94569992 in ADAMTS7. Die Zink-koordinierenden Histidine sind als Linien dargestellt, genau wie die jeweils mit dem Liganden wechselwirkenden Aminosäuren. Die Liganden werden durch orange Stäbchen dargestellt und die Zink-Ionen sind als graue Kugeln zu erkennen. Zink-koordinierende Bindungen sind durch grau gestrichelte Linien dargestellt, Wasserstoffbrücken durch gelb gestrichelte Linien. Die Bindetasche (Bereiche um S4, H4 und L1-Loop) ist in Cartoon-Darstellung visualisiert.

mit der Bindetasche. Im inversen Screening fehlt in etwa der Hälfte der Bindemuster die Wasserstoffbrücke zum Gly358-O-Analogon. Zudem wurden in den meisten Docking-Posen keine Wechselwirkungen mit dem L1-Loop-Analogon beobachtet, sodass die Bindeaffinität deutlich schlechter als in der ADAMTS7-Pose bewertet wird. In der Docking-Pose der Verbindung ZINC86853423 (siehe Abb. 6.10c) wurden zwei Wasserstoffbrücken zwischen ADAMTS7-Bindetasche (Pro417-O, Tyr421-OH) und Inhibitor (N-2) sowie hydrophobe Wechselwirkungen mit Leu357, Pro417 und Tyr421 beobachtet. In inversen Screening fehlt in den Docking-Posen die wahrscheinlich essenzielle Interaktion mit dem Gly358-Analogon oder überhaupt eine Wechselwirkung nahe dem Zink-Ion.

Die drei Bindemuster in Abbildung 6.10 weisen alle Wasserstoffbrücken zum L1-Loop (Pro417, Leu420, Tyr421) auf. Hydrophobe Wechselwirkungen zur Helix H4 wurden in zwei Posen beobachtet, zum L1-Loop in allen drei Docking-Posen.

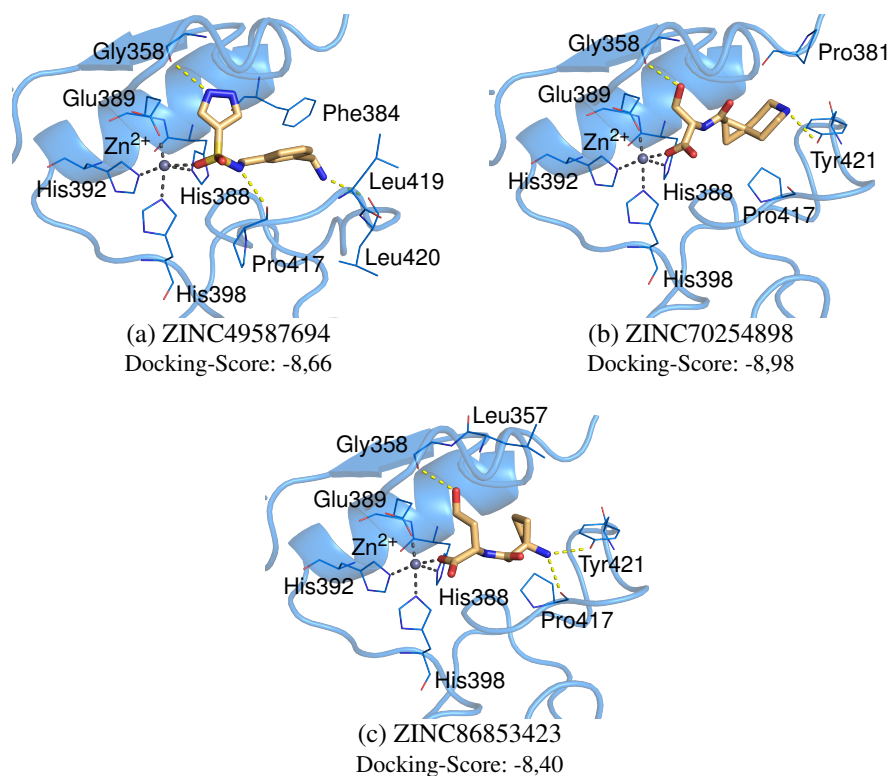


Abbildung 6.10: Docking-Posen der potenziell selektiven Hits ZINC49587694, ZINC70254898 und ZINC86853423 in ADAMTS7. Die Zink-koordinierenden Histidine sind als Linien dargestellt, genau wie die jeweils mit dem Liganden wechselwirkenden Aminosäuren. Die Liganden werden durch orange Stäbchen dargestellt und die Zink-Ionen sind als graue Kugeln zu erkennen. Zink-koordinierende Bindungen sind durch grau gestrichelte Linien dargestellt, Wasserstoffbrücken durch gelb gestrichelte Linien. Die Bindetasche (Bereiche um S4, H4 und L1-Loop) ist in Cartoon-Darstellung visualisiert.

Anhand der 18 analysierten Bindungsmuster konnte gezeigt werden, dass eine Vielzahl an Kombinationen von Interaktionen in der Bindetasche wahrscheinlich eine Selektivität für ADAMTS7 auszeichnen könnten. Interessant ist, dass im Vergleich zum inversen Screening die Wechselwirkung mit dem Aminosäurerest Gly358 für eine bessere Bindeaffinität zu sorgen scheint. Somit ist eine Donorstelle im Liganden, wie im Pharmacophore (D4) postuliert, eventuell essenziell, jedoch wahrscheinlich für keine selektive Interaktion verantwortlich. Wasserstoffbrücken zwischen den Inhibitoren und den Gly358- und Leu357-Analoga wurden bereits in den Röntgenkristallstrukturen von ADAMTS1, -4 und -5 beobachtet (vgl. Abschnitt 4.3). Außerdem sind Interaktionen mit dem L1-Loop potenziell wichtig für eine selektive Bindung. Obwohl kei-

ne spezifische Wechselwirkung ausgemacht werden konnte, wurden Wasserstoffbrücken zu Phe413, Ser416, Pro417, Leu420 und Tyr421 bei den nicht selektiven Hits nicht beobachtet. Donor- oder Akzeptor-Regionen des Liganden könnten in diesen Regionen somit eine Selektivität für ADAMTS7 in Kombination mit der Interaktion zu Gly358 darstellen. So ist wahrscheinlich eine Kombination aus der Zink-Koordination, der Bindung an Gly358-O und Wasserstoffbrücken zum L1-Loop (außer Gln418) entscheidend für eine hohe Bindeaffinität und wahrscheinlich auch selektive Bindung zu ADAMTS7. Somit lässt sich aus den Protein-Inhibitor-Interaktionen kein einheitliches Bindemuster für die potenziell selektiven ADAMTS7-Inhibitoren ableiten. Jedoch wird auch hier die Wichtigkeit der Interaktionen mit dem L1-Loop des Proteins deutlich.

6.3.4 MD-Simulationen der potenziell selektiven Hits

Nach der Untersuchung der Bindemuster der 18 potenziell selektiven ADAMTS7-Inhibitoren wurden die Protein-Inhibitor-Posen mittels kurzer MD-Simulationen untersucht. Diese sollten zeigen, ob die Interaktionen zwischen ADAMTS7-Bindetasche und Ligand stabil sind. Als Startstruktur für das ADAMTS7-Protein wurde jeweils das Homologiemodell oder das jeweilige Conformation Ensemble gewählt, in welchem die Pose vorlag. Dabei handelt es sich um das ADAMTS7-Homologiemodell für die 15 Verbindungen ZINC31961806, ZINC37707318, ZINC40470996, ZINC40558343, ZINC40603395, ZINC48847183, ZINC48847273, ZINC49571688, ZINC49587694, ZINC52012291, ZINC52329369 sowie ZINC75953566, ZINC80583404, ZINC94569517 und ZINC94569992 und Conformational Ensemble K2 für die drei Verbindungen ZINC52012042, ZINC70254898 und ZINC86853423.

Im ersten Schritt wurde die jeweilige Protein-Inhibitor-Konformation mit Schrödinger MacroModel (Version 10.4) und dem Kraftfeld OPLS2005 energieminiert. Die MD-Simulationen wurden mit dem Desmond [121]-Paket (Desmond Molecular Dynamics System, Version 3.6; Maestro-Desmond Interoperability Tools, Version 3.6) mit Standardeinstellungen und dem Kraftfeld OPLS2005 [104] durchgeführt. Die Parameter für die MD-Simulationen wurden analog zur Simulation des ADAMTS7-Homologiemodells (siehe Abschnitt 3.3.5) ausgewählt, um eine Vergleichbarkeit zu ermöglichen. Die 5 ns lange MD-Simulation wurde bei einer Temperatur von 300 K und 1 bar Druck durchgeführt. Sie erfolgte somit unter thermodynamischen Gleichgewichtsbedingungen bei konstantem Druck und konstanter Temperatur (*NPT-Ensemble*). Die nur 5 ns lange MD-Simulation sollte in diesem Fall ausreichen, um zu validieren, ob die Docking-Posen stabil sind.

Zur Auswertung der MD-Simulationen wurde jeweils der $\text{RMSD}_{C\alpha}$ in Ångström über die Simulationszeit im Vergleich zur Startstruktur gemessen (siehe Abb. 6.11a-6.11r). Die Analyse des $\text{RMSD}_{C\alpha}$ der Proteinstruktur (blau) zeigte für alle 18 Simulationen nach einer Äquilibrationsphase von etwa 4 ns einen stabilen Wert, welcher um weniger als 1 Å schwankt. Der $\text{RMSD}_{C\alpha}$ -Wert konvergiert gegen 1,8 Å für ZINC48847183, gegen 2,0 Å für ZINC48847273, gegen 2,1 Å für ZINC94569992, gegen 2,3 Å für ZINC40470996, ZINC40603395, ZINC75953566 und ZINC86853423, gegen 2,4 Å für ZINC31961806 und gegen 2,5 Å für ZINC49587694, ZINC52012042 und ZINC80583404. Für die Simulation von ZINC49571688 und ZINC94569517 konvergiert der $\text{RMSD}_{C\alpha}$ gegen 2,6 Å, für ZINC52012291 gegen 2,7 Å, für ZINC37707318, ZINC40558343 und ZINC70254898 gegen 2,8 Å und für ZINC52329369 gegen 4,2 Å. Auch die Analyse des $\text{RMSF}_{C\alpha}$ für die 18 MD-Simulationen zeigte nur für Loop-Regionen höhere Abweichungen ($\text{RMSF}_{C\alpha} > 1,5 \text{ \AA}$) auf (Daten nicht gezeigt). Die Beobachtungen in den Simulation deuten auf ein natives Verhalten der Komplexe hin. Folglich haben die Komplexe einen stabilen Zustand eingenommen und die Protein-Inhibitor-Interaktionen (von 4 ns – 5 ns der MD) konnten mit der Docking-Pose verglichen werden.

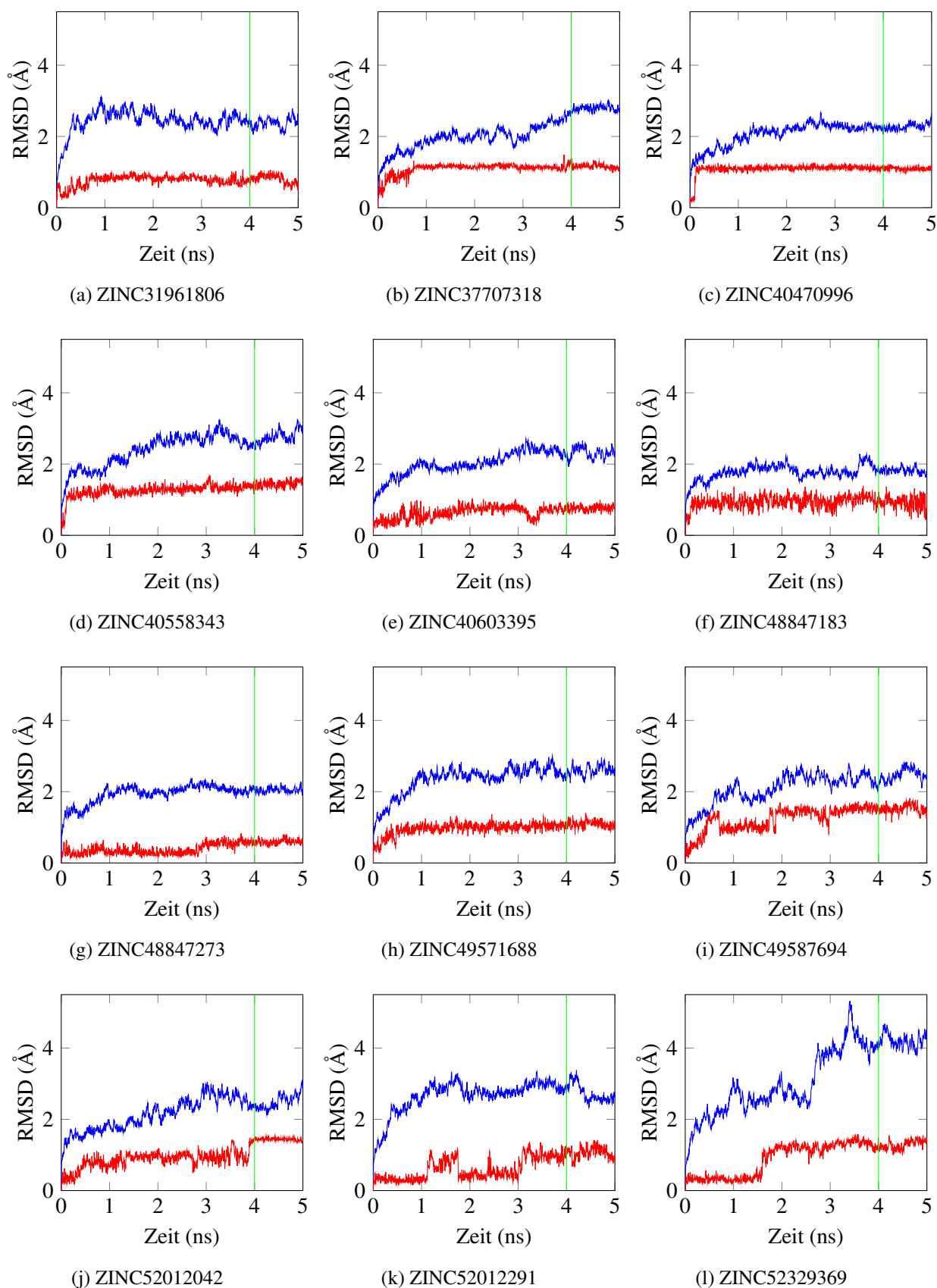


Abbildung 6.11: $\text{RMSD}_{\text{C}\alpha}$ des Proteins (blau) während der 5 ns dauernden MD-Simulationen der ADAMTS7-Docking-Posen im Vergleich zur Startstruktur. Der RMSD der Liganden im Verlauf der MD-Simulation ist als roter Graph gekennzeichnet. Der Beginn des stabilen Zeitraums bei 4 ns ist durch eine grüne, vertikale Linie markiert.

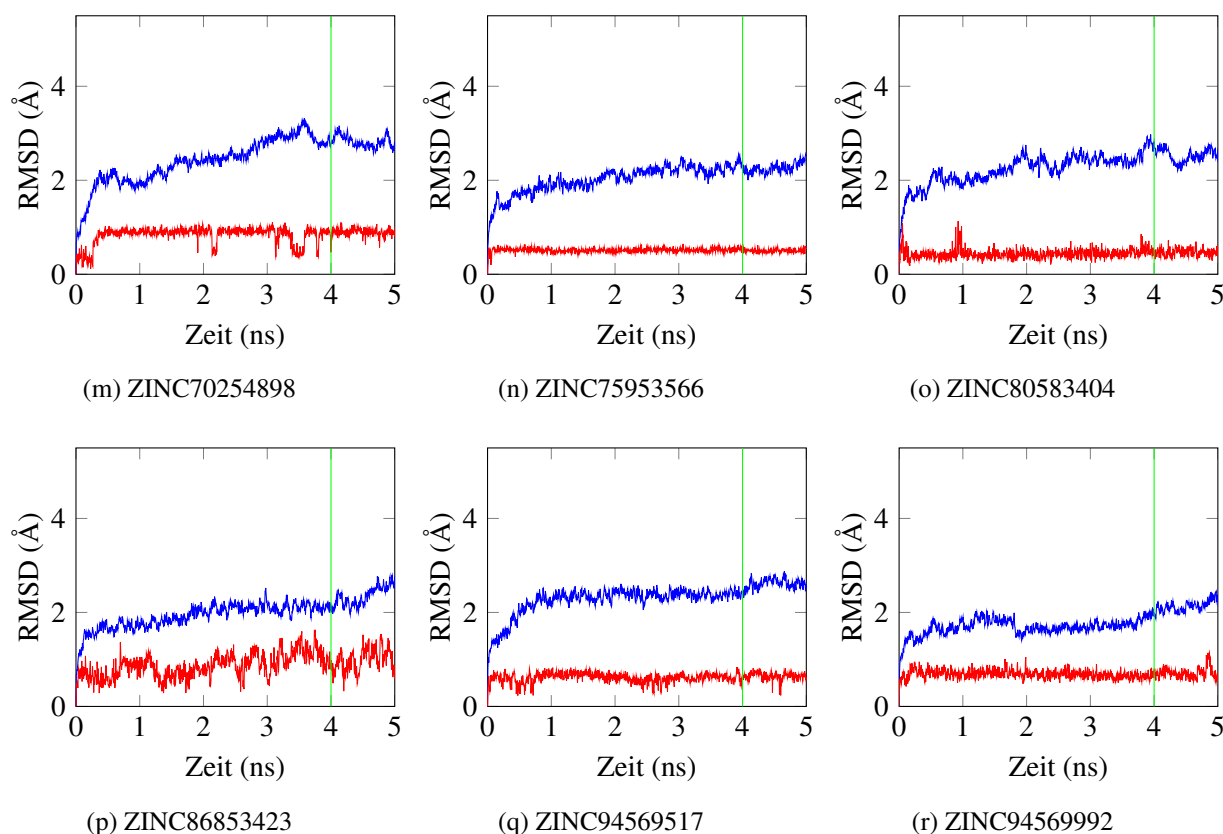


Abbildung 6.11 (Fortsetzung): $\text{RMSD}_{\text{C}\alpha}$ des Proteins (blau) während der 5 ns dauernden MD-Simulationen der ADAMTS7-Docking-Posen im Vergleich zur Startstruktur. Der RMSD der Liganden im Verlauf der MD-Simulation ist als roter Graph gekennzeichnet. Der Beginn des stabilen Zeitraums bei 4 ns ist durch eine grüne, vertikale Linie markiert.

Ein Blick auf den RMSD der potenziellen ADAMTS7-Inhibitoren (siehe Abb. 6.11, rote Graphen) im Verlauf der MD macht deutlich, dass sich die Position der fünf Verbindungen ZINC31961806, ZINC48847273, ZINC70254898, ZINC75953566 und ZINC94569517 nur minimal ändert ($\text{RMSD} \leq 1 \text{ \AA}$). Die RMSD-Werte von insgesamt zehn Liganden (ZINC37707318, ZINC40470996, ZINC40603395, ZINC48847183, ZINC49571688, ZINC52012291, ZINC52329369, ZINC80583404, ZINC86853423 und ZINC94569992) schwankt bis etwa $1,5 \text{ \AA}$. Des Weiteren beträgt der RMSD der drei potenziellen Inhibitoren ZINC40558343, ZINC49587694, ZINC52012042 weniger als $1,8 \text{ \AA}$. Dies macht deutlich, dass die Liganden stabil in der Bindetasche liegen. Protein-Inhibitor-Wechselwirkungen wurden für den Zeitraum von 4-5 ns ausgewertet. Wie oben festgestellt, nehmen die Proteinstrukturen in dieser Zeit eine stabile Konformation ein. Im Folgenden beziehen sich deshalb alle prozentualen Angaben zur Häufigkeit von Interaktionen nur auf den stabilen Auswertungszeitraum von 4-5 ns. Da keine experimentellen Daten über ADAMTS7-Inhibitoren und Bindemuster vorliegen, kann nicht abgeschätzt werden, wie stabil und somit wie stark eine Interaktion sein muss. Daher wird hier kein Grenzwert bei der Häufigkeit zur Abschätzung für eine stabile Interaktion festgelegt.

In allen Simulationen liegt die Zink-Koordination über den Liganden vor. Die Wasserstoffbrücken der 18 potenziellen Hits sowie deren Häufigkeit in den MD-Simulationen sind in Tabelle 6.7 zusammengefasst. Als Wasserstoffbrücke wurde eine Kombination aus Donor-(D), Wasserstoff-(H) und Akzeptor-(A) Atomen mit einer (D-H...A-X)-Konfiguration definiert, wenn die Distanz zwischen D und A $\leq 2,5 \text{ \AA}$, der Winkel D-H-A $\geq 120^\circ$ und der Winkel H-A-X $\geq 90^\circ$ war. In der Simulation von ZINC31961806 ist die Wasserstoffbrücke zwischen N-5 und Gly358-O in 96 % der Auswertungszeit vorhanden. Die Interaktion

aus der Docking-Pose mit Glu389 wurde nicht beobachtet, dafür aber aromatische Wechselwirkungen mit Phe384 (17 %) und hydrophobe Wechselwirkungen mit Leu419 (11 %). Die Simulation von ZINC37707318 zeigt die gleiche Interaktion auf, wie die Docking-Pose. Die Wasserstoffbrücke zwischen O-4 und Gly358-O wurde in 44 % der Zeit beobachtet, zusätzlich fanden sich noch eine Wasserstoffbrücke zwischen O-4 und Leu357-N (33 %) und zwischen N-1 und Asp422-O_{δ2} (80 %) im L1-Loop. Die Auswertung der Simulation für ZINC40470996 zeigte ein ähnliches Bindemuster wie in der Docking-Pose: Die Gly358-O-Wasserstoffbrücke mit O-4 (43 %), sowie jene zwischen N-2 und Ile414-O (83 %) bzw. Ser416-O (30 %) sind vorhanden. Des Weiteren wurden zwei weitere Bindungen dieser Art beobachtet: Eine zwischen O-4 und Leu357-N (64 %) und eine zwischen O-3 und Thr385-O_{γ1} (92 %).

Die Simulation von ADAMTS7 mit ZINC40558343 offenbarte ein Bindemuster, welches fast identisch zur Docking-Pose ist. Einzig die Wasserstoffbrücke zu Phe413-O fehlt. Es zeigen sich zudem noch zusätzliche Wasserstoffbrücken zwischen O-5 und Leu357-N (90 %), zwischen N-3 und Ser416-O (10 %) und zwischen N-3 und Pro417-O (95 %). Hier wird die Pose wahrscheinlich durch die anderen beiden Wasserstoffbrücken im L1-Loop stabilisiert. Die Bindung von ZINC40603395 zu ADAMTS7 veränderte sich in der Simulation. Die wahrscheinlich essenzielle Wasserstoffbrücke mit Gly358-O ist nicht vorhanden und in der Region um das Zink-Ion ist keine Bindung dieser Art beobachtet worden. Stabil bleibt die Interaktion zwischen O-1 und Ser416-O (77 %). Das Bindemuster des potenziellen Inhibitors ZINC48847183 bleibt in der MD-Simulation erhalten. Zusätzlich fanden sich zwei Wasserstoffbrücken zwischen O-4 und Leu357-N (86 %) bzw. Gly358-N (66 %). Die Simulation von ZINC48847273 zeigte, dass in der Bindetasche die Interaktion mit Gly358-O verloren geht. Eine (alternative) Wasserstoffbrücke findet sich in dem Bereich um S4 zwischen O-1 und Glu355-O (79 %). Die Interaktion im L1-Loop zwischen N-2 und Ser416-O ist vorhanden, jene zu Ile414-O und Gln418-O nicht. Diese werden „ersetzt“ durch zwei Wasserstoffbrücken zwischen N-2 (53 %) sowie N-1 (100 %) und Pro417-O im L1-Loop.

In der Simulation von ZINC49571688 sind keine Wasserstoffbrücken zum L1-Loop beobachtet worden (außer hydrophobe Wechselwirkungen zu Ile414 (1 %), Pro417 (24 %), Leu419 (5 %)). Die wahrscheinlich essenzielle Interaktion zu Gly358-O über O-2 liegt vor (53 %) und wird durch zwei weitere Wasserstoffbrücken (zwischen O-2 und Leu357-N bzw. Gly358-N) in diesem Bereich ergänzt. In der Simulation von ZINC49587694 wurde beobachtet, dass die Wasserstoffbrücke zu Gly358-O verschwindet. An deren Stelle treten Wasserstoffbrücken zwischen Leu357-N und N-3 (24 %) sowie zwischen Glu355-O und N-4 (64 %). Die Interaktion von ZINC49587694 mit Pro417 geht ebenfalls verloren und wird durch eine Wasserstoffbrücke zwischen N-1 und Leu420-O (48 %) bzw. Asp422-O_{δ2} (99 %) ersetzt. Die aromatischen Wechselwirkungen zu Phe384 (35 %) bleiben bestehen und es sind hydrophobe Wechselwirkungen zu Leu419 beobachtet worden. Der Position des Liganden in der MD ist demzufolge etwas verschoben.

Das Bindemuster der Docking-Pose von ZINC52012042 wurde in der MD-Simulation beobachtet. Zusätzlich fand sich noch eine Wasserstoffbrücke zwischen O-3 und Leu357-N. Dieser potenzielle Inhibitor zeigt in der Docking-Pose und in der MD-Simulation keine Interaktion mit dem L1-Loop. Die Untersuchung der Bindung von ZINC52012291 zeigt eine stabile Koordination in der Umgebung des Zink-Ions. Die Wasserstoffbrücke mit Ser416-O aus der Docking-Pose ist nicht vorhanden, dafür zwei Interaktionen dieser Art im L1-Loop zwischen O-5 und Gln418-N (70 %) bzw. Leu420-O (57 %). Der Ligand verschiebt sich somit im L1-Loop, wobei die Interaktionen um das Zink-Ion weiterhin vorhanden sind. ZINC52329369 scheint in der MD-Simulation die Interaktion zu Gly358-O zu verlieren, welche durch eine Wasserstoffbrücke zwischen O-1 und Glu355-O (84 %) ersetzt wird. Ein ähnliches Bild zeichnet die Simulation von

Tabelle 6.7: Beobachtete Wasserstoffbrückenbindungen in den Simulationen der ADAMTS7-Inhibitor-Komplexe mit Häufigkeit (%) in stabiler Simulationszeit. * - Wasserstoffbrücke ist in Docking-Pose vorhanden. Nicht beobachtete Interaktionen werden aus Gründen der Übersichtlichkeit nicht mit 0 gekennzeichnet. Eine Ausnahme bilden Wasserstoffbrücken, welche in der Docking-Pose beobachtet wurden. Diese sind als 0* aufgeführt.

Inhibitor	Wasserstoffbrückenbindungen zu ADAMTS7-Atom (Häufigkeit in %)													
	Glu355-O	Leu357-N	Gly358-O	Gly358-N	Thr385-O _{γ1}	Glu389-O _{ε1}	Phe413-O	Ile414-O	Ser416-O	Pro417-O	Gln418-N	Leu420-O	Tyr421-OH	Asp422-O _{δ2}
ZINC31961806			96*			0*								
ZINC37707318		33	44*											80
ZINC40470996		64	43*		92		2	83*	30*		3			
ZINC40558343		90	36*		100*				10	95	24*			
ZINC40603395							2	77*						
ZINC48847183		86	34*	66			1	97*						
ZINC48847273	79				67			59*	100;53					
ZINC49571688		94	53*	38										
ZINC49587694	64	24	0*						0*		48*			99
ZINC52012042	96*	85	25*											
ZINC52012291		25	43*	11	36						70	57		
ZINC52329369	84		0*									1		
ZINC70254898		96	0*	69									1*	
ZINC75953566		97	68*	2							100			98
ZINC80583404		77	26*	2	98			4			89*	59		
ZINC86853423		98	42*	28		6				0*			0*	
ZINC94569517		72	14*					63	55*		99*			97
ZINC94569992		93	48*	41				18*	74*		87			93

ZINC70254898: Auch hier ist die Wasserstoffbrücke zu Gly358-O nicht vorhanden, dafür aber zwei Wasserstoffbrücken: Zwischen O-2 und Leu357-N (96 %) bzw. Gly358-N (69 %).

Die MD-Simulation des potenziellen Inhibitors ZINC75953566 zeigt zusätzlich zu den Interaktionen in der Docking-Pose (Zink-Koordinierung und Wasserstoffbrücke Gly358-O) Wechselwirkungen mit dem S4- und dem L1-Loop-Bereich. Dabei handelt es sich um Wasserstoffbrücken zwischen N-2 und Leu357-N (97 %), zwischen N-1 und Gln418-N (100 %) bzw. Asp422-O_{δ2} (98 %). In der Simulation der Verbindung ZINC80583404 wurden alle Interaktionen der Docking-Pose beobachtet. Zusätzlich sind weitere fünf Wasserstoffbrücken (zwischen O-5 und Leu357-N bzw. Gly358-N; N-1 und Thr385-O_{γ1}, N-2 und Ser416-O; N-3 und Leu420-O) vorhanden, welche die Bindung zu stabilisieren scheinen. ZINC86853423 zeigt in der Simulation eine stärkere Bindung in der Umgebung des Zink-Ions. Zwei zusätzliche Wasserstoffbrücken zwischen O2 und Leu357-N bzw. Gly358-N treten hier auf, dabei gehen die Interaktionen im L1-Loop (Pro417, Tyr421) verloren. Es liegen zusätzlich hydrophobe Wechselwirkungen zu Pro417 vor, aber die Position ist im Vergleich zur Docking-Pose ein wenig gedreht. Die MD-Simulation von ZINC94569517 zeigt, dass die Interaktionen der Docking-Pose erhalten bleiben. Es wurden drei zusätzliche Wasserstoffbrücken zwischen O-5 und Leu357-N, N-2 und Ile414-O sowie N-2 und Asp422 im L1-Loop beobachtet. Die Analyse des Bindemusters von ZINC94569992 während der MD-Simulation zeigte konstante Wechselwirkungen zwischen Docking-Pose und Simulation (bis auf Phe413). Zusätzlich treten vier weitere Wasserstoffbrücken auf, von denen zwei im L1-Loop liegen (zwischen N-1 und Gln418-N bzw. Asp422-O_{δ2}).

Insgesamt ist die Zink-Koordinierung in allen Simulationen vorhanden. Die potenziell essenzielle Wasserstoffbrücke zwischen Inhibitor und Gly358-O wurde in 13 der 18 Simulationen beobachtet, wobei interessant ist, dass in allen Simulationen, außer bei ZINC31961806 und ZINC40603395, andere oder weitere

Interaktionen (Wasserstoffbrücken zwischen Inhibitor und Glu355-O, Leu357-N und Gly358-N) in dem Bereich beobachtet wurden. Somit ist die Interaktion um das Zink-Ion, welche wahrscheinlich wichtig für eine gute Bindeaffinität ist, sichergestellt. Nur in fünf Simulationen (ZINC31961806, ZINC49571688, ZINC52012042, ZINC52329369 und ZINC86853423) sind keine L1-Loop-Interaktionen vorhanden. Bei den vier erstgenannten Verbindungen könnte dieses Verhalten mit den Ringstrukturen im Liganden begründet sein. In 12 der 18 Simulationen wurden weitestgehend alle Interaktionen zum L1-Loop aus den Docking-Posen wieder gefunden bzw. finden sich weitere Wechselwirkungen zu dem Bereich, die die Bindung zu stabilisieren scheinen.

In den MD-Simulationen zeigte sich, dass ein paar Interaktionen im Bereich um das Zink-Ion und im L1-Loop variabel sind. Dies hängt damit zusammen, dass das Protein und der potenzielle Inhibitor keine starren Moleküle sind. Somit ist eine gewisse Abweichung von der Docking-Pose plausibel. Im Allgemeinen zeigte sich trotzdem, dass die Zink-Koordinierung der Liganden und die globale Position der potenziellen 18 ADAMTS7-Inhibitoren stets vorhanden ist. Experimentelle Untersuchungen (Assays) wären bzw. sind nötig, um die hier erlangten Erkenntnisse mit einer möglichen inhibitorischen Wirkung assoziieren zu können und um zu erklären, welche Wechselwirkungen im Bereich des Zink-Ions bzw. des L1-Loops für eine gute Bindeaffinität wichtig sind.

6.4 Zusammenfassung

In diesem Kapitel wurden die besten 79 Hits aus dem ADAMTS7-Virtual Screening genauer untersucht. Mithilfe eines inversen Screening sollte die Affinität der 79 potenziellen ADAMTS7-Inhibitoren zu den weiteren ADAMTS-Proteinen abgeschätzt werden.

Hier gelang es von den 79 Hits sieben Verbindungen zu identifizieren, welche eine Affinität zu mindestens neun anderen ADAMTS-Proteinen aufweisen. 54 Verbindungen zeigen neben ihrer Affinität zu ADAMTS7 eine gute Bindeaffinität zu weniger als neun anderen ADAMTS-Proteinen. Insgesamt 18 der 79 potenziellen ADAMTS7-Inhibitoren wurden als potenziell selektive ADAMTS7-Inhibitoren eingeordnet, da sie zu keinem weiteren ADAMTS-Protein eine ähnliche oder bessere Affinität als zu ADAMTS7 zeigen.

Die Analyse der 7 nicht selektiven und der 18 selektiven Verbindungen zeigte nicht nur, dass die potenziell selektiven Hits komplexer sind. So wurde postuliert, dass diese auch länger sind und mehr Volumen einnehmen als die nicht selektiven Hits. Mittels eines Klassifikationsverfahrens (SVM) gelang es 24 der 25 Liganden bezüglich ihrer Länge und ihres Volumen als nicht selektive bzw. selektive Hits zu klassifizieren.

Im Folgenden wurden die Docking-Posen der 7 nicht selektiven und der 18 selektiven Verbindungen in ADAMTS7 ausgewertet. Im Vergleich zu den ADAMTS-Docking-Posen des inversen Screenings zeigte sich, dass der Verlust der Zink-Koordinierung zu einem schlechteren Docking-Score und damit zu einer potentiell schlechteren Bindeaffinität führt. Somit kann bestätigt werden, dass die Zink-Koordinierung eine entscheidende Rolle bei der Inhibierung von ADAMTS-Proteinen spielt. In der α -Helix H4 ist das Zink-Bindemotiv lokalisiert. Diese Region ist sequenziell hoch konserviert ist und weist zusätzlich nur geringe strukturelle Abweichungen zwischen den ADAMTS-Proteinen auf (vergleiche Abschnitt 4.2). Zusätzlich zu der proteinseitigen Koordinierung des Zink-Ions finden sich hydrophobe Wechselwirkungen zu der konservierten Aminosäure Phe384 sowie der nicht konservierten Aminosäure Pro381 in den als selektiv klassifizierten Hits. Daher könnte diese Interaktion eine essenzielle Wechselwirkung selektiver Hits in ADAMTS-Proteinen darstellen.

Im Vergleich zum inversen Screening zeigte sich zudem, dass Liganden, denen eine Wasserstoffbrücke in der Nähe des Zink-Ions (ADAMTS7: Gly358-O) im S4-Faltblatt fehlte, auch eine geringere Bindeaffinität zur jeweiligen Bindetasche aufwiesen. Diese Eigenschaft ist, aufgrund der Konservierung von ADAMTS7-Gly358 nicht selektiv, aber wahrscheinlich essenziell für die Affinität. Bei der Analyse der bekannten ADAMTS-Inhibitor-Komplexe konnte bereits gezeigt werden, dass diese Wechselwirkung eine entscheidende Rolle bei der Bindeaffinität spielt (vergleiche Abschnitt 4.3).

In den sequenziellen und strukturellen Vergleichen der ADAMTS-Proteine (vergleiche Abschnitt 4.2) wurde außerdem gezeigt, dass der L1-Loop der ADAMTS-Bindetasche die geringste Konservierung und die größten strukturellen Unterschiede aufweist. Zwischen dem L1-Loop, welcher sich wahrscheinlich auch an Liganden anpassen kann, und den jeweiligen Inhibitoren konnte ein hoher Anteil an Wasserstoffbrücken in den betrachteten Docking-Posen in ADAMTS7 identifiziert werden. Es wurden mehrere Aminosäuren (u. a. Phe413, Ile414, Ser416, Pro417, Gln418, Leu420 und Tyr421) als wichtige Interaktionspartner identifiziert. Diese wurden in den Docking-Posen als Wasserstoffbrückendonor oder -akzeptor beobachtet. Angemerkt werden muss dabei, dass einzelne Wasserstoffbrücken zu Gln418 von Inhibitoren ausgebildet werden, welche als nicht selektiv klassifiziert wurden. Selektive ADAMTS7-Inhibitoren benötigen somit Donor- bzw. Akzeptorgruppen in diesem Bereich, um hier eine stabile Bindung einzugehen. Das wurde bereits bei der Modellierung des ADAMTS7-Pharmacophores (siehe Abschnitt 5.1) durch die Pharmacophore-Features D1, D2 und A1 postuliert.

Mittels MD-Simulationen der ADAMTS7-Docking-Posen der 18 potenziell selektiven Inhibitoren konnte gezeigt werden, dass die Zink-Koordinierung der Liganden und die globale Position der potenziellen 18 ADAMTS7-Inhibitoren vorhanden ist. Zusammengefasst führt wahrscheinlich vor allem die Kombination von Wechselwirkungen in der ADAMTS7-Bindetasche zum S4-Bereich und zum L1-Loop zu einer hohen Bindeaffinität und möglicherweise zu einer selektiven Bindung an ADAMTS7.

Die 79 potenziellen ADAMTS7-Inhibitoren wurden zur experimentellen Testung vorgeschlagen. Die Ergebnisse standen bis zur Fertigstellung der Arbeit noch aus. Weiterhin wird eine *in vitro* Überprüfung bezüglich der Selektivität der 18 potenziell selektiven Hits und der 7 potenziell nicht selektiven Hits benötigt. Die Ergebnisse sind von großem Interesse, da aus ihnen ein Struktur-Aktivitäts-Modell für selektive ADAMTS7-Inhibitoren abgeleitet werden könnte.

7 Zusammenfassung und Ausblick

ADAMTS7 ist ein Mitglied der humanen ADAMTS-Proteinfamilie, welche 19 Mitglieder umfasst. Diese Proteine besitzen eine Metalloproteinase-Domäne, welche eine Bindetasche mit einem katalytischen Zink-Ion enthält [41]. Weiterhin ist bekannt, dass die Inhibition der Funktion in dieser Proteinfamilie über die Bindung von Liganden in diese Bindetasche bei gleichzeitiger Koordinierung des Zink-Ions geschieht [38]. Das proteinkodierende Gen *ADAMTS7* wird mit KHK assoziiert [17, 18, 28]. Tierversuchsstudien zeigten, dass *Adamts7* bei einer KHK überexprimiert wird [31]. ADAMTS7 ist somit ein erfolgversprechendes Target der kardiovaskulären Forschung. Eine Inhibierung dieses Proteins ist ein neuer Ansatz für eine mögliche KHK-Therapie [30]. Bisher sind keine Liganden oder Inhibitoren für dieses Protein bekannt [30]. Daher war das Ziel dieser Arbeit die Suche nach und die Identifizierung von selektiven ADAMTS7-Inhibitoren mittels *in silico*-Analysen.

Aufgrund des Fehlens struktureller Daten für ADAMTS7 wurde eine Modellierung vorgenommen, welche weitere *in silico*-Analysen ermöglichen sollte. Zu diesem Zweck wurde zuallererst die 3D-Struktur des Targetproteins ADAMTS7 mittels des Homology Modelings modelliert (siehe Kapitel 3). Als Template für die Modellierung wurden 2 der 15 bekannten ADAMTS-Röntgenkristallstrukturen ausgewählt (siehe Abschnitt 3.1). Bei den ADAMTS7-Templates handelt es sich um die ADAMTS5-Strukturen mit PDB-ID 3HYG und 2RJQ. Mit ersterer konnte die Metalloproteinase-Domäne modelliert werden, mit zweiterer auch die Disintegrin-ähnliche Domäne des Proteins. Nach Abschluss des Homology Modelings wurden im Dezember 2014 drei weitere ADAMTS4-Strukturen (PDB-ID: 4WK7, 4WKE und 4WKI [177]) veröffentlicht, die eine bessere Auflösung (1,24 – 1,62 Å) aufweisen als die damals bereits bekannten ADAMTS-Röntgenkristallstrukturen. Aufgrund des Zeitpunkts der Veröffentlichung und den bereits begonnenen Analysen wurden diese drei Röntgenkristallstrukturen nicht in die weiteren Analysen einbezogen. Das Homology Modeling wurde zum Vergleich mit drei verschiedenen Programmen (Schrödinger Prime, SWISS-MODEL und YASARA) durchgeführt. Die Modellvalidierung der erstellten vier Modelle zeigte, dass das Homologiemodell, welches mit Schrödinger Prime erzeugt wurde, von allen Modellen die beste Qualität aufweist (siehe Abschnitt 3.3). Daher wurde dieses Modell von ADAMTS7 für die weiteren Analysen genutzt. Somit gelang es im ersten Schritt ein Teilmodell von ADAMTS7 zu erstellen, welches die Metalloproteinase- sowie die Disintegrin-ähnliche Domäne abdeckt. Eine MD-Simulation dieses Homologiemodells bestätigte, dass es sich bei dem 3D-Modell um eine native Struktur von ADAMTS7 handeln kann (siehe Abschnitt 3.3.5).

In der MD-Simulation des ADAMTS7-Homologiemodells (siehe Abschnitt 3.3.5) konnte weiterhin beobachtet werden, dass in der Bindetasche des Proteins zwei Loop-Bereiche deutlich flexibler sind. Publikationen zu den Röntgenkristallstrukturen von ADAMTS4 und ADAMTS5 [47, 164, 165] deuten zudem darauf hin, dass es flexible Bereiche innerhalb der ADAMTS-Bindetasche gibt, welche die Anpassung an Liganden ermöglichen. Bei den flexiblen Regionen handelt es sich um einen Teil des L1-Loops (S1'-Loop) sowie einen Loop, welcher vor der Faltblattstruktur S4 (S2'-Loop) liegt. Die in den Publikationen beschriebenen Regionen sind jene, die in der MD-Simulation des ADAMTS7-Homologiemodells eine im Vergleich erhöhte

Flexibilität aufwiesen. Aus diesem Grund wurde mit einem Sampling der S1'- sowie der S2'-Region der Flexibilität dieser Regionen in der ADAMTS7-Bindetasche Rechnung getragen (siehe Abschnitt 3.4). Daraus resultierten vier weitere Konformationen des S1'-Loops. Für den S2'-Loop konnten für ADAMTS7 keine weiteren Konformationen aufgedeckt werden. Somit konnten keine Veränderung in dieser Region nachgewiesen werden, welche die Flexibilität der Bindetasche beeinflussen. Das ADAMTS7-Conformational Ensemble stellt somit insgesamt fünf Konformationen der ADAMTS7-Bindetasche dar. Dies erlaubte die Modellierung der möglichen Anpassung der Bindetasche an verschiedene (noch unbekannte) Liganden. Mithilfe dieser fünf ADAMTS7-Konformationen sollte in den weiteren Analysen die Flexibilität der Bindetasche berücksichtigt werden.

Der Schwerpunkt dieser Arbeit war die Suche nach ADAMTS7-Inhibitoren, die als zusätzliche Eigenschaft eine Selektivität für das Target aufweisen. Untersuchungen an den ADAMTS-Proteinen (siehe Kapitel 4) sollten Bereiche im ADAMTS7-Protein aufdecken, die bei der Suche nach selektiven Inhibitoren von Bedeutung sein könnten. Hierfür wurden unter anderem Sequenzvergleiche der humanen ADAMTS-Familie vorgenommen (siehe Abschnitt 4.2). Bereits der Vergleich über die gesamten ADAMTS-Sequenzen zeigte eine hohe Konservierung innerhalb dieser Proteingruppe. Mit einem detaillierten Sequenzvergleich der Metalloproteinase-Domänen der ADAMTS-Proteine konnte eine noch stärkere sequenzielle Konservierung in dieser Domäne im Vergleich zur kompletten Sequenz gezeigt werden. Die Sequenzidentität der Metalloproteinase-Domänen im Vergleich zu ADAMTS7 beträgt zwischen 35,7 % (ADAMTS13) und 76,8 % (ADAMTS12). Daher wurde die Bindetasche der ADAMTS-Proteine im Detail untersucht (siehe Abschnitt 4.2.2). Diese wurde hierfür in drei Bereiche unterteilt. Es zeigte sich, dass die Loop-Region L1 in der Bindetasche sequenziell weniger stark konserviert (durchschnittlich 25,2 %) ist als die anderen beiden Regionen um die α -Helix H4 (durchschnittlich 63,3 %) und das β -Faltblatt S4 (durchschnittlich 46,5 %). Aufgrund der starken Konservierung der Metalloproteinase-Domänen sowie der Bindetaschen der ADAMTS-Proteine sollte ein Strukturvergleich Aufschluss über Unterschiede in den ADAMTS-Proteinen geben. Zu diesem Zweck wurden Homologiemodelle der Metalloproteinase-Domänen von allen ADAMTS-Proteinen angefertigt, für die es keine experimentellen Strukturen gab (siehe Abschnitt 4.1). Röntgenkristallstrukturen sowie die Homologiemodelle ermöglichten den Strukturvergleich der Metalloproteinase-Domänen (siehe Abschnitt 4.2.1) sowie im Speziellen der Bindetasche der ADAMTS-Proteine (siehe Abschnitt 4.2.2). Anhand des Strukturvergleichs der Metalloproteinase-Domänen im Gesamten konnte eine durchschnittliche Abweichung ($\text{RMSD}_{\text{C}\alpha}$) von 8,3 Å identifiziert werden (siehe Abschnitt 4.2.1). Hiermit wurde nachgewiesen, dass diese Domänen in den ADAMTS-Proteinen trotz hoher sequenzieller Konservierung Unterschiede in der Struktur aufweisen. Eine detaillierte Analyse der drei Bereiche in der Bindetasche, analog zum Sequenzvergleich, zeigte, dass sich der Bereich um die α -Helix H4 wenig unterscheidet (durchschnittlicher $\text{RMSD}_{\text{C}\alpha}$ von 1,3 Å). In den Regionen um S4 und L1 konnten stärkere strukturelle Abweichungen beobachtet werden. Vor allem für die L1-Loop-Region konnte gezeigt werden, dass alle ADAMTS-Proteine untereinander eine konstante Abweichung (durchschnittlicher $\text{RMSD}_{\text{C}\alpha}$ von 6,5 Å) aufweisen. Die strukturellen Unterschiede in der Bindetasche geben somit einen Hinweis auf Interaktionsbereiche (L1-Loop) für selektive ADAMTS7-Inhibitoren.

Im Anschluss wurden bekannte ADAMTS-Inhibitor-Komplexe untersucht, die Aufschluss über Gemeinsamkeiten und Unterschiede in den Bindemustern geben sollten. Ziel war herauszufinden, ob es essenzielle Wechselwirkungen zwischen Bindetasche und Inhibitoren in dieser Familie gibt. Die Analyse der Komplexe und ihrer Bindemuster (siehe Abschnitt 4.3) verifizierte die Koordinierung des Zink-Ions in der Bindetasche

über den jeweiligen Inhibitor in allen untersuchten Komplexen. Des Weiteren treten Wasserstoffbrückenbindungen (ADAMTS7: Leu357, Gly358) in der Umgebung des Zink-Ions auf, welche die Bindung der Liganden in dieser Position zu stabilisieren scheinen. Diese Wechselwirkungen könnten unabhängig von der Selektivität essenziell für eine Bindung an ADAMTS-Proteine sein. In diesem Fall ist nur die Aminosäure Gly358 über alle ADAMTS-Proteine konserviert. Außerdem wurden Wasserstoffbrücken im flexiblen L1-Loop beobachtet, was die oben aufgestellte These unterstützt, dass dieser Bereich ein wichtiger Interaktionsbereich für selektive ADAMTS7-Inhibitoren sein könnte. Sequenz- und Strukturvergleiche sowie die beobachteten Bindemuster lassen vermuten, dass der L1-Loop eine wichtige Rolle bei der selektiven Interaktion mit Inhibitoren spielen könnte.

Basierend auf den vorigen Ergebnissen sollten anhand eines ADAMTS7-Virtual Screenings mögliche Inhibitoren für das Target ADAMTS7 identifiziert werden. Vor dem Virtual Screening für ADAMTS7 (siehe Kapitel 5) wurde die Bindetasche des Proteins, auch basierend auf den bereits erlangten Ergebnissen, charakterisiert werden. Mithilfe eines strukturbasierten Pharmacophores konnten Eigenschaften von Liganden definiert werden, welche eine Bindung an das Targetprotein ermöglichen sollten (Abschnitt 5.1). Hierbei wurden mittels der *e-Pharmacophore*-Methode vier Features (drei Wasserstoffbrückendonor-Features D1, D2 und D3 sowie ein Wasserstoffbrückenakzeptor-Feature A1) sowie die Grenzen der Bindetasche definiert. Eine Evaluierung der Pharmacophore-Features (siehe Abschnitt 5.1.1) zeigte, dass die assoziierten Aminosäuren dreier Features (D1: Ser416, D2: Asp422, A1: Leu420) sequenziell nicht konserviert sind. Zudem wurden die Bindetaschen der ADAMTS-Familie auf das Vorhandensein der Features an den gleichen Positionen untersucht. Es konnte gezeigt werden, dass die Features A1 und D2 nur noch in jeweils einer anderen ADAMTS-Bindetasche in der gleichen Region zu finden sind. Außerdem liegen sie im ADAMTS7-Protein in der strukturell diversen L1-Loop-Region. Sie könnten somit wichtig für die selektive Bindung von ADAMTS7-Inhibitoren sein. Ein Manko der *e-Pharmacophore*-Methode ist, dass die Koordinierung der Zink-Ionen nicht erkannt wird. Daher wurde eine manuelle Anpassung des ADAMTS7-Pharmacophores vorgenommen (siehe Abschnitt 5.1.2). Hierfür wurden zwei weitere Features (A2 und A3) hinzugefügt. Außerdem fand eine Anpassung in Anlehnung an die Auswertung der ADAMTS-Bindemuster von bekannten Protein-Inhibitor-Komplexen statt. Dazu wurde ein Wasserstoffbrückendonor-Feature bei der Aminosäure Gly358 (D4) hinzugefügt. Dies diente dazu, sicherzustellen, dass die ADAMTS-Bindemuster Interaktionen in der Nähe des Zink-Ions ausbilden. Diese Wechselwirkung könnte potenziell zur Stabilisierung der Bindung in ADAMTS-Proteinen wichtig sein, auch wenn sie für keine selektiven Eigenschaften steht. Zusammengefasst gelang es, ein ADAMTS7-Pharmacophore zu definieren, welches aus sieben Features besteht, die mögliche selektive Eigenschaften von Inhibitoren sowie die Zink-Koordinierung beinhalten. Das Pharmacophore stellte einen Filterschritt im anschließenden Virtual Screening nach selektiven ADAMTS7-Inhibitoren dar. Die Ergebnisse der Modellierung einer räumlichen ADAMTS7-Struktur sowie der Pharmacophoremmodellierung sind bereits in [173] publiziert.

Zur Evaluierung, welches Dockingprogramm die Bindung an ADAMTS(7) am besten abschätzen kann und daher bei der Suche nach ADAMTS7-Inhibitoren genutzt werden sollte, wurde ein Redocking mit den zehn bekannten ADAMTS-Inhibitor-Komplexen durchgeführt (siehe Abschnitt 5.2). Es standen die vier Dockingprogramme AutoDock4, AutoDockVina, PLANTS und Schrödinger Glide zur Verfügung. Es konnte gezeigt werden, dass das Programm Schrödinger Glide in der jeweils besten Docking-Pose mit einer durchschnittlichen Abweichung von 1,3 Å von der nativen Ligandenkonformation die besten Ergebnisse

über alle getesteten ADAMTS-Strukturen erreichte. Aus diesem Grund wurde Schrödinger Glide als Programm für alle weiteren Docking-Analysen genutzt.

Das anschließende Virtual Screening nach potenziellen ADAMTS7-Inhibitoren umfasste mehrere Filterschritte (Abschnitt 5.3). Die ZINC-Datenbank (*Drug-Like*) [153] wurde als Datenbank und damit als Grundlage für mögliche Inhibitoren ausgewählt. Rund 23 Millionen Verbindungen wurden mit dem ADAMTS7-Pharmacophore im ersten Filterschritt gescreent. Die verbleibenden Moleküle wurden in das ADAMTS7-Conformational Ensemble gedockt (2. Filterschritt, siehe Abschnitt 5.3.3). Die resultierenden Docking-Posen wurden anhand mehrerer Kriterien bewertet (siehe Abschnitt 5.3.4): Zink-Koordinierung (3. Filterschritt), Wasserstoffbrücke zu Gly358-O (4. Filterschritt) sowie Docking-Score (5. Filterschritt). Dies diente der Bewertung der Bindung bzw. der Abschätzung der Bindeaffinität der Verbindungen zu ADAMTS7. Anhand der ersten vier Schritte des Virtual Screenings gelang es, eine große Anzahl an Liganden zu identifizieren, welche potenziell in die Bindetasche passen und binden könnten. Im fünften Filterschritt wurde eine Anreicherung von potenziellen ADAMTS7-Inhibitoren mit hoher Bindeaffinität zu ADAMTS7 in den besten 5 % der Verbindungen (bzgl. des Docking-Scores) prognostiziert. Dies entspricht 79 Verbindungen mit den niedrigsten Docking-Scores. Diese 79 Verbindungen wurden weiteren Analysen unterzogen, um abzuschätzen, wieso diese als potenzielle beste Binder auftreten und ob diese selektiv für ADAMTS7 sein könnten.

Weiterführende Analysen mit Proteinen aus der ADAMTS-Familie lassen mögliche Rückschlüsse über die Selektivität der gefundenen Hits zu. Mithilfe eines inversen Screenings (siehe Abschnitt 6.2) sollte die Bindeaffinität der 79 Verbindungen zu den 18 weiteren ADAMTS-Proteinen abgeschätzt werden. Dies sollte die Charakterisierung von nicht selektiven und selektiven Hits ermöglichen. Hierbei gelang es, 18 Verbindungen zu identifizieren, die im inversen Screening in den anderen 18 ADAMTS-Strukturen mindestens eins der folgenden drei Kriterien erfüllen: keine Zink-Koordinierung, keine Wasserstoffbrücke zum Gly358-O-Analogon von ADAMTS7 oder Docking-Score höher als der Docking-Score in ADAMTS7. Es wird vermutet, dass dies auf möglicherweise selektive Verbindungen hindeutet. Ebenso konnten sieben wahrscheinlich nicht selektive Hits innerhalb dieser Analysen identifiziert werden, welche in mindestens neun weiteren ADAMTS-Proteinen die oben genannten Kriterien nicht erfüllten (siehe Abschnitt 6.2.1).

Die Analyse der Wechselwirkungen zur ADAMTS7-Bindetasche der 79 Verbindungen (siehe Abschnitt 6.1) sowie der potenziell selektiven Inhibitoren (siehe Abschnitt 6.3.2) zeigte kein eindeutiges Bindemuster auf dem eine Affinität bzw. Selektivität zu ADAMTS7 basieren könnte. Es wurde beobachtet, dass bei potenziell nicht selektiven Verbindungen (siehe Abschnitt 6.3.2), neben der Zink-Koordinierung und der Interaktion zu Aminosäuren nahe des Zink-Ions, häufig eine Wasserstoffbrücke zu Gln418 vorliegt. Außerdem sind die funktionellen Gruppen der nicht selektiven Hits vor allem OH-Gruppen. Die Liganden sind im Vergleich zu den potenziell selektiven Molekülen kleiner und nehmen in der Docking-Pose ein kleineres Volumen ein. Daher wird postuliert, dass Verbindungen mit kurzen, linearen Kohlenstoffketten und mehreren –OH-Gruppen sich an viele ADAMTS-Proteine anpassen und somit keine selektive Bindung eingehen.

Für die 18 potenziell selektiven Verbindungen konnte beobachtet werden, dass diese vergleichsweise groß sind und ein größeres Volumen im Gegensatz zu den als nicht selektiv klassifizierten Hits einnehmen. Mittels eines Klassifikationsverfahrens (SVM) konnte ein Zusammenhang zwischen Größe und Volumen der potenziellen Inhibitoren und der Selektivität zu ADAMTS7 bestätigt werden (siehe Abschnitt 6.3.1). Daher wurde postuliert, dass die 18 potenziell selektiven ADAMTS7-Inhibitoren aufgrund ihrer Größe nicht in die Bindetaschen der anderen ADAMTS-Proteine passen. Außerdem sind diese Hits strukturell vielfältiger.

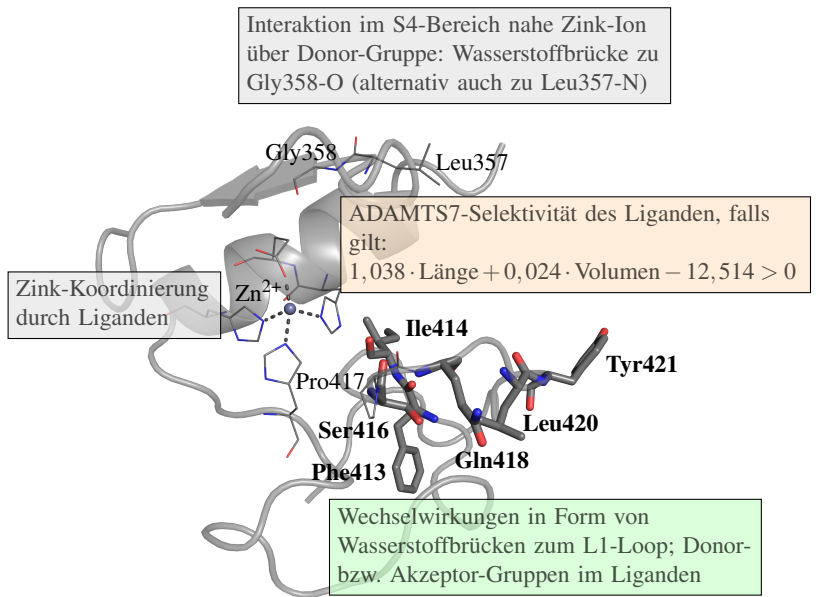
Bei der Auswertung der Posen in ADAMTS7 zeigte sich für diese 18 Hits, dass in fast allen Docking-Posen Wasserstoffbrücken zum L1-Loop vorliegen (siehe Abschnitte 6.2.2 und 6.3.3). MD-Simulationen der ADAMTS7-Docking-Posen der 18 potenziell selektiven Inhibitoren zeigten, dass die Zink-Koordinierung der Liganden und die globale Position der potenziellen 18 ADAMTS7-Inhibitoren stabil ist. Hier wurde beobachtet, dass zusätzlich zur Gly358-O-Wasserstoffbrücke auch Wechselwirkungen dieser Art zu Leu357-N und Glu355-O ausgebildet werden.

Schlussendlich wurde postuliert, dass ADAMTS7-Inhibitoren an das Zink-Ion koordinieren, eine Wasserstoffbrücke mit der Aminosäure Gly358, welche räumlich nah am Zink-Ion liegt, oder deren Nachbar-Aminosäuren ausbilden sollten sowie mit dem L1-Loop über Wasserstoffbrücken wechselwirken, um eine hohe Bindeaffinität zum Targetprotein aufzuweisen (siehe Abschnitt 6.3.1). Die Selektivität für das Protein ADAMTS7 wird dabei wahrscheinlich durch das Wechselwirken mit dem sequenziell weniger stark konservierten und strukturell differenziellen L1-Loop (außer Gln418) erreicht. Die Ergebnisse zeigen, dass wahrscheinlich wichtige Eigenschaften potenzieller Inhibitoren bereits im ADAMTS7-Pharmacophore definiert wurden und dieses somit eine gute Basis für die Identifizierung von ADAMTS7-Inhibitoren darstellt(e). In Abbildung 7.1 wurden die Eigenschaften potenzieller (selektiver) ADAMTS7-Inhibitoren zusammenfassend dargestellt. Zum Vergleich ist weiterhin das ADAMTS7-Pharmacophore abgebildet. Final konnten Erkenntnisse zum ADAMTS7-Protein, deren Bindetasche und möglicher (selektiver) Inhibitoren erlangt werden. Experimentelle Untersuchungen, z. B. Mutations-Analysen, sind nötig, um die Wichtigkeit der einzelnen Wechselwirkungen evaluieren zu können. Die *in silico*-Analysen geben Hinweise, die durch weitere Tests bestätigt werden müssen.

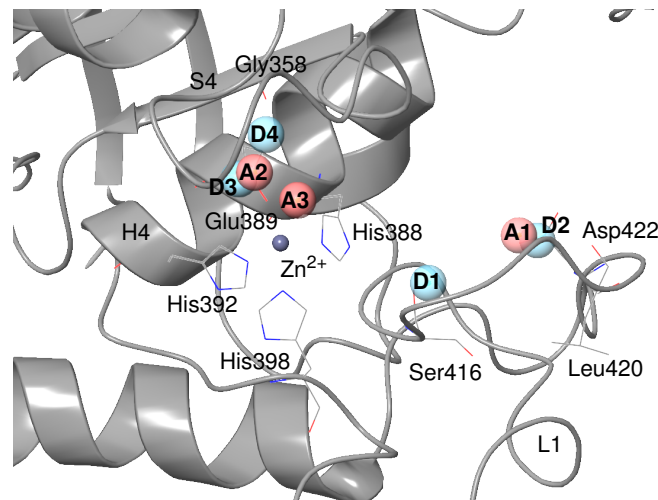
ADAMTS-Proteine sind, genau wie ADAM- und MMP-Proteine, aufgrund ihrer Assoziation mit zahlreichen Erkrankungen [36] ein attraktives Target. Zukünftige Analysen zur Selektivität von Inhibitoren sollten daher auch die ADAM-Proteine und MMPs mit einbeziehen. Diese Familien gehören zu den Metalloproteasen und aufgrund der Ähnlichkeiten dieser Proteine, die alle eine Metalloproteinase-Domäne mit katalytischem Zink-Ion beinhalten, sind Kreuzreaktionen mit diesen Proteinen zu erwarten, die unter anderem durch weitere inverse Screenings abgeschätzt werden könnten. Mit den Informationen über mögliche Kreuzreaktionen wäre es möglich, bereits im Vorfeld mögliche Nebenwirkungen der Verbindungen auszuschließen.

In dieser Arbeit wurde nach Liganden bzw. Inhibitoren, welche in die bekannte Bindetasche von ADAMTS7 binden können, gesucht. Ein weiterer Ansatz wäre die Suche nach allosterischen Inhibitoren von ADAMTS7. Für die Frage nach potenziellen allosterischen Zentren von ADAMTS7 und allosterischen Inhibitoren ist der modellierbare Bereich des Proteins jedoch nicht ausreichend. Auch wenn dieser weitere Aspekt möglicherweise Vorteile bei der Suche nach selektiven Inhibitoren gebracht hätte, konnte ihm in dieser Arbeit und mit den momentan zur Verfügung stehenden Daten nicht auf den Grund gegangen werden. Das Gleiche gilt für die Entwicklung von Antikörpern für ADAMTS7. Kann zukünftig die (komplette) ADAMTS7-Struktur bzw. ein größerer Bereich des Proteins bestimmt oder modelliert werden, empfehlen sich diese Analysen. Sie können ein weiterer Aspekt bei der Suche nach selektiven Inhibitoren für dieses Target sein.

Basierend auf den vorangegangenen Analysen rund um das Protein ADAMTS7 sollen die 79 ausgewählten Verbindungen des Virtual Screenings zur *in vitro*-Testung vorgeschlagen werden, wobei ein Augenmerk auf die erwähnten 18 potenziell selektiven Verbindungen gelegt werden sollte. Es konnte nicht eindeutig geklärt werden, ob die Interaktion über eine Wasserstoffbrücke zu Gly358-O ein essenzielles Merkmal von ADAMTS-Inhibitoren darstellt. Daher wäre eine Möglichkeit im weiteren Verlauf bei *in vitro*-



(a) Eigenschaften potenzieller ADAMTS7-Inhibitoren



(b) ADAMTS7-Pharmacophore

Abbildung 7.1: (a) Wechselwirkungen und Eigenschaften potenzieller ADAMTS7-Inhibitoren in der Bindetasche im Vergleich zum ADAMTS7-Pharmacophore (b). Grün markiert in (a) ist das wahrscheinlich selektive Merkmal: Interaktionen mit dem L1-Loop. Interagierende Aminosäuren in diesem Bereich aus den untersuchten Bindemustern sind als Linien (konserviert) bzw. Stäbchen (nicht konserviert) dargestellt. Das ADAMTS7-Modell (grau) ist in der Cartoon-Darstellung, die Zink-koordinierenden Histidine als Linien veranschaulicht. Das Zink-Ion ist als graue Kugel zu erkennen.

Untersuchungen auch die besten Liganden, die das Zink-Ion koordinieren, aber keine Wasserstoffbrücke zu Gly358-O jedoch z. B. zu Leu357-N aufweisen, zu testen. Hierzu könnte der Filter um eine Wasserstoffbrücke mit der Aminosäure Leu357 erweitert werden.

Bis zum Abschluss der Arbeit konnten die *in vitro*-Versuche noch nicht durchgeführt werden. Daher bedürfen alle hier gemachten Aussagen einer Validierung durch Labor-Untersuchungen. Mit Ergebnissen aus *in vitro*-Analysen können weiterführende Aussagen bezüglich der potenziellen ADAMTS7-Inhibitoren und möglichen Variationen vorgenommen werden. In anknüpfenden Untersuchungen könnten, z. B. mittels einer QSAR [178] (Abk. für *Quantitative Structure-Activity Relationship*), die validierten wichtigen Eigenschaften von ADAMTS7-Inhibitoren herauskristallisiert werden. Hierdurch könnten Merkmale für eine verbesserte nächste Generation von potenziellen ADAMTS7-Inhibitoren abgeleitet werden. Nach diesen Schritten wäre man dem Ziel, durch die Inhibition von ADAMTS7 (das Fortschreiten einer) KHK zu verhindern, ein großes Stück näher. Schlussendlich ist die Hoffnung, dass eine Inhibition von ADAMTS7 in der Zukunft Einzug in die Therapie bei KHK hält.

A Anhang

A.1 Ergebnisse: Modellierung der ADAMTS7-Proteinstruktur

A.1.1 Wahl des/der Template(s)

	215	220	225	230	235	240	245	250	255	260	265	270	275	280	285	290
ADAMTS7	RRERWEQRQQWRRPRLRLRHQR	SVSK	EWVETLVV	ADAKMVEYHGQP	QVESYVLTIMNMVAGLFHDPS	IGNPIHIT										
2JIH
2V4B
3Q2G
3Q2H
2RJ2
3B2Z
2RJQ
3B8Z
3HY7
3HY9
3HYG
3LJT
3GHM
3GHN
3VN4
	295	300	305	310	315	320	325	330	335	340	345	350	355			
ADAMTS7	IVRLVLL.....	EDDEEDLKI	THHADNTLKS	FCWKQKSI	NMKGDÄH	PLHHDTA	ILLTRKDL	CAAMNRP	CETLG							
2JIH
2V4B
3Q2G
3Q2H
2RJ2
3B2Z
2RJQ
3B8Z
3HY7
3HY9
3HYG
3LJT
3GHM
3GHN
3VN4
	360	365	370	375	380	385	390	395	400	405	410	415	420	425	430	
ADAMTS7	LSHVAGMCPHR	SCSINEDT	GLPLAFTVA	HELGHSGFI	QHDSGND	CEPVG...	KRPFIM	SQOLLY	DAAPLT	WSRC						
2JIH
2V4B
3Q2G
3Q2H
2RJ2
3B2Z
2RJQ
3B8Z
3HY7
3HY9
3HYG
3LJT
3GHM
3GHN
3VN4
	435	440	445	450	455	460	465	470	475	480	485	490	495	500		
ADAMTS7	SROYITRFLDR	GWGLCLDD	PPAKDI	IDFPSVPP	GLVDVSHO	CRLOYG	AYSACF	EDMNDV	CVHTLW	CSV...	GTTC					
2JIH
2V4B
3Q2G
3Q2H
2RJ2
3B2Z
2RJQ
3B8Z
3HY7
3HY9
3HYG
3LJT
3GHM
3GHN
3VN4

Abbildung A.1: Multiples Alignment von der ADAMTS7-Sequenz und den zum Zeitpunkt des Homology Modelings bekannten ADAMTS-Röntgenkristallstrukturen. Die ADAMTS7-Sequenz wurde aus Übersichtlichkeitsgründen auf den Bereich des Alignments eingekürzt.

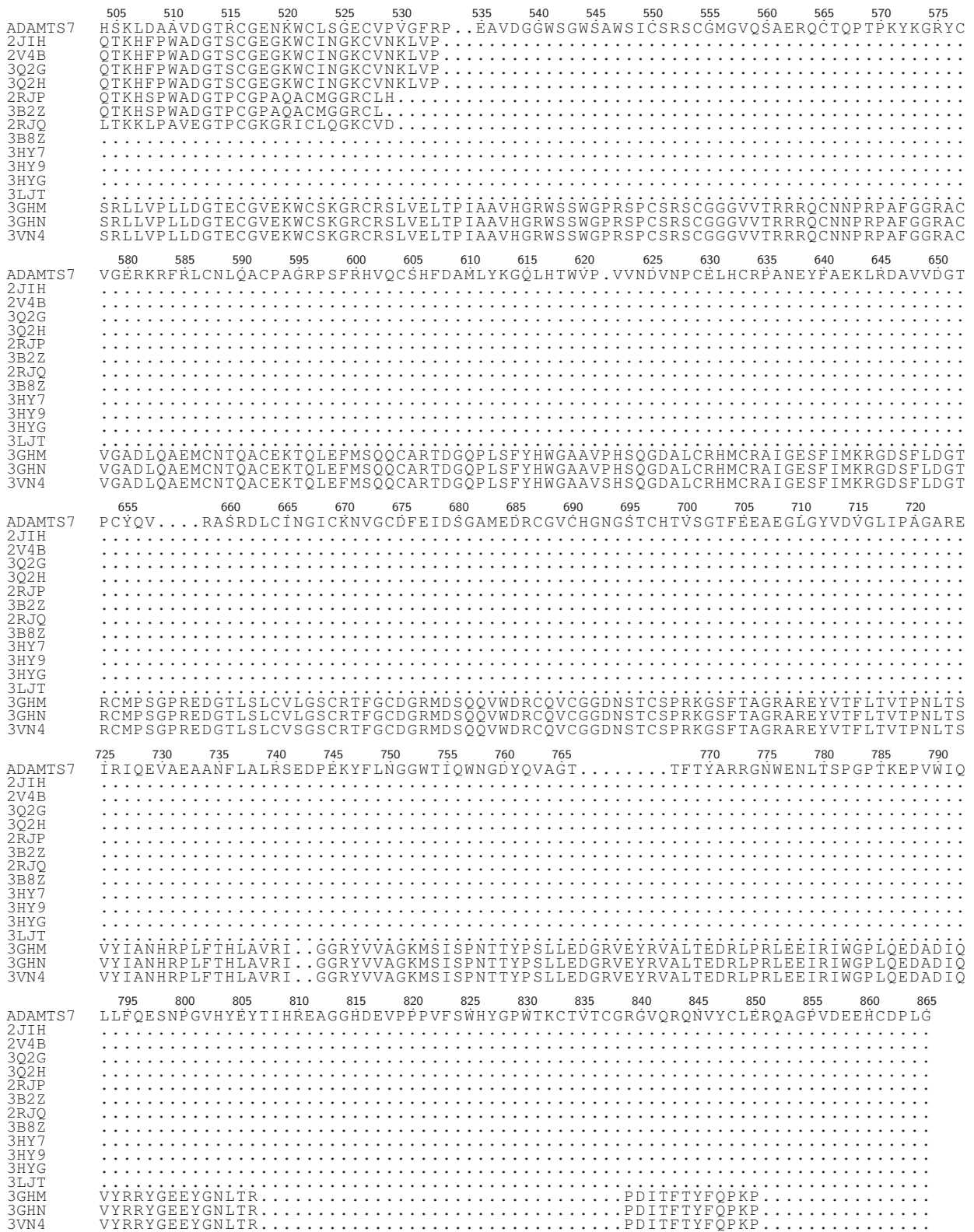


Abbildung A.1 (Fortsetzung): Multiples Alignment von der ADAMTS7-Sequenz und den zum Zeitpunkt des Homology Modelings bekannten ADAMTS-Röntgenkristallstrukturen. Die ADAMTS7-Sequenz wurde aus Übersichtlichkeitsgründen auf den Bereich des Alignments eingekürzt.

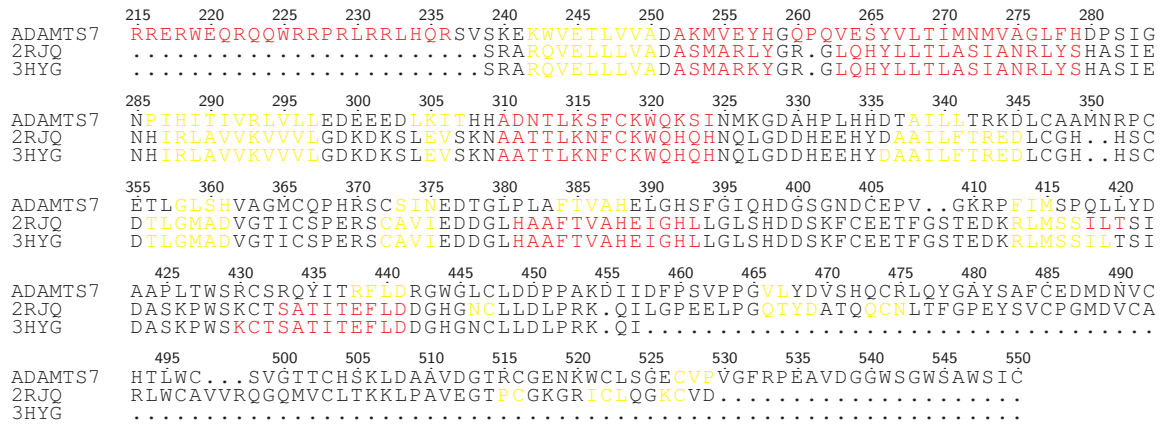


Abbildung A.2: Multiples Alignment von der ADAMTS7-Sequenz und den ADAMTS5-Template-Strukturen (PDB-ID: 3HYG, 2RJQ). Die ADAMTS7-Sequenz wurde aus Gründen der Übersichtlichkeit auf den Bereich des Alignments eingekürzt. Die Sekundärstrukturen bzw. die vorhergesagten Sekundärstrukturen für ADAMTS7 sind folgendermaßen eingefärbt: α -Helix - rot und β -Faltblatt - gelb.

A.1.2 Modellvalidierung der ADAMTS7-Modelle

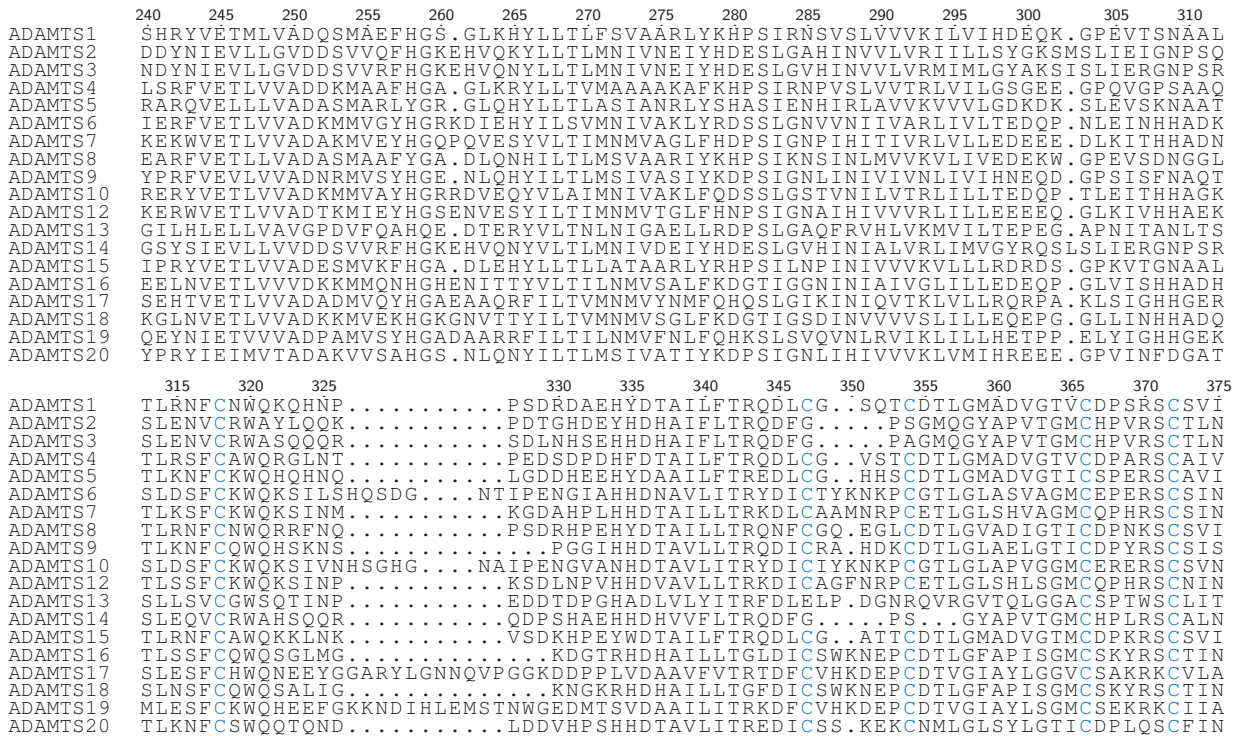


Abbildung A.3: Multiples Sequenzalignment der Metalloproteinase- und Disintegrin-ähnlichen Domänen der ADAMTS-Proteine. Die Sequenzen dieser ADAMTS-Domänen stammen aus der UniProt-Datenbank [169]. Die Nummerierung des Alignments bezieht sich auf ADAMTS7. Cysteine sind blau markiert.

	380	385	390	395	400	405	410	415	420	425	430	435	440	445
ADAMTS1	EDDGLQAAFTTAAHELGHVFNMPHDDAKQCASLNG.VN	QDSHMMASMLSNLDHSQPWSF	CSAYMITSFLDNGHGE											
ADAMTS2	HEDGFSSAFVVAHETGHVLMGMEHDGQGNRCGDEV...	RLGSIAPLVQAAFHFRHWSR	CSQOELSRYLHSYDCL											
ADAMTS3	HEDGFSSAFVVAHETGHVLMGMEHDGQGNRCGDEV...	AMGSVMAPLVQAAFHRYHWSR	CSGOELKRYIHSYDCL											
ADAMTS4	EDDGLQSAFTAAHELGHVFNMLHDNSKPCISLNGPLST	SRHVMAVMAHVDPPEEPWSP	CSARFITDFLDNGYGH											
ADAMTS5	EDDGLHAFTVAHEIGHLLGLSHDSSKFCBETFG.STED	KRLMSSILTSIDASKPWSK	CTSATITEFLDDGHGN											
ADAMTS6	EDIGLGSAFTIAHEIGHNFGMNHHDGIGNSCGTKG...	HEAAKLMAAHITANTNPFWSA	CSRDIYITSLDSSGRGT											
ADAMTS7	EDTGLPLAFTVAHELGHSFGIOHDGSGNDCEPVG...	KRPFIMSPQLLYDAAPLWWSR	CSRQYITRFLDRGWGL											
ADAMTS8	EDEGLQAAHTLAHELGHVLSMPHDDSKPCTRLFG.PMG	KHHVMAPLFVHLNQTLPWSF	CSAMYLTEFLDGGHGD											
ADAMTS9	EDSGLSTAFTIAHELGHVFNMPHDDNNKCKEEGV...	KSPQHVMAPTLNFYTNPNWWSK	CSRKYITEFLDTGYGE											
ADAMTS10	EDIGLATAFTIAHEIGHTFGMNHHDGIVGNSCGARG...	QDPAKLMAAHITMKTNPFWWSK	CSRDIYITSLDSSGLGL											
ADAMTS12	EDSGLPLAFTIAHELGHSFGIOHDGKENDCEPVG...	RHPYIMSRQLQYDPTPLWWSK	CSSEYITRFLDRGWGF											
ADAMTS13	EDTGFDLGVTIAHEIGHSFGLSHDGGAPGSGCG...	PSGHVMSADGAAPRAGLAWSP	CSRQELLSLLSAGRAR											
ADAMTS14	HEDGFSSAFVVAHETGHVLMGMEHDGQGNRCGDEV...	SLGSVMAPLVQAAFHFRHWSR	CSKLELSRYLPSYDCL											
ADAMTS15	EDTGLPLAFTVAHELGHVFNMPHHDNVKVEEVFG...	KLRANHMMSPTLIQIDRANPWSA	CSAAIITDFLDSGHGD											
ADAMTS16	EDTGLGLAFTIAHESGHNFGMIHHDGEGNMCKK...	SEGNIMSPTLAGRNGVFWWSF	CSRQYLHKFLSTAQAI											
ADAMTS17	EDNGLNLAFTIAHELGHNLMNHDDHSSCAGR...	SHIMSGEWVKGRNPSDLWSK	CSRQDLLENFLKSKVST											
ADAMTS18	EDTGLGLAFTIAHESGHNFGMIHHDGEGNPCRK...	AEGNIMSPTLTGNGVFWWSK	CSRQYLLKFLSTPQAS											
ADAMTS19	EDNGLNLAFTIAHEGMHNMGINHDDHNSPCADG...	LHIMSGEWIKQNLGDVSWWSR	CSRQEDLERFLRSKASN											
ADAMTS20	EEKGLISAFTIAHELGHTLGVQHDNDNPRCKEMKV...	TKYHVMAFALSFMSPWWSN	CSRKYVTEFLDTGYGE											
	450	455	460	465	470	475	480	485	490	495	500	505		
ADAMTS1	CLMDKPKQ...NP	IQLPGD..LPG	TSYDANRQCQ	FTFGEDSKHCP	DAAS...T	CSTLWCTGT..	SGGVLV	CQTK						
ADAMTS2	LDDPFA...HD	WPALPQ..LP	GLHYSMNEQ	CRFD	FGLGYMCT	AFRT...F	DPCKQLWCS	HDPNPF	YF	CKTK				
ADAMTS3	LDDPFD...HD	WPKLPE..LP	GINYSMDEQ	CRFD	FVGVMCT	AFRT...F	DPCKQLWCS	HDPNPF	YF	CKTK				
ADAMTS4	CLLDKPE...A	PLHLPVT..F	PKDYADRQC	QLTF	GPDSRHCP	QLPP...P	CAALWCSG	H..L	NGHAM	CQTK				
ADAMTS5	CLLDLPR...K	QILGPEE..L	PGQTYDATQ	CNLT	TFGPEYSV	CPGMDV...C	ARLWCAV	V..R	QGMV	CLTK				
ADAMTS6	CLDNEPP...K	RDFLYPAV..A	PGVYDADEQ	CRFY	QATSRO	CKYGEV...C	RELWCLS	SKS.N	RCV	TNSIPA				
ADAMTS7	CLDDPPA...K	DIIDFVSV..P	GVLYDVSHQ	CRLO	YGAYSAF	CEMDN...V	CHTLWCS	...V	GT	CHSK				
ADAMTS8	CLLDAPA...A	ALPLPTGL	PRMALYQLDQ	CRIF	FGDFRHC	PNNTSAQ.DV	CAQLWCH	TDG.A	EPL	CHTKNG				
ADAMTS9	CLLNEPE...S	RYPLPVO..L	PGLYVNVNK	CELIF	FGPGSQV	CPYMMQ...C	RLWCSNN	V..N	GVH	KGCRTO				
ADAMTS10	CLNRRP...R	QDFVYPTV..A	PGQAYDADEQ	CRFO	HGVKSRQ	CKYGEV...C	SELWCLS	SKS.N	RC	TNSIPA				
ADAMTS12	CLDDIPK...K	KGLKSKVI..A	PGVIYDVHHQ	CQLQ	YGPNATF	CQEVEN...V	CQTLWCS	...V	K	GF	CRSL			
ADAMTS13	CWDDPPRP	QPGSAGHPDA..Q	PGLYYSANEQ	CRVA	FGPKAVAC	TFAREHLD	MCQALS	CHTD..P	L	DQSS	CSRL			
ADAMTS14	LDDPFD...P	AWPQPEE..L	GINYSMNEQ	CRFD	FGSGYQT	CLAFRT...F	DPCKQLWCS	HDPNPF	YF	CKTK				
ADAMTS15	CLLDQPS...K	PISLPED..L	PASYSLSQ	CELAF	FGVSKP	CPYMQY...C	TKLWCTG	K..A	K	GMV	CQTR			
ADAMTS16	CLADQPK...P	VKEYKYPEK..L	PELYDANTQ	CKWQ	FGAKAKL	CMLDFK...K	DIKALWC	..H	R	GRK	CETK			
ADAMTS17	CLLVTDK...R	SQHTVRLPHK..L	PGMYSANEQ	CQIL	FGMNATF	CRNMEH...L	MCAGLW	LC..V	E	GT	CKTK			
ADAMTS18	CLVDEPK...Q	AGQYKYPDK..L	PGIYDADTQ	CKWQ	FGAKAKL	CSLGFV...K	DIKALWC	..H	R	GRK	CETK			
ADAMTS19	CLLQTNP...Q	SVNSVMVPSK..L	PGMTYADEQ	CQIL	FGLASF	CQEMQH...V	ITGLWCK	..V	E	GE	KE	CRTK		
ADAMTS20	CLLDKPD...E	EIYNLPSE..L	PSRYDGNKQ	CELAF	FGGSGM	CPHINI...C	MHLWCT	ST..E	K	LH	KG	CFTQ		
	510	515	520	525										
ADAMTS1	HFPWADGTS	CGEGKWC	INGKCVN											
ADAMTS2	KGPPLDGTM	CAPGKHC	FKGHC	CIW										
ADAMTS3	KGPPLDGTE	CAAGKWC	YKGC	CMW										
ADAMTS4	HSPWADGTP	CGPAQAC	MGGRC	CLH										
ADAMTS5	KLPAVEGTP	CGKGRIC	LQK	CVD										
ADAMTS6	AEGTLCQT	GNIEKGC	YQGD	CVP										
ADAMTS7	LDAAVDGT	RCGENKWC	CLSGE	CVP										
ADAMTS8	SLPWADGTP	CGPGHLC	SEGS	CLP										
ADAMTS9	HTPWADGTE	CEPGKHC	YGF	CVP										
ADAMTS10	AEGTLCQT	HTIDKGC	YKRV	CVP										
ADAMTS12	LDAAADGT	QCGEKKWC	MAGK	CIT										
ADAMTS13	LVPLLDGTE	CGVEKWC	SKGR	CRS										
ADAMTS14	KGPPLDGTE	CAPGKWC	FKGHC	CIW										
ADAMTS15	HFPWADGTS	CGEGKLC	LKG	CVE										
ADAMTS16	FMPAAEGTI	CGHDMWC	RRGQ	CVK										
ADAMTS17	LDPPLDGT	ECGADKWC	RAGE	CVS										
ADAMTS18	FMPAAEGTV	CGLSMWC	RQGG	CVK										
ADAMTS19	LDPMDGTD	CDLGGKWC	RAGE	CTS										
ADAMTS20	HVPADGTD	CGPGMHC	RHGL	CVN										

Abbildung A.3 (Fortsetzung): Multiples Sequenzalignment der Metalloproteinase- und Disintegrin-ähnlichen Domänen der ADAMTS-Proteine. Die Sequenzen dieser ADAMTS-Domänen stammen aus der UniProt-Datenbank [169]. Die Nummerierung des Alignments bezieht sich auf ADAMTS7. Cysteine sind blau markiert.

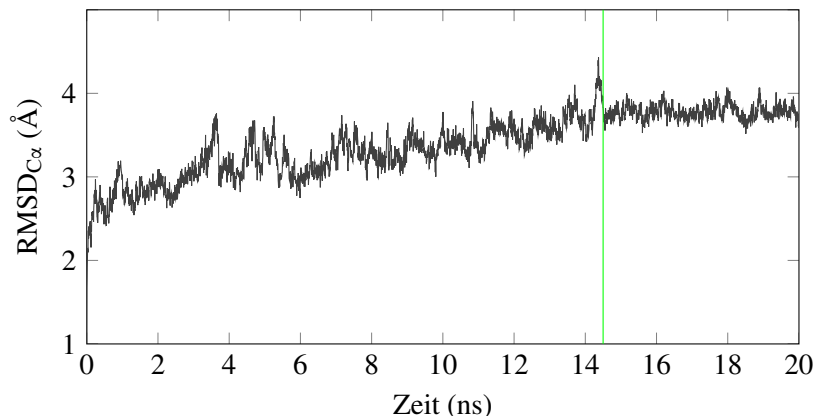


Abbildung A.4: $RMSD_{C\alpha}$ über die 20 ns lange MD-Simulation des ADAMTS7-Homologiemodells im Vergleich zur Startstruktur. Der Beginn des stabilen Zeitraums bei 14,5 ns ist durch eine grüne, vertikale Linie markiert.

A.2 Ergebnisse: Untersuchungen der ADAMTS-Familie

A.2.1 Sequenz- und Strukturvergleich

```

245   250   255   260   265   270   275   280   285   290   295   300   305   310
ADAMTS1 RYVETMLVADQSMAEFHG.SGLKHYLLTLFVSAARLYKHPSIRNSVSLVVVKILVIHDEQK.GPEVTSNA
ADAMTS2 YNIEVLLGVDDSVVQFHGKEHVQKYLTLTLMNIVNEIYHDESLGAHINVVLVRIILLSYGKSMSLIEIGNP
ADAMTS3 YNIEVLLGVDDSVVRFHGKEHVQNYLLTLMNIVNEIYHDESLGVHINVVLVRMIMLGYAKSISLIERGNP
ADAMTS4 RFVETLVVADDKMAAFHG.AGLKRYLLTVMAAAAKAFKHPSIRNPVSLVVTRLVILGSGEE.GPQVGPSPA
ADAMTS5 RQVELLLVADASMARLYG.RGLQHYLLTLASIANRLYSHASIEHRI LAVVKKVVVLGDKDK.SLEVSKNA
ADAMTS6 RFVETLVVADKMMVGYHGRKDI EHYILSVMNIVAKLYRDSSSLGNVNNIIVARLIVLTEDQP.NLEINHHA
ADAMTS7 KWVETLVVADAKMVEYHGQPQVESYVLTIMNMVAGLFHDPSIGNP IHITIVRLVLLLEDEEE.DLKITHHA
ADAMTS8 RFVETLVVADASMAAFYG.ADLQNHILTLMSVAARIYKHPSIKNSINLMVVKVLIVEDEKW.GPEVSDNG
ADAMTS9 RFVEVLVADNRMVSYHG.ENLQHYILTLMNIVASIKDPSIGNLINIVNVLVIVHNEQD.GPSISFNA
ADAMTS10 RYVETLVVADKMMVAYHGRRDVEQYVLAIMNIVAKLFQDSSLGSTVNIILVTRLILLTEDEQP.TLEITHHA
ADAMTS12 RWVETLVVADTKMIEYHGSENVESYILTIMNMVGTGLFHNPSIGNA IHI VVVRLILLLEEEEQ.GLKIVHHA
ADAMTS13 LHLELLVAVGPDVFOAHO.EDTERYVLTNLNIGAELLRDP SLGAQFRVHLVKMVLITPEEG.APNITANL
ADAMTS14 YSIEVLLVDDSVVRFHGKEHVQNYVLTLMNIVDEIYHDESLGVHINIALVRLIMVGRYQSLSLIERGNP
ADAMTS15 RQVETLVVADKMMVEKMGKGNVTYILTVMNVSGLFKDGTIGSDINVVVVVLLLEQEPG.GLLINHHA
ADAMTS16 LNVETLVVVDKMMQNHGHENITTYVLTILNMSVAFKDGITGGNINIAIVGLLLEDEQP.GLVISHHA
ADAMTS17 HTVETLVVADADMVQYHGAEAAQRFILTVMMNVMVNFQHOSLGIKINIQTVKLVLLRORPA.KLSIGHHG
ADAMTS18 LNVETLVVADKMMVEKMGKGNVTYILTVMNVSGLFKDGTIGSDINVVVVVLLLEQEPG.GLLINHHA
ADAMTS19 YNIETVVVADPAMVSYHGADAARRFILTILNMFVNLFOHKSLSVQVNLRVIKLILLHETPP.ELYIGHHG
ADAMTS20 RYIETMVADAKVVS AHG.SNLQNYILTLMNIVATIKDPSIGNLIHIVVVKLVMIHREEE.GPVINFDG

315   320   325   330   335   340   345   350   355   360   365
ADAMTS1 ALTLRNF CNWQKQHNP . . . . . SDRDAEHYDTAIFLTRQDL CGS . . . . . QTCDDTLGMADVGTVC DPS
ADAMTS2 SQSLENVCRWAYLQQKP . . . . . DTGDEYHDHAI FLTRQDFGP . . . . . SGMQGYAPVTGMCHPV
ADAMTS3 SRSLENVCRWASQQORS . . . . . DLNHEHHDHAI FLTRQDFGP . . . . . AGMQGYAPVTGMCHPV
ADAMTS4 AOTLRSFCAWQRGLNTP . . . . . EDSDEPHFDTAIFLTRQDL CGV . . . . . STCDTLGMADVGTVC DPS
ADAMTS5 ATTLKNFCWQHQNQL . . . . . GDDHEEHYDAAILFTR EDL CGH . . . . . HSCDTLGMADVGTIC SPS
ADAMTS6 DKSLDSFCWKQKSILSHQSD . . . . . GNTIPENGAIHHDNAVLI TRYDI CTYKKNKPCGTLGLASVAGMCEPE
ADAMTS7 DNTLKSFCKWKQKSINMK . . . . . GDHPLHHDTAI LLTRKDLCAAMNRP CETLLGLSHVAGMCPQH
ADAMTS8 GLTLRNF CNWQRNFNP . . . . . SDRHPEHYDTA ILLTRQDFCGO . EGLCDTLGVADIGTICDPN
ADAMTS9 QTTLRNF CQWQHSKNSP . . . . . GGIHHD TAVLLTRQDI CRA . HDKCDTLGLAELGTICDPY
ADAMTS10 GKSLDSFCWKQKSIVNHS GH . . . . . GNAIPENGVANHD TAVLITRYDI CIYKKNKPCGTLGLAPVGGMCERE
ADAMTS12 EKTLSDFCKWKQKSINPK . . . . . SDLNPHHDVAVLLTRKDI CAGFNRP CETLLGLSHLSGMCPQH
ADAMTS13 TSSLLSFCWQWQSQTINPE . . . . . DDTDEPHADLVLYITRDFLELP . DGNRQVRVGTQVLTGGCASP T
ADAMTS14 SRSLEQVCRWAHSQORQ . . . . . DSPHAEHHDHVFLTRQDFGP . . . . . SGYAPVTGMCHPL
ADAMTS15 ALTLRNFCAWQKKNKV . . . . . SDKHPEYWDTA IFLTRQDL CGA . . . . . TTCDDTLGMADVGTMC DPK
ADAMTS16 DHTLSSFCWQWQSGLMGK . . . . . DGRHDHAI LLTGLDICSWKNEPCDTLGFAPISGMCSKY
ADAMTS17 ERSLESFCWQWNEEYGGARYLGNQVPGGKDDPPLVDA AVFVIRTD F CVHKDEPCDTVGIAYLGGVCSAK
ADAMTS18 DQSLNSFCWQWQSALIGK . . . . . NGKRHDHAI LLTGFDCISWKNEPCDTLGFAPISGMCSKY
ADAMTS19 EKMLSEFCWQWHEEF GKKNDIHEMSTNWGEDMTSVDAAILITR KDFCVHKDEPCDTVGIAYLGGVCSAK
ADAMTS20 ATTLKNFCSWQQTQNDL . . . . . DDVHP SHHD TAVLITR EDICSS . KEKCNMLGLSYLGTICDPL

370   375   380   385   390   395   400   405   410   415   420   425   430
ADAMTS1 RSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQCASLNGVNQ.DSHMMASMLSN.LDHSQPWSPCSAY
ADAMTS2 RSCTLNHEDGFSSAFVVAHETGHVVLGMEHDGQGNRCGD.EVR.LGSIMAPLVQA.AFHRFHWSRCSOQ
ADAMTS3 RSCTLNHEDGFSSAFVVAHETGHVVLGMEHDGQGNRCGD.ETA.MGSVMAPLVQA.AFHRYHWSRCSGO
ADAMTS4 RSCAIVEDDGLQSAFTA AHELGHVFNMLHDNSKPCISLNGPLSTRHVMA PVMAH.VDFEPEWSPCSAR
ADAMTS5 RSCAIVEDDGLHAAFTVAHEIGHLLGLSHDHSKFCFEETFGSTE.DKRLMSILTS.IDASKPWSKCTSA
ADAMTS6 RSCSINEDIGLGSAFTIAHEIGHNFGMNHDDIGNSCGT.KGHE.AAKLMAAHITA.NNPFWSWACS RD
ADAMTS7 RSCSINEDTGLPLAFTVAHELGHVSFGIQHDGSGNDCEP.VGK.RPFIMSPQLLY.DAAPLTWSCRSA R
ADAMTS8 KCSVIEDEGLQAAHTLAHELGHVLSMPHDDSKPCTRFLFGPMG.KHHVMAPLFVH.LNQTLPWSPCSAM
ADAMTS9 RSCSINEDSGLSTAFTIAHELGHVFNMPHDDNNKCKE.E.GVKS.PQHVMAPTLNF.YTNPWMSKCSRK
ADAMTS10 RSCSVNEDIGLATAFTIAHEIGHTFGMNHDDVGNSCGA.RGQD.PAKLMAAHITM.KTNPFVWSSCSR D
ADAMTS12 RSCNINEDSGLPLAFTIAHELGHVSFGIQHDGNDCEP.VGR.HPYIMSRQLQY.DPTPLTWSCSRE
ADAMTS13 WSCLIITEDTGFDLGVTIAHEIGHVSFGI QHDGAPGSGCGP.SGHVMSADGAA.PRAGLWSPCSR R
ADAMTS14 RSCALNHEDGFSSAFVIAHETGHVVLGMEHDGQGNCCAD.ETS.LGSVMAPLVQA.AFHRFHWSRCSKL
ADAMTS15 RSCSVIEDDGLPSAFTTAHELGHVFNMPHDDNVRKCEEVFGKLR.ANHMMSPTLIQ.IDRANPWSACSAA
ADAMTS16 RSCTINEDTGLGLAFTIAHESGHNFMIHDGEGNMCKKS.EGNIMSPTLAG.RNGVFSWSPCSRQ
ADAMTS17 RKCVLAEDNGLNLAFTIAHELGHNLGMNHDDHSSCAG.RSHIMSGEWEVKGRNPDLWSWSSCSR D
ADAMTS18 RSCTINEDTGLGLAFTIAHESGHNFMIHDGEGNPRKA.EGNIMSPTLTG.NNGVFSWSSCSRQ
ADAMTS19 RKCIIAEDNGLNLAFTIAHEMGNMGINHDNDHPSCAD.GLHIMSGEWIKQNLGDVSWSRCSKE
ADAMTS20 QSCFINEEKGLISAFTIAHELGHITLGVQHDNDNPRCKE.MKVT.KYHVMAPALSF.HMSPWWSNSCSR K

435   440   445   450
ADAMTS1 MITSFLDNGHGECCLMDKP
ADAMTS2 ELSRYLHSY.DCLLDDP
ADAMTS3 ELKRYIHSY.DCLLDDP
ADAMTS4 FITDFLDNGYGHCLLDDK
ADAMTS5 TITFDLDDGHGNCLLDLP
ADAMTS6 YITFDLDSGRGTCLDNPE
ADAMTS7 YITRFLDRGWGLCLDDPP
ADAMTS8 YLTELDDGGHGDCCLLDDAP
ADAMTS9 YITEFLDTGYGECCLLNEP
ADAMTS10 YITFDLDSGLGLCLNRRP
ADAMTS12 YITRFLDRGWGFCLDDIP
ADAMTS13 QLLSLLSAGRARCVDWPP
ADAMTS14 ELSRYLPSY.DCLLDDP
ADAMTS15 IITDFLD SGHGDCCLLDDQ
ADAMTS16 YLHKFLSTAQAICLADQP
ADAMTS17 DLENFLKSKVSTCLLVTD
ADAMTS18 YLKKFLSTPQAGCLVDEP
ADAMTS19 DLERFLRSKASNCLLQTN
ADAMTS20 YVTEFLDTGYGECCLLDDK

```

Abbildung A.5: Multiples Sequenzalignment der Metalloproteinase-Domänen aller ADAMTS-Proteine. Die Sequenzen dieser ADAMTS-Domänen stammen aus der UniProt-Datenbank [169]. Die Nummerierung des Alignments bezieht sich auf ADAMTS7.

Tabelle A.1: Paarweiser Vergleich der Sequenzidentitäten des Bereichs um das β -Faltblatt S4 (ADAMTS7: Met350-Met365) in der ADAMTS-Bindetasche (in Prozent). Die Identitätswerte sind farbkodiert: 100 % identisch - orange, 0 % identisch - weiß.

Sequenzidentität	ADAMTS1	ADAMTS2	ADAMTS3	ADAMTS4	ADAMTS5	ADAMTS6	ADAMTS7	ADAMTS8	ADAMTS9	ADAMTS10	ADAMTS12	ADAMTS13	ADAMTS14	ADAMTS15	ADAMTS16	ADAMTS17	ADAMTS18	ADAMTS19	ADAMTS20
ADAMTS1	100	30,8	30,8	92,9	80,0	46,7	40,0	66,7	60,0	53,3	33,3	20,0	40,0	86,7	46,7	53,3	46,7	40,0	40,0
ADAMTS2	30,8	100	92,3	23,1	30,8	53,8	38,5	23,1	23,1	61,5	30,8	23,1	90,0	38,5	46,2	30,8	46,2	38,5	23,1
ADAMTS3	30,8	92,3	100	23,1	30,8	53,8	38,5	23,1	23,1	61,5	30,8	23,1	90,0	38,5	46,2	30,8	46,2	38,5	23,1
ADAMTS4	92,9	23,1	23,1	100	78,6	42,9	35,7	64,3	57,1	50,0	28,6	14,3	30,0	85,7	42,9	50,0	42,9	35,7	35,7
ADAMTS5	80,0	30,8	30,8	78,6	100	46,7	40,0	73,3	66,7	53,3	33,3	20,0	40,0	80,0	46,7	46,7	46,7	40,0	46,7
ADAMTS6	46,7	53,8	53,8	42,9	46,7	100	70,6	37,5	43,8	88,2	58,8	18,8	60,0	53,3	64,7	47,1	64,7	52,9	31,3
ADAMTS7	40,0	38,5	38,5	35,7	40,0	70,6	100	31,3	37,5	64,7	82,4	18,8	50,0	46,7	52,9	35,3	52,9	41,2	37,5
ADAMTS8	66,7	23,1	23,1	64,3	73,3	37,5	31,3	100	62,5	43,8	31,3	31,3	30,0	66,7	50,0	43,8	50,0	37,5	43,8
ADAMTS9	60,0	23,1	23,1	57,1	66,7	43,8	37,5	62,5	100	50,0	43,8	25,0	30,0	60,0	43,8	50,0	43,8	43,8	62,5
ADAMTS10	53,3	61,5	61,5	50,0	53,3	88,2	64,7	43,8	50,0	100	58,8	25,0	70,0	60,0	70,6	52,9	70,6	52,9	37,5
ADAMTS12	33,3	30,8	30,8	28,6	33,3	58,8	82,4	31,3	43,8	58,8	100	25,0	40,0	40,0	58,8	41,2	58,8	52,9	43,8
ADAMTS13	20,0	23,1	23,1	14,3	20,0	18,8	18,8	31,3	25,0	25,0	25,0	100	30,0	20,0	18,8	37,5	18,8	31,3	25,0
ADAMTS14	40,0	90,0	90,0	30,0	40,0	60,0	50,0	30,0	30,0	70,0	40,0	30,0	100	50,0	60,0	40,0	60,0	50,0	20,0
ADAMTS15	86,7	38,5	38,5	85,7	80,0	53,3	46,7	66,7	60,0	60,0	40,0	20,0	50,0	100	53,3	46,7	53,3	46,7	40,0
ADAMTS16	46,7	46,2	46,2	42,9	46,7	64,7	52,9	50,0	43,8	70,6	58,8	18,8	60,0	53,3	100	58,8	100	70,6	31,3
ADAMTS17	53,3	30,8	30,8	50,0	46,7	47,1	35,3	43,8	50,0	52,9	41,2	37,5	40,0	46,7	58,8	100	58,8	88,2	43,8
ADAMTS18	46,7	46,2	46,2	42,9	46,7	64,7	52,9	50,0	43,8	70,6	58,8	18,8	60,0	53,3	100	58,8	100	70,6	31,3
ADAMTS19	40,0	38,5	38,5	35,7	40,0	52,9	41,2	37,5	43,8	52,9	52,9	31,3	50,0	46,7	70,6	88,2	70,6	100	37,5
ADAMTS20	40,0	23,1	23,1	35,7	46,7	31,3	37,5	43,8	62,5	37,5	43,8	25,0	20,0	40,0	31,3	43,8	31,3	37,5	100

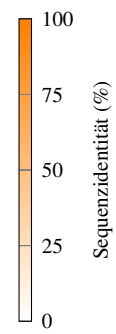


Tabelle A.2: Paarweiser Vergleich der Sequenzidentitäten des Bereichs um die α -Helix H4 (ADAMTS7: Leu380-Gly400) in der ADAMTS-Bindetasche (in Prozent). Die Identitätswerte sind farbkodiert: 100 % identisch - orange, 0 % identisch - weiß.

Sequenzidentität	ADAMTS1	ADAMTS2	ADAMTS3	ADAMTS4	ADAMTS5	ADAMTS6	ADAMTS7	ADAMTS8	ADAMTS9	ADAMTS10	ADAMTS12	ADAMTS13	ADAMTS14	ADAMTS15	ADAMTS16	ADAMTS17	ADAMTS18	ADAMTS19	ADAMTS20
ADAMTS1	100	52,4	52,4	81,0	61,9	61,9	61,9	81,0	85,7	61,9	61,9	42,9	52,4	85,7	61,9	66,7	61,9	52,4	61,9
ADAMTS2	52,4	100	100	57,1	57,1	61,9	57,1	52,4	57,1	57,1	52,4	52,4	95,2	57,1	57,1	57,1	57,1	47,6	57,1
ADAMTS3	52,4	100	100	57,1	57,1	61,9	57,1	52,4	57,1	57,1	52,4	52,4	95,2	57,1	57,1	57,1	57,1	47,6	57,1
ADAMTS4	81,0	57,1	57,1	100	52,4	66,7	61,9	66,7	76,2	61,9	61,9	42,9	57,1	85,7	61,9	61,9	61,9	57,1	61,9
ADAMTS5	61,9	57,1	57,1	52,4	100	61,9	61,9	61,9	57,1	61,9	57,1	52,4	52,4	52,4	57,1	66,7	57,1	57,1	66,7
ADAMTS6	61,9	61,9	61,9	66,7	61,9	100	66,7	52,4	66,7	85,7	71,4	61,9	66,7	66,7	85,7	76,2	85,7	71,4	66,7
ADAMTS7	61,9	57,1	57,1	61,9	61,9	66,7	100	52,4	61,9	66,7	95,2	61,9	52,4	66,7	71,4	66,7	71,4	66,7	66,7
ADAMTS8	81,0	52,4	52,4	66,7	61,9	52,4	52,4	100	71,4	52,4	52,4	38,1	52,4	66,7	52,4	66,7	52,4	47,6	61,9
ADAMTS9	85,7	57,1	57,1	76,2	57,1	66,7	61,9	71,4	100	71,4	66,7	47,6	61,9	81,0	66,7	71,4	66,7	57,1	66,7
ADAMTS10	61,9	57,1	57,1	61,9	61,9	85,7	66,7	52,4	71,4	100	71,4	61,9	61,9	61,9	76,2	71,4	76,2	66,7	66,7
ADAMTS12	61,9	52,4	52,4	61,9	57,1	71,4	95,2	52,4	66,7	71,4	100	66,7	57,1	66,7	76,2	71,4	76,2	71,4	71,4
ADAMTS13	42,9	52,4	52,4	42,9	52,4	61,9	61,9	38,1	47,6	61,9	66,7	100	57,1	42,9	61,9	52,4	61,9	52,4	47,6
ADAMTS14	52,4	95,2	95,2	57,1	52,4	66,7	52,4	52,4	61,9	61,9	57,1	57,1	100	57,1	61,9	61,9	61,9	52,4	61,9
ADAMTS15	85,7	57,1	57,1	85,7	52,4	66,7	66,7	66,7	81,0	61,9	66,7	42,9	57,1	100	61,9	61,9	61,9	57,1	61,9
ADAMTS16	61,9	57,1	57,1	61,9	57,1	85,7	71,4	52,4	66,7	76,2	76,2	61,9	61,9	61,9	100	76,2	100	71,4	61,9
ADAMTS17	66,7	57,1	57,1	61,9	66,7	76,2	66,7	66,7	71,4	71,4	71,4	52,4	61,9	61,9	76,2	100	76,2	81,0	76,2
ADAMTS18	61,9	57,1	57,1	61,9	57,1	85,7	71,4	52,4	66,7	76,2	76,2	61,9	61,9	61,9	100	76,2	100	71,4	61,9
ADAMTS19	52,4	47,6	47,6	57,1	57,1	71,4	66,7	47,6	57,1	66,7	71,4	52,4	52,4	57,1	71,4	81,0	71,4	100	61,9
ADAMTS20	61,9	57,1	57,1	61,9	66,7	66,7	66,7	61,9	66,7	66,7	71,4	47,6	61,9	61,9	61,9	76,2	61,9	61,9	100

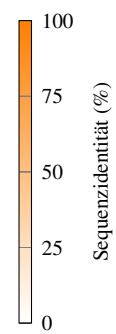
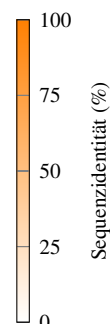


Tabelle A.3: Paarweiser Vergleich der Sequenzidentitäten der Loop-Region L1 (ADAMTS7: Ser401-Arg430) in der ADAMTS-Bindetasche (in Prozent). Die Identitätswerte sind farbkodiert: 100 % identisch - orange, 0 % identisch - weiß.

Sequenzidentität	ADAMTS1	ADAMTS2	ADAMTS3	ADAMTS4	ADAMTS5	ADAMTS6	ADAMTS7	ADAMTS8	ADAMTS9	ADAMTS10	ADAMTS12	ADAMTS13	ADAMTS14	ADAMTS15	ADAMTS16	ADAMTS17	ADAMTS18	ADAMTS19	ADAMTS20
ADAMTS1	100	16,7	16,7	43,8	37,5	12,9	16,7	37,5	29,0	12,9	16,7	28,6	16,7	34,4	17,9	17,2	14,3	13,8	26,7
ADAMTS2	16,7	100	83,3	16,7	10,0	33,3	30,0	20,0	16,7	30,0	26,7	21,4	83,3	16,7	35,7	17,2	35,7	24,1	20,0
ADAMTS3	16,7	83,3	100	20,0	13,3	30,0	26,7	23,3	20,0	26,7	20,0	25,0	83,3	13,3	28,6	13,8	28,6	20,7	20,0
ADAMTS4	43,8	16,7	20,0	100	28,1	12,9	16,7	53,1	32,3	12,9	10,0	32,1	23,3	34,4	21,4	13,8	14,3	17,2	26,7
ADAMTS5	37,5	10,0	13,3	28,1	100	22,6	26,7	28,1	25,8	16,1	23,3	17,9	13,3	43,8	17,9	13,8	21,4	13,8	20,0
ADAMTS6	12,9	33,3	30,0	12,9	22,6	100	23,3	12,9	22,6	67,7	20,0	21,4	30,0	16,1	28,6	24,1	35,7	24,1	20,0
ADAMTS7	16,7	30,0	26,7	16,7	26,7	23,3	100	20,0	26,7	23,3	66,7	17,9	26,7	23,3	35,7	27,6	35,7	24,1	23,3
ADAMTS8	37,5	20,0	23,3	53,1	28,1	12,9	20,0	100	29,0	12,9	13,3	28,6	23,3	31,3	21,4	20,7	21,4	17,2	30,0
ADAMTS9	29,0	16,7	20,0	32,3	25,8	22,6	26,7	29,0	100	25,8	26,7	21,4	23,3	32,3	21,4	13,8	21,4	13,8	50,0
ADAMTS10	12,9	30,0	26,7	12,9	16,1	67,7	23,3	12,9	25,8	100	20,0	17,9	26,7	12,9	25,0	20,7	32,1	17,2	16,7
ADAMTS12	16,7	26,7	20,0	10,0	23,3	20,0	66,7	13,3	26,7	20,0	100	14,3	20,0	23,3	28,6	27,6	28,6	20,7	20,0
ADAMTS13	28,6	21,4	25,0	32,1	17,9	21,4	17,9	28,6	21,4	17,9	14,3	100	25,0	14,3	21,4	25,0	14,3	17,9	25,0
ADAMTS14	16,7	83,3	83,3	23,3	13,3	30,0	26,7	23,3	23,3	26,7	20,0	25,0	100	13,3	32,1	17,2	32,1	24,1	20,0
ADAMTS15	34,4	16,7	13,3	34,4	43,8	16,1	23,3	31,3	32,3	12,9	23,3	14,3	13,3	100	25,0	17,2	25,0	20,7	30,0
ADAMTS16	17,9	35,7	28,6	21,4	17,9	28,6	35,7	21,4	21,4	25,0	28,6	21,4	32,1	25,0	100	25,0	78,6	28,6	21,4
ADAMTS17	17,2	17,2	13,8	13,8	13,8	24,1	27,6	20,7	13,8	20,7	27,6	25,0	17,2	17,2	25,0	100	32,1	65,5	20,7
ADAMTS18	14,3	35,7	28,6	14,3	21,4	35,7	35,7	21,4	21,4	32,1	28,6	14,3	32,1	25,0	78,6	32,1	100	32,1	21,4
ADAMTS19	13,8	24,1	20,7	17,2	13,8	24,1	24,1	17,2	13,8	17,2	20,7	17,9	24,1	20,7	28,6	65,5	32,1	100	17,2
ADAMTS20	26,7	20,0	20,0	26,7	20,0	20,0	23,3	30,0	50,0	16,7	20,0	25,0	20,0	30,0	21,4	20,7	21,4	17,2	100

Tabelle A.4: Strukturvergleich des Bereiches um das β -Faltblatt S4 (ADAMTS7: Met350-Met365) anhand des $RMSD_{C\alpha}$ in Å. Die strukturellen Unterschiede zwischen den Proteinen sind mittels eines Farbverlaufs dargestellt: Von 0 Å Unterschied - blau zu 10 Å Unterschied - weiß bis 20 Å Unterschied - rot. * - Röntgenkristallstruktur.

$RMSD_{C\alpha}$	ADAMTS1*	ADAMTS2	ADAMTS3	ADAMTS4*	ADAMTS5*	ADAMTS6	ADAMTS7	ADAMTS8	ADAMTS9	ADAMTS10	ADAMTS12	ADAMTS13	ADAMTS14	ADAMTS15	ADAMTS16	ADAMTS17	ADAMTS18	ADAMTS19	ADAMTS20
ADAMTS1*	0,0	6,0	6,2	0,5	0,9	6,1	6,0	3,8	4,1	6,5	6,5	4,2	14,3	1,0	6,8	7,0	7,5	7,0	3,9
ADAMTS2	6,0	0,0	1,5	6,1	5,9	10,3	10,2	8,4	8,5	10,7	10,7	8,1	9,9	6,3	11,2	11,2	11,7	11,4	8,0
ADAMTS3	6,2	1,5	0,0	6,2	6,1	10,5	10,5	8,6	8,8	11,0	11,0	8,5	10,1	6,5	11,5	11,6	12,0	11,7	8,3
ADAMTS4*	0,5	6,1	6,2	0,0	0,8	6,1	6,1	3,9	4,2	6,5	6,5	4,3	14,3	1,2	6,9	7,0	7,6	7,1	3,9
ADAMTS5*	0,9	5,9	6,1	0,8	0,0	6,2	6,1	4,2	4,3	6,8	6,7	4,4	13,8	1,1	6,9	7,0	7,6	7,1	4,2
ADAMTS6	6,1	10,3	10,5	6,1	6,2	0,0	1,3	4,1	3,9	2,7	2,4	4,1	17,0	5,8	1,6	1,9	2,4	1,8	4,3
ADAMTS7	6,0	10,2	10,5	6,1	6,1	1,3	0,0	4,0	3,8	2,4	2,1	4,0	16,8	5,8	1,7	1,8	2,3	1,7	4,2
ADAMTS8	3,8	8,4	8,6	3,9	4,2	4,1	4,0	0,0	1,7	4,1	4,2	2,0	16,4	3,7	4,5	4,7	5,2	4,7	1,6
ADAMTS9	4,1	8,5	8,8	4,2	4,3	3,9	3,8	1,7	0,0	3,8	3,9	1,1	16,0	3,9	4,0	4,1	4,6	4,2	1,2
ADAMTS10	6,5	10,7	11,0	6,5	6,8	2,7	2,4	4,1	3,8	0,0	1,1	4,0	17,6	6,3	2,8	3,0	2,9	2,7	3,9
ADAMTS12	6,5	10,7	11,0	6,5	6,7	2,4	2,1	4,2	3,9	1,1	0,0	4,2	17,5	6,3	2,5	2,8	2,7	2,6	4,1
ADAMTS13	4,2	8,1	8,5	4,3	4,4	4,1	4,0	2,0	1,1	4,0	4,2	0,0	15,6	3,9	4,2	4,3	4,6	4,4	1,7
ADAMTS14	14,3	9,9	10,1	14,3	13,8	17,0	16,8	16,4	16,0	17,6	17,5	15,6	0,0	14,2	17,9	17,6	18,0	17,9	15,6
ADAMTS15	1,0	6,3	6,5	1,2	1,1	5,8	5,8	3,7	3,9	6,3	6,3	3,9	14,2	0,0	6,5	6,6	7,2	6,7	3,8
ADAMTS16	6,8	11,2	11,5	6,9	6,9	1,6	1,7	4,5	4,0	2,8	2,5	4,2	17,9	6,5	0,0	1,3	1,3	1,4	4,7
ADAMTS17	7,0	11,2	11,6	7,0	7,0	1,9	1,8	4,7	4,1	3,0	2,8	4,3	17,6	6,6	1,3	0,0	1,8	0,7	4,8
ADAMTS18	7,5	11,7	12,0	7,6	7,6	2,4	2,3	5,2	4,6	2,9	2,7	4,6	18,0	7,2	1,3	1,8	0,0	1,8	5,3
ADAMTS19	7,0	11,4	11,7	7,1	7,1	1,8	1,7	4,7	4,2	2,7	2,6	4,4	17,9	6,7	1,4	0,7	1,8	0,0	4,8
ADAMTS20	3,9	8,0	8,3	3,9	4,2	4,3	4,2	1,6	1,2	3,9	4,1	1,7	15,6	3,8	4,7	4,8	5,3	4,8	0,0

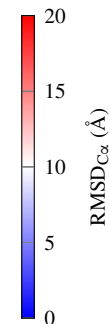


Tabelle A.5: Strukturvergleich des Bereiches um die α -Helix H4 (ADAMTS7: Leu380-Gly400) anhand des $RMSD_{C\alpha}$ in Å. Die strukturellen Unterschiede zwischen den Proteinen sind mittels eines Farbverlaufs dargestellt: Von 0 Å Unterschied - blau zu 10 Å Unterschied - weiß bis 20 Å Unterschied - rot. * - Röntgenkristallstruktur.

$RMSD_{C\alpha}$	ADAMTS1*	ADAMTS2	ADAMTS3	ADAMTS4*	ADAMTS5*	ADAMTS6	ADAMTS7	ADAMTS8	ADAMTS9	ADAMTS10	ADAMTS12	ADAMTS13	ADAMTS14	ADAMTS15	ADAMTS16	ADAMTS17	ADAMTS18	ADAMTS19	ADAMTS20
ADAMTS1*	0,0	0,9	1,2	0,9	1,0	0,6	0,8	0,6	0,5	3,8	0,9	0,5	1,1	0,5	0,9	0,7	1,0	0,9	0,8
ADAMTS2	0,9	0,0	0,8	1,2	1,1	0,7	0,7	0,9	0,9	3,8	1,2	0,9	1,2	0,9	1,1	1,0	1,2	1,0	0,9
ADAMTS3	1,2	0,8	0,0	1,4	1,3	0,8	0,9	1,1	1,3	4,0	1,3	1,0	1,1	1,1	1,2	1,3	1,2	1,3	1,2
ADAMTS4*	0,9	1,2	1,4	0,0	1,4	1,0	1,0	0,8	1,2	3,8	1,0	0,9	1,2	0,9	1,2	0,8	0,9	1,1	0,7
ADAMTS5*	1,0	1,1	1,3	1,4	0,0	0,9	1,1	1,3	1,0	3,9	1,7	1,0	1,5	1,3	1,4	1,1	1,6	1,3	1,3
ADAMTS6	0,6	0,7	0,8	1,0	0,9	0,0	0,8	0,8	0,9	3,9	1,1	0,6	1,1	0,8	1,0	1,0	1,1	1,1	0,9
ADAMTS7	0,8	0,7	0,9	1,0	1,1	0,8	0,0	0,8	0,8	3,8	1,1	0,8	0,9	0,9	0,7	0,8	0,9	1,0	0,7
ADAMTS8	0,6	0,9	1,1	0,8	1,3	0,8	0,8	0,0	0,8	3,9	0,8	0,7	1,2	0,5	0,9	0,8	1,0	1,0	0,8
ADAMTS9	0,5	0,9	1,3	1,2	1,0	0,9	0,8	0,8	0,0	3,9	1,0	0,7	1,3	0,7	0,9	0,8	1,2	1,0	0,8
ADAMTS10	3,8	3,8	4,0	3,8	3,9	3,9	3,8	3,9	3,9	0,0	3,8	4,0	3,9	3,9	4,1	3,8	4,0	3,7	3,8
ADAMTS12	0,9	1,2	1,3	1,0	1,7	1,1	1,1	0,8	1,0	3,8	0,0	1,1	1,2	0,8	1,0	1,2	1,1	1,3	0,9
ADAMTS13	0,5	0,9	1,0	0,9	1,0	0,6	0,8	0,7	0,7	4,0	1,1	0,0	1,1	0,6	0,8	0,8	1,0	1,0	0,8
ADAMTS14	1,1	1,2	1,1	1,2	1,5	1,1	0,9	1,2	1,3	3,9	1,2	1,1	0,0	1,3	0,9	1,1	0,7	1,2	1,0
ADAMTS15	0,5	0,9	1,1	0,9	1,3	0,8	0,9	0,5	0,7	3,9	0,8	0,6	1,3	0,0	1,0	0,8	1,0	1,0	0,8
ADAMTS16	0,9	1,1	1,2	1,2	1,4	1,0	0,7	0,9	0,9	4,1	1,0	0,8	0,9	1,0	0,0	1,0	0,9	1,2	1,0
ADAMTS17	0,7	1,0	1,3	0,8	1,1	1,0	0,8	0,8	0,8	3,8	1,2	0,8	1,1	0,8	1,0	0,0	1,0	0,6	0,7
ADAMTS18	1,0	1,2	1,2	0,9	1,6	1,1	0,9	1,0	1,2	4,0	1,1	1,0	0,7	1,0	0,9	1,0	0,0	1,0	0,9
ADAMTS19	0,9	1,0	1,3	1,1	1,3	1,1	1,0	1,0	1,0	3,7	1,3	1,0	1,2	1,0	1,2	0,6	1,0	0,0	0,9
ADAMTS20	0,8	0,9	1,2	0,7	1,3	0,9	0,7	0,8	0,8	3,8	0,9	0,8	1,0	0,8	1,0	0,7	0,9	0,9	0,0

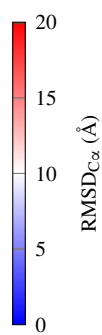
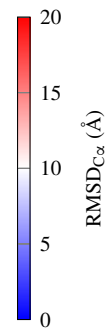
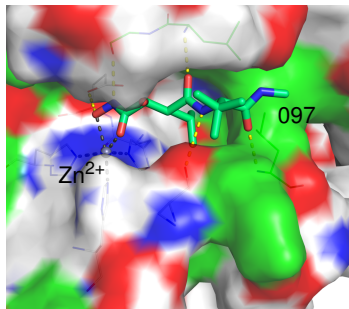


Tabelle A.6: Strukturvergleich des Bereiches um den L1-Loop (ADAMTS7: Ser401-Arg430) anhand des $RMSD_{C\alpha}$ in Å. Die strukturellen Unterschiede zwischen den Proteinen sind mittels eines Farbverlaufs dargestellt: Von 0 Å Unterschied - blau zu 10 Å Unterschied - weiß bis 20 Å Unterschied - rot. * - Röntgenkristallstruktur.

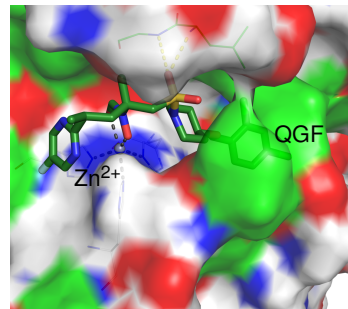
$RMSD_{C\alpha}$	ADAMTS1*	ADAMTS2	ADAMTS3	ADAMTS4*	ADAMTS5*	ADAMTS6	ADAMTS7	ADAMTS8	ADAMTS9	ADAMTS10	ADAMTS12	ADAMTS13	ADAMTS14	ADAMTS15	ADAMTS16	ADAMTS17	ADAMTS18	ADAMTS19	ADAMTS20
ADAMTS1*	0,0	5,3	5,8	8,6	2,5	3,7	6,8	2,6	3,6	3,3	5,2	8,9	5,5	3,0	8,3	8,8	8,8	8,9	5,8
ADAMTS2	5,3	0,0	3,9	9,5	5,6	4,0	5,6	5,6	4,6	4,1	3,8	7,4	4,2	6,4	6,6	7,0	7,1	7,1	4,0
ADAMTS3	5,8	3,9	0,0	10,5	5,9	4,5	3,3	5,3	4,2	4,4	2,9	6,1	3,2	6,7	6,0	6,0	6,6	6,4	3,3
ADAMTS4*	8,6	9,5	10,5	0,0	8,5	9,5	11,4	9,3	10,0	9,4	10,5	11,5	10,7	8,1	11,1	12,0	11,5	12,1	10,6
ADAMTS5*	2,5	5,6	5,9	8,5	0,0	3,3	7,3	3,3	4,5	4,1	5,6	9,5	6,4	3,2	8,7	9,0	9,3	9,5	6,3
ADAMTS6	3,7	4,0	4,5	9,5	3,3	0,0	6,0	3,6	3,2	2,8	4,0	8,5	4,9	4,5	7,7	8,0	8,2	8,2	4,8
ADAMTS7	6,8	5,6	3,3	11,4	7,3	6,0	0,0	6,0	5,1	5,6	4,3	5,9	4,3	7,5	6,6	5,8	7,1	6,3	4,6
ADAMTS8	2,6	5,6	5,3	9,3	3,3	3,6	6,0	0,0	3,7	3,7	4,7	8,0	5,2	3,3	7,5	7,8	7,9	7,9	5,4
ADAMTS9	3,6	4,6	4,2	10,0	4,5	3,2	5,1	3,7	0,0	2,4	3,8	7,5	3,9	5,2	7,3	7,1	7,8	7,5	4,0
ADAMTS10	3,3	4,1	4,4	9,4	4,1	2,8	5,6	3,7	2,4	0,0	3,7	7,6	3,8	4,7	7,0	7,6	7,4	7,7	4,1
ADAMTS12	5,2	3,8	2,9	10,5	5,6	4,0	4,3	4,7	3,8	3,7	0,0	6,1	2,6	6,2	5,6	6,1	6,2	6,3	2,4
ADAMTS13	8,9	7,4	6,1	11,5	9,5	8,5	5,9	8,0	7,5	7,6	6,1	0,0	5,6	9,3	3,0	4,3	3,5	4,4	5,8
ADAMTS14	5,5	4,2	3,2	10,7	6,4	4,9	4,3	5,2	3,9	3,8	2,6	5,6	0,0	6,7	5,2	6,0	5,5	6,0	2,1
ADAMTS15	3,0	6,4	6,7	8,1	3,2	4,5	7,5	3,3	5,2	4,7	6,2	9,3	6,7	0,0	8,6	9,2	9,1	9,3	7,0
ADAMTS16	8,3	6,6	6,0	11,1	8,7	7,7	6,6	7,5	7,3	7,0	5,6	3,0	5,2	8,6	0,0	5,1	1,5	4,8	5,1
ADAMTS17	8,8	7,0	6,0	12,0	9,0	8,0	5,8	7,8	7,1	7,6	6,1	4,3	6,0	9,2	5,1	0,0	5,6	2,6	6,2
ADAMTS18	8,8	7,1	6,6	11,5	9,3	8,2	7,1	7,9	7,8	7,4	6,2	3,5	5,5	9,1	1,5	5,6	0,0	5,1	5,4
ADAMTS19	8,9	7,1	6,4	12,1	9,5	8,2	6,3	7,9	7,5	7,7	6,3	4,4	6,0	9,3	4,8	2,6	5,1	0,0	6,2
ADAMTS20	5,8	4,0	3,3	10,6	6,3	4,8	4,6	5,4	4,0	4,1	2,4	5,8	2,1	7,0	5,1	6,2	5,4	6,2	0,0



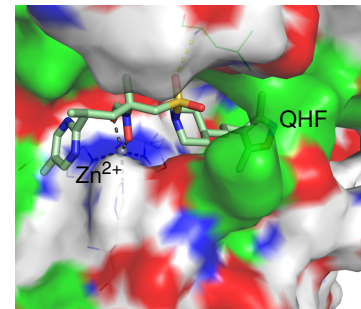
A.2.2 Bindemuster von bekannten Inhibitoren in ADAMTS-Proteinen



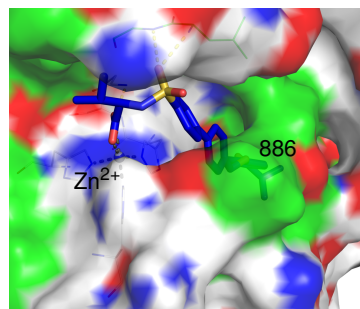
(a) ADAMTS1 (PDB-ID: 2JIH)



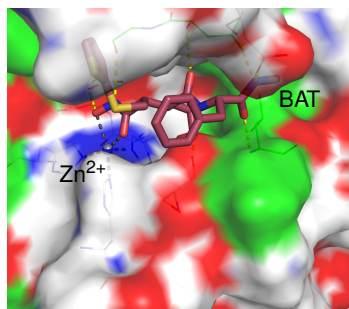
(b) ADAMTS1 (PDB-ID: 3Q2G)



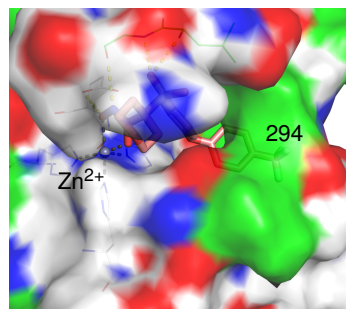
(c) ADAMTS1 (PDB-ID: 3Q2H)



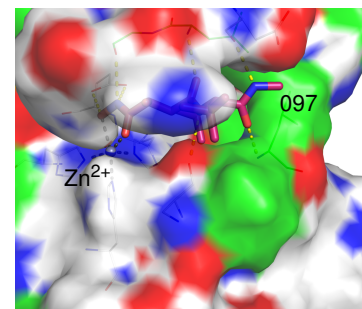
(d) ADAMTS4 (PDB-ID: 2RJP)



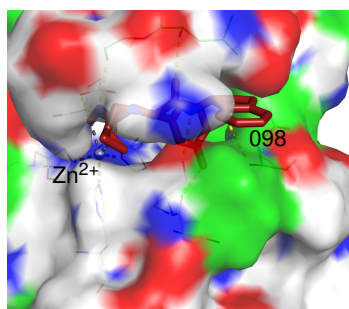
(e) ADAMTS5 (PDB-ID: 2RJQ)



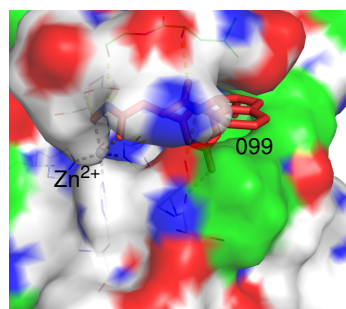
(f) ADAMTS5 (PDB-ID: 3B8Z)



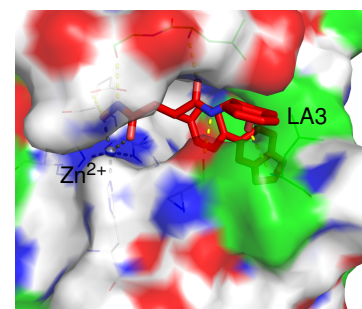
(g) ADAMTS5 (PDB-ID: 3HY7)



(h) ADAMTS5 (PDB-ID: 3HY9)



(i) ADAMTS5 (PDB-ID: 3HYG)



(j) ADAMTS5 (PDB-ID: 3LJT)

Abbildung A.6: Oberflächendarstellung (grün: hydrophober Bereich; rot: Akzeptorbereich; blau: Donorbereich) der bekannten mit Inhibitor kokristallisierten ADAMTS-Röntgenkristallstrukturen. Die Bilder sind analog zu denen in Abbildung 4.2 um 15° nach oben gedreht. Die Zink-koordinierenden Histidine sind als Linien veranschaulicht, genau wie die jeweils mit dem Inhibitor wechselwirkenden Aminosäuren des Rezeptors. Die Inhibitoren werden durch Stäbchen dargestellt und die Zink-Ionen sind als graue Kugeln zu erkennen. Zink-koordinierende Bindungen sind durch grau gestrichelte Linien dargestellt, Wasserstoffbrücken durch gelb gestrichelte Linien.

A.3 Ergebnisse: Virtual Screening

A.3.1 Redocking zur Evaluierung von Dockingmethoden

Redocking mit AutoDock4 und AutoDockVina

AutodDock4: Docking-Parameter-Datei (dpf-Datei):

```

1 autodock_parameter_version 4.2      # used by autodock to validate parameter set
2 outlev 1                            # diagnostic output level
3 intelec                             # calculate internal electrostatics
4 seed pid time                       # seeds for random generator
5 ligand_types C OA HD N             # atoms types in ligand
6 fld Rezeptor_H.maps.fld            # grid_data_file
7 map Rezeptor_H.C.map               # atom-specific affinity map
8 map Rezeptor_H.OA.map              # atom-specific affinity map
9 map Rezeptor_H.HD.map              # atom-specific affinity map
10 map Rezeptor_H.N.map              # atom-specific affinity map
11 elecmap Rezeptor_H.e.map          # electrostatics map
12 desolvmap Rezeptor_H.d.map         # desolvation map
13 move Ligand.pdbqt                 # small molecule
14 about x1 y1 z1                    # small molecule center
15 tran0 random                      # initial coordinates/A or random
16 quaternion0 random                # initial orientation
17 dihe0 random                      # initial dihedrals (relative) or random
18 torsdof n                         # torsional degrees of freedom
19 rmstol 2.0                        # cluster_tolerance/A
20 extnrg 1000.0                     # external grid energy
21 e0max 0.0 10000                   # max initial energy; max number of retries
22 ga_pop_size 150                   # number of individuals in population
23 ga_num_evals 2500000              # maximum number of energy evaluations
24 ga_num_generations 27000          # maximum number of generations
25 ga_elitism 1                      # number of top individuals to survive to next generation
26 ga_mutation_rate 0.02             # rate of gene mutation
27 ga_crossover_rate 0.8             # rate of crossover
28 ga_window_size 10                 #
29 ga_cauchy_alpha 0.0               # Alpha parameter of Cauchy distribution
30 ga_cauchy_beta 1.0               # Beta parameter Cauchy distribution
31 set_ga                             # set the above parameters for GA or LGA
32 sw_max_its 300                    # iterations of Solis & Wets local search
33 sw_max_succ 4                     # consecutive successes before changing rho
34 sw_max_fail 4                     # consecutive failures before changing rho
35 sw_rho 1.0                        # size of local search space to sample
36 sw_lb_rho 0.01                   # lower bound on rho
37 ls_search_freq 0.06               # probability of performing local search on individual
38 set_pswl                           # set the above pseudo-Solis & Wets parameters
39 unbound_model bound               # state of unbound ligand
40 ga_run 10                         # do this many hybrid GA-LS runs
41 analysis                           # perform a ranked cluster analysis

```

Folgend sind die Befehle für die Erstellung von Grid-Dateien (gpf-Datei) und das Docking mit AutoDock4 aufgeführt:

```

1 autogrid4 -p Rezeptor.gpf -l Rezeptor.glg
2 autodock4 -p Rezeptor.dpf -l Rezeptor.dlg

```

AutoDockVina: Aufruf für das Docking:

```

1 vina --config config.txt --log log.txt

```

Die *config.txt* beinhaltet die Informationen über Protein- und Liganddateien, sowie Parameter für das Grid:

```

1 receptor = Rezeptor.pdbqt
2 ligand = Ligand.pdbqt
3 out = Rezeptor-Ligand-Ausgabe.pdbqt
4 center_x = x
5 center_y = y
6 center_z = z
7 size_x = 22.5
8 size_y = 22.5
9 size_z = 22.5

```

Redocking mit PLANTS

Skript für das Redocking mit PLANTS:

```

1 aco_ants 20
2 aco_evap 0.25
3 aco_sigma 1.0
4 flip_amide_bonds 0
5 flip_planar_n 1
6 cluster_structures 10
7 cluster_rmsd 2.0
8 protein_file Rezeptor.mol2
9 ligand_file Ligand.mol2
10 bindingsite_center x y z
11 bindingsite_radius 15
12 scoring_function chemplp/plp
13 ligand_intra_score clash
14 output_dir Rezeptor-Ligand-Ausgabe

```

Protein und *Ligand* bezeichnet das jeweilige Protein und den zu dockenden Liganden. Die Koordinaten (x y z) dienen der Definition der Bindestelle und bezeichnen das Zentrum dieser. In dieser Arbeit war das C α -Atom des ersten Histidins des Zink-Bindemotivs. Außerdem wird die jeweilige Scoring-Funktion festgelegt werden (*chemplp/plp*).

Redockingergebnisse

Tabelle A.7: Redockingergebnisse (numerisch) für die jeweiligen ADAMTS-Protein-Inhibitor-Komplexe. Für jeden Komplex ist der RMSD zwischen nativer und gedockter Pose angegeben sowie die Mittelwerte über alle Komplexe für das jeweilige Programm.

Protein	PDB	AutoDock4 RMSD (Å)	AutoDockVina RMSD (Å)	Plants (CHEMPLP) RMSD (Å)	PLANTS (PLP) RMSD (Å)	Schrödinger Glide RMSD (Å)
ADAMTS1	2JIH	2.50	3.31	2.03	2.09	0.55
	3Q2G	3.37	3.80	2.90	4.29	1.59
	3Q2H	9.66	4.55	6.47	6.38	3.59
ADAMTS4	2RJP	13.37	1.69	2.54	2.37	1.48
ADAMTS5	2RJQ	5.55	1.50	7.14	7.12	0.40
	3B8Z	9.21	1.24	0.70	2.08	0.39
	3HY7	3.86	2.46	1.71	1.12	0.90
	3HY9	5.00	1.60	1.39	2.03	1.63
	3HYG	6.20	2.41	1.65	0.46	2.20
	3LJT	1.42	5.01	2.03	2.20	0.25
Mittelwert		6.01	2.76	2.86	3.02	1.30

A.3.2 Virtual Screening

Tabelle A.8 (Fortsetzung): Strukturformelverzeichnis der besten 79 Verbindungen des ADAMTS7-Virtual Screenings. Es ist jeweils die ZINC-ID, die Strukturformel, der Trivialname und der Docking-Score für die beste Pose angegeben. Die Liste ist sortiert nach Docking-Score.

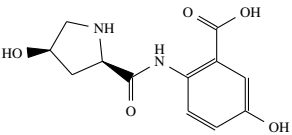
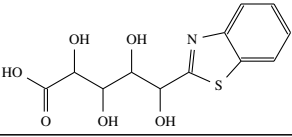
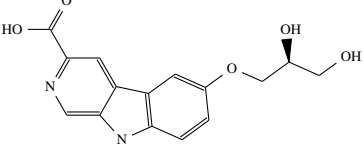
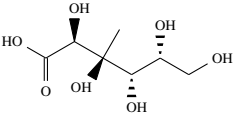
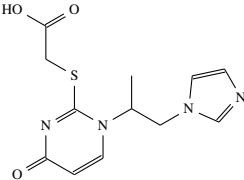
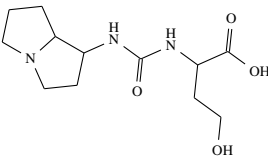
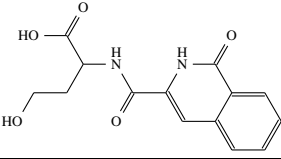
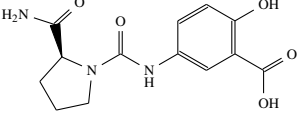
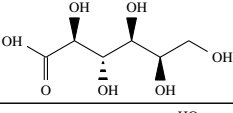
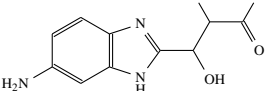
Nr.	Zinc-ID (CAS)	Struktur	Trivialname	Docking-Score
1	ZINC48847273 (-)		5-hydroxy-2-[(4-hydroxypyrrolidine-2-carbonyl)amino]benzoic acid	-10,67
2	ZINC03954014 (-)		5-benzothiazol-2-yl-2,3,4,5-tetrahydroxy-pentanoic acid	-10,43
3	ZINC62424080 (-)		6-[(2S)-2,3-dihydroxypropoxy]-9H-pyrido[3,4-b]indole-3-carboxylic acid	-10,30
4	ZINC45076865 (299-27-4)		2,3,4,5,6-pentahydroxy-3-methylhexanoate	-10,15
5	ZINC52329369 (-)		2-[1-(1-imidazol-1-ylpropan-2-yl)-4-oxopyrimidin-2-yl]sulfanylacetic acid	-10,13
6	ZINC49571688 (-)		2-(2,3,5,6,7,8-hexahydro-1H-pyrrolizin-1-ylcarbamoylamino)-4-hydroxybutanoic acid	-10,06
7	ZINC40697441 (-)		4-hydroxy-2-[(1-oxo-2H-isoquinoline-3-carbonyl)amino]butanoic acid	-9,89
8	ZINC40558343 (-)		5-[(2-carbamoylpyrrolidine-1-carbonyl)amino]-2-hydroxybenzoic acid	-9,78
9	ZINC03869786 (526-97-6)		(2S,3S,4S,5R)-2,3,4,5,6-pentahydroxyhexanoic acid	-9,74
10	ZINC06445419 (957314-21-5)		3-(5-amino-1H-benzimidazol-2-yl)-2,3-dihydroxypropanoic acid	-9,66

Tabelle A.8 (Fortsetzung): Strukturformelverzeichnis der besten 79 Verbindungen (5 %) des ADAMTS7-Virtual Screenings. Es ist jeweils die ZINC-ID, die Strukturformel, der Trivialname und der Docking-Score für die beste Pose angegeben. Die Liste ist sortiert nach Docking-Score.

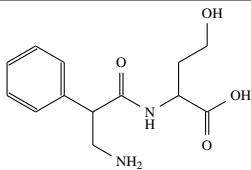
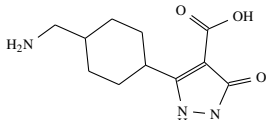
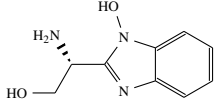
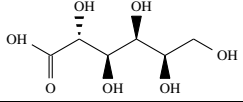
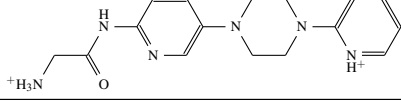
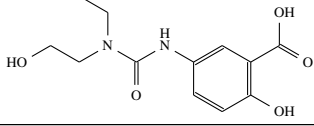
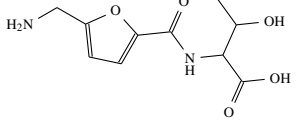
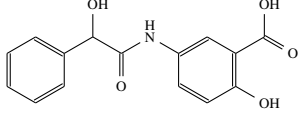
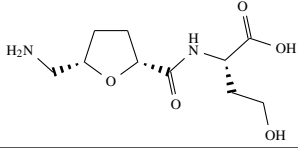
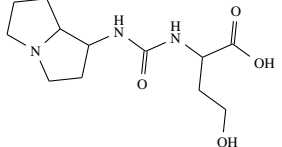
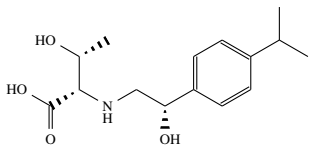
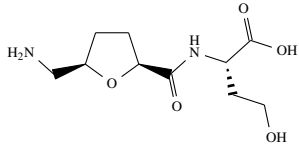
Nr.	Zinc-ID (CAS)	Struktur	Trivialname	Docking-Score
11	ZINC61612585 (-)		2-[(3-amino-2-phenylpropanoyl)amino]-4-hydroxybutanoic acid	-9,62
12	ZINC75953566 (-)		3-[4-(aminomethyl)cyclohexyl]-5-oxo-1,2-dihydropyrazole-4-carboxylic acid	-9,57
13	ZINC76136192 (-)		(2R)-2-amino-2-(1-hydroxybenzimidazol-2-yl)ethanol	-9,48
14	ZINC05131990 (20246-33-7)		(2R,3R,4S,5R)-2,3,4,5,6-penta-hydroxyhexanoic acid	-9,44
15	ZINC31961806 (-)		[2-oxo-2-[[5-(4-pyridin-1-ium-2-yl)piperazin-1-yl]pyridin-2-yl]amino]ethyl]azanium	-9,37
16	ZINC40603395 (-)		5-[[ethyl(2-hydroxyethyl)carbamoyl]amino]-2-hydroxybenzoic acid	-9,33
17	ZINC94569517 (-)		2-[[5-(aminomethyl)furan-2-carbonyl]amino]-3-hydroxy-butanoic acid	-9,33
18	ZINC37374912 (-)		2-hydroxy-5-[(2-hydroxy-2-phenylacetyl)amino]benzoic acid	-9,32
19	ZINC94569989 (-)		(2S)-2-[[[(2R,5S)-5-(aminomethyl)tetrahydrofuran-2-carbonyl]amino]-4-hydroxy-butanoic acid	-9,30
20	ZINC49571689 (-)		2-(2,3,5,6,7,8-hexahydro-1H-pyrrolizin-1-ylcarbamoylamino)-4-hydroxybutanoic acid	-9,29
21	ZINC71689288 (-)		(2S,3R)-3-hydroxy-2-[[[(2S)-2-hydroxy-2-(4-isopropylphenyl)ethyl]amino]butanoic acid	-9,28
22	ZINC94569992 (-)		(2S)-2-[[[(2S,5R)-5-(aminomethyl)tetrahydrofuran-2-carbonyl]amino]-4-hydroxy-butanoic acid	-9,22

Tabelle A.8 (Fortsetzung): Strukturformelverzeichnis der besten 79 Verbindungen (5 %) des ADAMTS7-Virtual Screenings. Es ist jeweils die ZINC-ID, die Strukturformel, der Trivialname und der Docking-Score für die beste Pose angegeben. Die Liste ist sortiert nach Docking-Score.

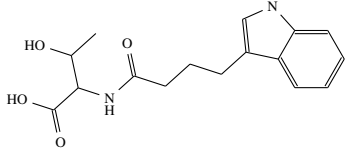
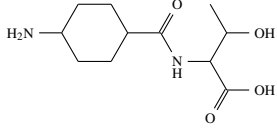
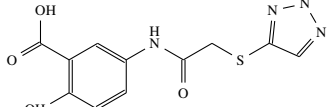
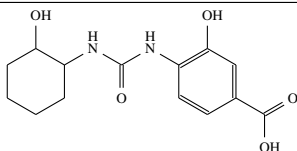
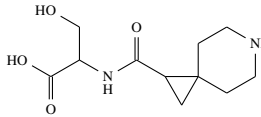
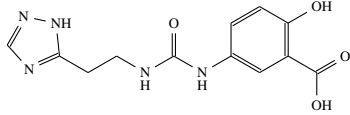
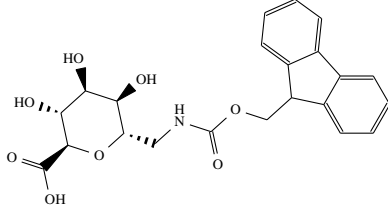
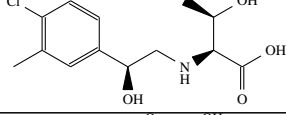
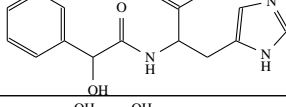
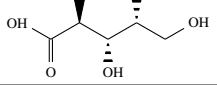
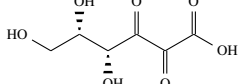
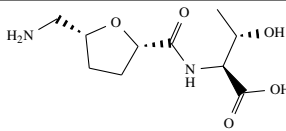
Nr.	Zinc-ID (CAS)	Struktur	Trivialname	Docking-Score
23	ZINC05426499 (-)		3-hydroxy-2-[4-(1H-indol-3-yl)-butanoylamino]butanoic acid	-9,19
24	ZINC40470996 (-)		2-[(4-aminocyclohexanecarbonyl)-amino]-3-hydroxybutanoic acid	-9,09
25	ZINC04912692 (-)		2-hydroxy-5-[2-(1H-1,2,3-triazol-5-ylsulfanyl)acetamido]benzoic acid	-9,03
26	ZINC52012042 (-)		3-hydroxy-4-[(2-hydroxycyclohexyl)-carbamoylamino]benzoic acid	-9,03
27	ZINC70254898 (-)		2-(6-azaspiro[2.5]octane-2-carbonylamino)-3-hydroxypropanoic acid	-8,98
28	ZINC80629149 (-)		2-hydroxy-5-[2-(1H-1,2,4-triazol-5-yl)ethylcarbamoylamino]benzoic acid	-8,93
29	ZINC71788171 (-)		(2R,3R,4S,5R,6S)-6-[(9H-fluoren-9-ylmethoxycarbonylamino)methyl]-3,4,5-trihydroxyoxane-2-carboxylic acid	-8,93
30	ZINC71692517 (-)		(2S,3R)-2-[[2-(2S)-2-(4-chloro-3-methyl-phenyl)-2-hydroxy-ethyl]amino]-3-hydroxy-butanoic acid	-8,92
31	ZINC37374946 (-)		2-[(2-hydroxy-2-phenylacetyl)amino]-3-(1H-imidazol-5-yl)propanoic acid	-8,91
32	ZINC00895290 (28223-40-7)		(2S,3S,4R)-2,3,4,5-tetrahydroxypentanoic acid	-8,91
33	ZINC30725382 (-)		(4R,5S)-4,5,6-trihydroxy-2,3-dioxo-hexanoate	-8,90
34	ZINC94569514 (-)		(2R,3S)-2-[[2-(2S,5R)-5-(aminomethyl)tetrahydrofuran-2-carbonyl]amino]-3-hydroxy-butanoic acid	-8,85

Tabelle A.8 (Fortsetzung): Strukturformelverzeichnis der besten 79 Verbindungen (5 %) des ADAMTS7-Virtual Screenings. Es ist jeweils die ZINC-ID, die Strukturformel, der Trivialname und der Docking-Score für die beste Pose angegeben. Die Liste ist sortiert nach Docking-Score.

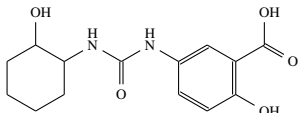
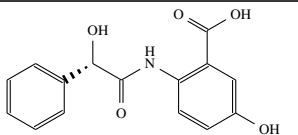
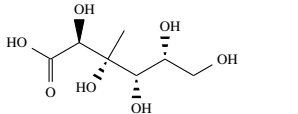
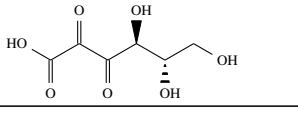
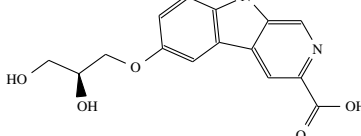
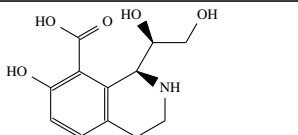
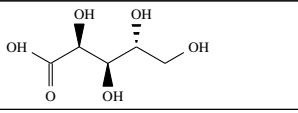
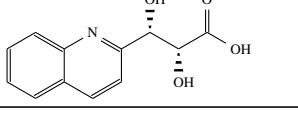
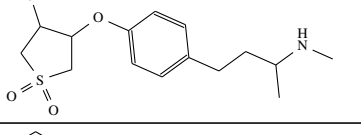
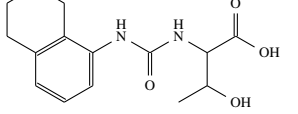
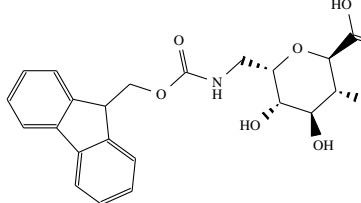
Nr.	Zinc-ID (CAS)	Struktur	Trivialname	Docking-Score
35	ZINC52011379 (-)		2-hydroxy-5-[(2-hydroxycyclohexyl)-carbamoylamino]benzoic acid	-8,85
36	ZINC48847183 (-)		5-hydroxy-2-[(2-hydroxy-2-phenylacetyl)amino]benzoic acid	-8,84
37	ZINC45076864 (299-27-4)		(2S,3R,4S,5R)-2,3,4,5,6-pentahydroxy-3-methylhexanoate	-8,82
38	ZINC01530152 (-)		(4R,5S)-4,5,6-trihydroxy-2,3-dioxohexanoic acid	-8,81
39	ZINC62424078 (-)		6-[(2R)-2,3-dihydroxypropoxy]-9H-pyrido[3,4-b]indole-3-carboxylic acid	-8,78
40	ZINC75821893 (-)		(1S)-1-[(1R)-1,2-dihydroxyethyl]-7-hydroxy-1,2,3,4-tetrahydroisoquinoline-8-carboxylic acid	-8,77
41	ZINC00895294 (488-30-2)		(2S,3R,4R)-2,3,4,5-tetrahydroxypentanoic acid	-8,77
42	ZINC62722318 (-)		(2R,3S)-2,3-dihydroxy-3-(2-quinolyl)propanoic acid	-8,76
43	ZINC37707318 (-)		4-[4-[3-(methylamino)butyl]phenoxy]-1,1-dioxothiolan-3-ol	-8,75
44	ZINC40536338 (-)		3-hydroxy-2-(5,6,7,8-tetrahydronaphthalen-1-ylcarbamoylamino)butanoic acid	-8,72
45	ZINC71788172 (-)		(2R,3R,4S,5S,6S)-6-[(9H-fluoren-9-ylmethoxycarbonylamino)methyl]-3,4,5-trihydroxy-tetrahydropyran-2-carboxylic acid	-8,72

Tabelle A.8 (Fortsetzung): Strukturformelverzeichnis der besten 79 Verbindungen (5 %) des ADAMTS7-Virtual Screenings. Es ist jeweils die ZINC-ID, die Strukturformel, der Trivialname und der Docking-Score für die beste Pose angegeben. Die Liste ist sortiert nach Docking-Score.

Nr.	Zinc-ID (CAS)	Struktur	Trivialname	Docking-Score
46	ZINC40478993 (-)		(2S)-5-amino-5-oxo-2-[[[(4S)-thiazolidine-4-carbonyl]amino]pentanoic acid	-8,72
47	ZINC76136191 (-)		(2S)-2-amino-2-(1-hydroxybenzimidazol-2-yl)ethanol	-8,71
48	ZINC42430101 (-)		2-[(4-aminopiperidine-1-carbonyl)amino]-3-hydroxypropanoic acid	-8,68
49	ZINC18141652 (6814-36-4)		(2S,3S,4S,5R)-2,3,4,5-tetrahydroxy-6-oxohexanoic acid	-8,68
50	ZINC37838771 (-)		5-[bis(2-hydroxyethyl)sulfamoyl]-2-hydroxybenzoic acid	-8,68
51	ZINC74459502 (-)		(3S,4S)-4-hydroxy-3-(2-hydroxyethylamino)-4-(2-oxo-3H-1,3-benzoxazol-6-yl)butanoic acid	-8,67
52	ZINC49587694 (-)		N-[[[3-(3-aminoprop-1-ynyl)phenyl]methyl]-1H-pyrazole-4-sulfonamide	-8,66
53	ZINC52012291 (-)		5-hydroxy-2-[(2-hydroxycyclohexyl)carbamoylamino]benzoic acid	-8,65
54	ZINC40699189 (-)		(2S)-2-(1-adamantylcarbamoylamino)-4-hydroxybutanoic acid	-8,64
55	ZINC52011629 (-)		2-hydroxy-4-[(2-hydroxycyclohexyl)carbamoylamino]benzoic acid	-8,62
56	ZINC52012220 (-)		4-hydroxy-1-[(2-hydroxycyclohexyl)carbamoyl]pyrrolidine-2-carboxylic acid	-8,61

Tabelle A.8 (Fortsetzung): Strukturformelverzeichnis der besten 79 Verbindungen (5 %) des ADAMTS7-Virtual Screenings. Es ist jeweils die ZINC-ID, die Strukturformel, der Trivialname und der Docking-Score für die beste Pose angegeben. Die Liste ist sortiert nach Docking-Score.

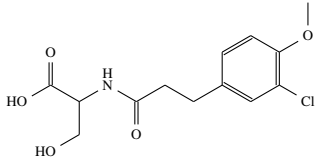
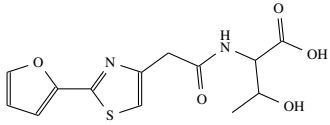
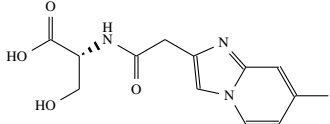
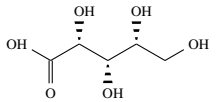
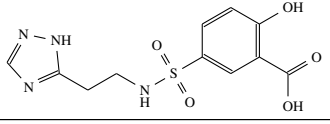
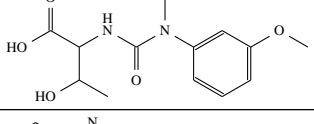
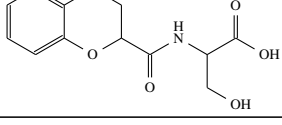
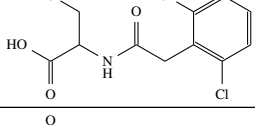
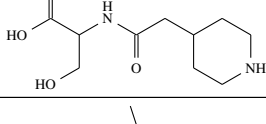
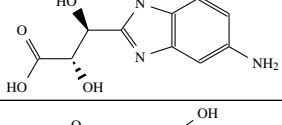
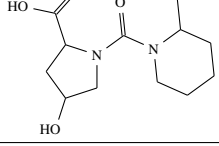
Nr.	Zinc-ID (CAS)	Struktur	Trivialname	Docking-Score
57	ZINC41638695 (-)		2-[3-(3-chloro-4-methoxyphenyl)propanoylamino]-3-hydroxypropanoic acid	-8,58
58	ZINC36971837 (-)		2-[[2-[2-(furan-2-yl)-1,3-thiazol-4-yl]acetyl]amino]-3-hydroxybutanoic acid	-8,55
59	ZINC71541252 (-)		(2R)-3-hydroxy-2-[[2-(7-methylimidazo[1,2-a]pyridin-2-yl)acetyl]amino]propanoic acid	-8,53
60	ZINC08551175 (526-91-0)		(2R,3S,4R)-2,3,4,5-tetrahydroxypentanoic acid	-8,52
61	ZINC80583404 (-)		2-hydroxy-5-[2-(1H-1,2,4-triazol-5-yl)ethylsulfamoyl]benzoic acid	-8,51
62	ZINC49974870 (-)		3-hydroxy-2-[[3-(3-methoxyphenyl)methylcarbamoyl]amino]butanoic acid	-8,51
63	ZINC40473067 (-)		2-(3,4-dihydro-2H-1,4-benzoxazine-2-carbonylamino)-3-hydroxypropanoic acid	-8,51
64	ZINC37414218 (-)		2-[[2-(2,6-dichlorophenyl)acetyl]amino]-3-hydroxypropanoic acid	-8,47
65	ZINC40473185 (-)		3-hydroxy-2-[(2-piperidin-4-ylacetyl)amino]propanoic acid	-8,46
66	ZINC59116063 (-)		(2S,3S)-3-(5-amino-1-methylbenzimidazol-2-yl)-2,3-dihydroxypropanoic acid	-8,45
67	ZINC48858643 (-)		4-hydroxy-1-[2-(hydroxymethyl)piperidine-1-carbonyl]pyrrolidine-2-carboxylic acid	-8,45

Tabelle A.8 (Fortsetzung): Strukturformelverzeichnis der besten 79 Verbindungen (5 %) des ADAMTS7-Virtual Screenings. Es ist jeweils die ZINC-ID, die Strukturformel, der Trivialname und der Docking-Score für die beste Pose angegeben. Die Liste ist sortiert nach Docking-Score.

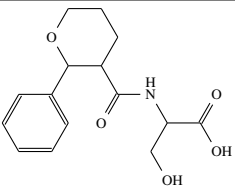
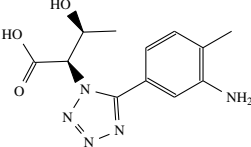
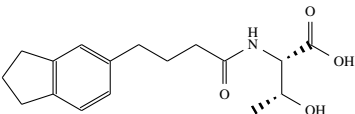
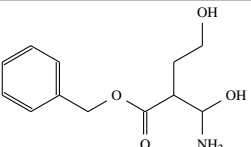
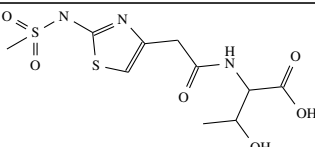
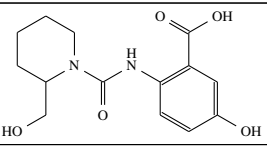
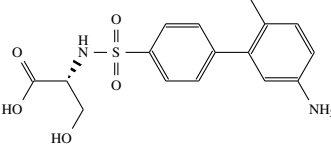
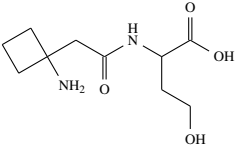
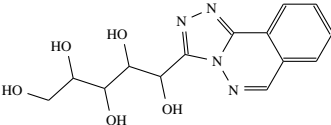
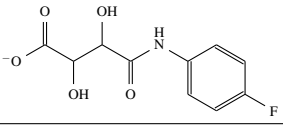
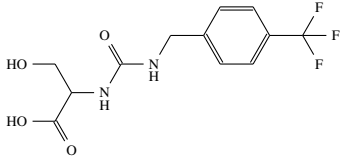
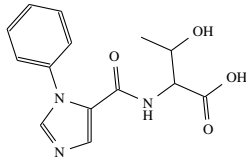
Nr.	Zinc-ID (CAS)	Struktur	Trivialname	Docking-Score
68	ZINC37921499 (-)		3-hydroxy-2-[(2-phenyloxane-3-carbonyl)amino]propanoic acid	-8,44
69	ZINC62468257 (-)		(2R,3S)-2-[5-(3-amino-4-methylphenyl)tetrazol-1-yl]-3-hydroxybutanoic acid	-8,43
70	ZINC62458532 (-)		(2S,3R)-3-hydroxy-2-(4-indan-5-ylbutanoylamino)butanoic acid	-8,42
71	ZINC71186797 (118219-23-1)		benzyl 2-[amino(hydroxy)methyl]-4-hydroxybutanoate	-8,41
72	ZINC37431718 (-)		3-hydroxy-2-[[2-[2-(methanesulfonamido)-1,3-thiazol-4-yl]acetyl]amino]butanoic acid	-8,41
73	ZINC48859135 (-)		5-hydroxy-2-[[2-(hydroxymethyl)piperidine-1-carbonyl]amino]benzoic acid	-8,41
74	ZINC71710083 (-)		(2R)-2-[[4-(5-amino-2-methylphenyl)phenyl]sulfonylamino]-3-hydroxypropanoic acid	-8,40
75	ZINC86853423 (-)		(2S)-2-[[2-(1-aminocyclobutyl)acetyl]amino]-4-hydroxybutanoic acid	-8,40
76	ZINC35644432 (79364-50-4)		1-([1,2,4]triazolo[3,4-a]phthalazin-3-yl)pentane-1,2,3,4,5-pentol	-8,39
77	ZINC02545176 (206761-65-1)		4-(4-fluoroanilino)-2,3-dihydroxy-4-oxobutanoic acid	-8,38

Tabelle A.8 (Fortsetzung): Strukturformelverzeichnis der besten 79 Verbindungen (5 %) des ADAMTS7-Virtual Screenings. Es ist jeweils die ZINC-ID, die Strukturformel, der Trivialname und der Docking-Score für die beste Pose angegeben. Die Liste ist sortiert nach Docking-Score.

Nr.	Zinc-ID (CAS)	Struktur	Trivialname	Docking-Score
78	ZINC61627384 (-)		3-hydroxy-2-[[4-(trifluoromethyl)phenyl]methylcarbamoylamino]propanoic acid	-8,37
79	ZINC37431585 (-)		3-hydroxy-2-[(3-phenylimidazole-4-carbonyl)amino]butanoic acid	-8,37

A.4 Ergebnisse: Weiterführende Untersuchungen der potenziellen ADAMTS7-Inhibitoren

A.4.1 Inverses Screening

Tabelle A.9: Normalisierte Docking-Scores der 79 Verbindungen aus dem inversen Screening. Die normalisierten Docking-Scores für ADAMTS7 resultieren aus dem Virtual Screening und sind fett markiert. Docking-Posen sind mit ** markiert, wenn die Zink-Koordinierung und die Wasserstoffbrücke zum Gly358-O-Analogon vorlag, mit *, wenn nur die Zink-Koordinierung vorlag und ohne weitere Markierung, wenn keines der beiden Kriterien zutrif.

ADAMTS7 (VS)	-0,12**	0,12**	-0,39**	0,75**	-1,50**	-2,63**	-0,32**	-0,99**	-0,59**	-1,35**	0,52**	0,25**	-0,46**	-2,04**
ADAMTS1	-2,25*	-2,91**	-2,16*	-3,95**	-3,76**	-7,93**	-2,81*	-3,73**	-6,34*	-3,57*	-2,88**	-3,80**	-1,79**	0,78
ADAMTS2	1,55**	1,81**	1,03*	0,68*	0,39**	-0,87*	2,34*	1,65**	-2,46**	1,32*	1,51**	-0,09**	1,42**	-0,80*
ADAMTS3	0,93*	0,71*	1,31*	0,03*	0,28*	-1,88**	1,96**	0,92*	-2,16*	0,71*	0,93**	0,76**	0,95**	0,04
ADAMTS4	0,55**	0,77**	0,29**	-1,12*	-1,59**	-5,42**	0,37**	-2,06**	-3,36*	-0,82*	-1,61**	0,11**	0,44**	-0,82
ADAMTS5	0,94*	-0,84**	-0,60*	-2,48**	-1,71**	-2,80*	-1,91**	-3,50*	0,54*	-1,66**	-2,14**	-2,27**	0,01*	-0,15
ADAMTS6	0,06**	-0,72**	0,30**	0,16*	-0,02**	-3,24**	-0,44**	-0,36**	-2,72**	0,65**	0,15**	-0,12**	0,80**	-0,71**
ADAMTS8	1,34*	1,00**	0,39*	-0,18*	-1,65**	-3,26**	0,69*	-1,75**	-1,37*	-1,75**	-1,53**	-1,69**	0,82*	-1,08
ADAMTS9	-0,59**	0,04**	0,60*	2,25*	-1,43**	-0,96**	2,86*	-2,57**	-1,04*	1,22*	0,64*	1,10*	0,52*	1,03
ADAMTS10	-0,89**	-0,70**	-0,54**	3,60**	-1,67**	-2,79**	2,48**	-3,40**	1,03**	3,97	0,87*	0,63*	0,31**	-0,95
ADAMTS12	-1,28**	-0,85**	-1,24**	2,28*	-2,91**	-1,72**	2,26**	-2,79**	-0,34*	0,93**	-0,77**	-1,87**	-0,81**	1,73**
ADAMTS13	1,28**	1,02*	1,34*	1,44*	0,27*	-1,15**	1,30*	-0,62**	0,69*	1,64*	1,10**	0,89**	0,83**	0,03**
ADAMTS14	-0,59*	-0,31**	0,99**	2,77*	-0,86**	-0,60**	1,10*	-0,53*	0,07*	1,14*	0,17**	0,69*	0,02**	-0,66
ADAMTS15	0,01**	-0,35**	-0,53**	0,68*	-1,36**	-4,15**	1,33**	-0,78*	-1,45*	-0,62**	0,37**	-0,02**	0,61*	-0,09
ADAMTS16	-1,06**	-1,32**	-1,28**	1,59*	-2,44**	-2,03**	1,56*	-2,47**	2,20*	0,72*	-0,04**	0,69**	-1,95**	1,22
ADAMTS17	0,63**	2,42**	0,26**	4,99*	-2,28**	1,54*	3,80*	-2,04**	-0,08**	2,88*	1,01**	-0,78**	-0,02**	0,71
ADAMTS18	0,26**	-0,55**	-1,22**	-0,15*	-2,85*	4,38**	-0,25*	-4,97**	-2,38*	1,35*	-2,07**	-1,88*	-1,12**	1,60*
ADAMTS19	-0,36**	1,11**	1,07*	6,35	-0,27*	-0,50**	4,65*	-2,57**	2,52**	5,98*	0,53**	1,90**	-1,49**	-0,24
ADAMTS20	-0,41**	0,36*	0,40*	3,11**	-2,30**	-1,37**	6,61*	-1,60*	0,53*	-0,35**	-0,24*	0,19*	0,90*	0,40
ZINC31961806														
ZINC30725382														
ZINC18141652														
ZINC08551175														
ZINC06445419														
ZINC05426499														
ZINC05131990														
ZINC04912692														
ZINC03954014														
ZINC03869786														
ZINC02545176														
ZINC01530152														
ZINC00895294														
ZINC00895290														

Tabelle A.9 (Fortsetzung): Normalisierte Docking-Scores der 79 Verbindungen aus dem inversen Screening. Die normalisierten Docking-Scores für ADAMTS7 resultieren aus dem Virtual Screening und sind fett markiert. Docking-Posen sind mit ** markiert, wenn die Zink-Koordinierung und die Wasserstoffbrücke zum Gly358-O-Analogon vorlag, mit *, wenn nur die Zink-Koordinierung vorlag und ohne weitere Markierung, wenn keines der beiden Kriterien zutrif.

ZINC35644432	ZINC36971837	ZINC37374912	ZINC37374946	ZINC37414218	ZINC37431585	ZINC37431718	ZINC37707318	ZINC37838771	ZINC37921499	ZINC40470996	ZINC40473067	ZINC40473185	ZINC40478993	ZINC40536338	ZINC40558343	ZINC40603395	ZINC40697441	ZINC40699189		
ADAMTS7 (VS)	0,73**	-0,09**	-0,65**	0,11**	-0,42**	0,07**	-0,30**	-2,00**	-1,27	-1,26*	0,06**	-1,15**	-0,05**	-0,17**	-0,55**	-0,59**	-1,78**	-0,97**	-1,24**	-0,66**
ADAMTS1	-0,82**	-2,33*	-0,11**	-1,05*	-2,04*	-0,15*	-2,72**	-1,27	-1,26*	-2,65*	-2,41*	-2,16*	-1,76**	-1,67**	1,13	-2,47*	-2,12**	-2,83**		
ADAMTS2	2,14**	0,11**	-0,51*	0,86*	0,26*	1,69*	0,93**	-0,16**	-0,39**	1,24*	0,51**	-0,86*	0,53**	0,48**	0,20*	-0,23**	-0,44**	-0,12**	-1,10**	
ADAMTS3	1,45	0,24*	-0,05*	-1,03*	-0,01*	-0,44*	0,65*	0,75**	1,01*	0,43*	-0,30*	-0,40*	0,54*	0,55**	0,67*	-0,45*	0,41**	-0,16*	-0,34*	
ADAMTS4	-1,01**	-0,68**	-0,98*	-0,56*	-0,96*	0,88*	-0,98**	-1,11	-0,84*	-0,79*	-0,29**	-0,80*	0,11*	-0,56**	-0,01	-1,66**	-1,01*	-1,83**	-0,52*	
ADAMTS5	-3,99	-1,40*	0,49	0,02*	-0,85*	-2,35*	-0,73*	-0,90	-0,60*	-1,74*	0,19*	-1,32*	-2,18*	-0,45*	-1,78*	-0,95*	-1,32*	-0,60*	0,65	
ADAMTS6	-0,05**	0,24*	-0,42*	-1,29*	-0,23**	0,09**	0,66**	-0,15*	-0,94*	-0,06*	0,04**	-0,57**	0,37**	0,66*	-0,40*	-0,80*	-0,25*	-0,70*	-0,78*	
ADAMTS8	0,70**	-0,55*	-0,73*	-0,99*	-0,65*	0,54*	0,50*	0,03*	-0,95**	0,12*	0,09**	-0,25*	-0,20*	1,10*	-0,29*	-0,51*	-0,37*	-0,13*	0,85*	
ADAMTS9	2,04**	0,89*	0,38*	-0,14*	1,34*	1,14*	1,09*	-0,17*	2,83*	1,40*	0,31**	1,41*	-1,09**	-1,37*	0,48*	1,80*	0,40*	1,37*	1,62**	
ADAMTS10	3,19**	-0,06**	1,18	2,44	2,24	-0,24**	-0,08**	0,07	1,02*	0,95**	2,84*	0,26**	0,40**	1,29*	1,07	0,92**	1,69	1,48	0,59**	
ADAMTS12	2,69	1,82**	0,25**	1,13**	0,30**	0,24*	0,18*	-1,47**	-0,84**	0,25**	0,24**	0,82**	1,35**	0,96**	1,16**	0,75*	0,39**	0,42*	0,02**	
ADAMTS13	3,09**	-0,27*	0,30*	0,78**	0,31*	0,39*	0,32**	-0,14*	0,23*	0,51*	0,01*	0,14**	-1,04*	0,22**	0,69*	-0,45*	0,16*	-0,14*	0,06*	
ADAMTS14	3,47**	1,14*	0,58*	0,69*	0,99*	1,18*	1,58*	1,39*	-0,01*	1,08*	-0,96*	0,54*	0,80*	-1,09*	1,57*	0,26*	0,05*	0,54*	0,89*	
ADAMTS15	0,12**	-0,19*	-0,75**	-0,77*	-0,69**	0,20**	0,87**	0,04**	-0,89*	-0,53**	-0,32**	-0,33**	0,40*	0,62*	-0,12*	1,10*	-0,38*	-0,32*	0,00*	
ADAMTS16	3,39**	1,41*	0,39*	-0,54**	0,26**	-0,51*	0,57**	0,72	0,06**	0,24*	0,72**	0,67*	-0,85*	-1,81**	0,25*	0,40*	0,50**	0,59*	0,17*	
ADAMTS17	3,83**	-0,80**	1,75*	-0,85**	0,58**	-0,37*	-1,24**	0,84	0,72*	0,37**	-0,06**	1,89*	1,76**	0,38**	-0,39**	0,97*	1,49**	1,28**	-0,04**	
ADAMTS18	4,16**	-0,74*	-0,67*	-0,49**	-1,39*	-1,73*	-0,26*	1,57*	0,78*	-0,54*	0,23*	-0,42*	0,09*	-0,76**	0,09*	-1,06*	0,54*	0,75*	0,65*	
ADAMTS19	1,51**	0,28*	1,82*	1,35*	0,80**	-1,25**	-0,60**	0,90	0,69*	-0,62**	0,67**	1,53**	0,53**	0,73**	0,59**	0,62*	1,19*	0,62**	-0,53**	
ADAMTS20	3,71**	0,97*	0,02*	0,31*	0,14*	0,62**	-0,44**	1,03	-0,27**	0,27**	-0,37*	-0,10*	0,42*	1,27*	-0,51**	-0,05*	0,38*	0,31**	1,29**	

Tabelle A.9 (Fortsetzung): Normalisierte Docking-Scores der 79 Verbindungen aus dem inversen Screening. Die normalisierten Docking-Scores für ADAMTS7 resultieren aus dem Virtual Screening und sind fett markiert. Docking-Posen sind mit ** markiert, wenn die Zink-Koordinierung und die Wasserstoffbrücke zum Gly358-O-Analogon vorlag, mit *, wenn nur die Zink-Koordinierung vorlag und ohne weitere Markierung, wenn keines der beiden Kriterien zutraf.

ADAMTS7 (VS)	0,06**	-1,06**	0,02**	-2,17**	-0,70**	-1,37**	-0,90**	-0,83**	-1,64**	-1,19**	-2,22**	-0,51**	-0,15**	-1,03**	-0,67**	-0,73**	-1,66**	0,64**
ADAMTS1	-2,63*	-2,49*	-2,32**	-3,68**	-2,60*	-1,12*	-2,06*	-2,13*	-2,21*	-2,16**	-0,28	-2,03**	-1,94*	0,27**	-2,38*	-2,92*	-1,31*	-2,68*
ADAMTS2	0,49**	0,33**	0,88**	0,11**	-0,20*	1,07*	0,19*	0,60	0,38*	-0,72**	0,85	-0,05*	-0,32	-0,59*	-0,11*	-0,64**	0,19	0,93*
ADAMTS3	0,39*	-0,52*	0,29*	-0,01**	-0,46**	0,35**	-0,08*	-0,39**	-0,10*	0,28*	-0,63	-0,19*	0,15**	0,19*	0,10*	0,16*	-0,13**	0,13*
ADAMTS4	-0,99*	-1,17**	-0,20**	-2,39**	0,47	-0,72**	-0,26**	0,24	-0,40**	-0,71**	-1,22	-0,81*	-0,99**	-1,06**	-1,43*	-1,22**	-0,66**	0,03*
ADAMTS5	-0,88*	-0,55*	-2,70**	-3,32*	-0,04	-0,41*	-1,15*	0,00	1,63*	-0,96*	0,30	-0,79*	0,30	-1,38*	-1,24*	-0,02*	0,41	-0,17*
ADAMTS6	-0,52**	0,24**	-1,18**	-0,60**	-0,92*	-1,67*	-0,32*	-1,37**	-0,39**	-0,36**	-0,62**	-0,46**	-0,58*	-0,26*	-1,50*	-0,53**	-0,68**	0,29**
ADAMTS8	-0,22**	-0,07**	-0,57*	-0,70*	-0,22*	-0,76**	-0,11*	0,60*	-0,24*	0,11**	-0,33	-0,64*	-0,59*	-0,43*	-0,49*	-0,14*	0,23**	-0,67*
ADAMTS9	0,23*	1,14*	-0,96**	-2,00**	1,40	1,29*	1,35*	1,17*	1,08*	1,68*	-0,45	1,39*	0,71*	0,40*	1,03**	0,51*	0,97*	1,09*
ADAMTS10	0,84**	1,58*	-0,55*	-1,57**	1,05	1,50	1,08**	1,98	1,68	0,76**	0,04	2,35*	1,94	1,32**	0,10*	2,31	1,44	2,22
ADAMTS12	1,66**	0,27**	-0,40**	-1,57**	2,18	0,65**	-1,05**	0,47*	-0,49**	1,56*	-0,67	1,09*	0,50**	0,50*	0,17*	-0,03**	0,41*	0,89**
ADAMTS13	-0,12*	-1,13*	-1,00**	-0,14**	0,30*	0,34*	0,67*	0,68*	0,61*	0,67*	0,27*	-0,28*	0,32*	0,19*	0,32*	0,25**	-0,18*	-1,14*
ADAMTS14	0,87**	1,12*	0,22*	0,25*	0,14*	-0,84*	0,19**	0,39*	0,82*	0,59*	1,31	0,40*	0,35*	-0,05*	2,05*	0,53*	0,46*	0,81*
ADAMTS15	-0,97*	-0,47*	-1,26*	-0,27*	-0,33*	1,42*	-0,36*	-0,23*	-0,37*	-0,37*	0,13*	-0,79*	-0,90**	-0,79*	-0,90**	-1,02**	-0,61*	-0,92*
ADAMTS16	0,67*	0,66**	-2,88**	-2,47**	-0,60**	-1,24**	-0,49*	-1,49**	0,10*	1,08*	0,78	0,41*	0,93*	0,39*	0,17*	0,07**	0,02*	0,18**
ADAMTS17	-0,62**	0,71**	0,79**	-2,45**	0,38**	0,82**	1,72*	0,57*	-0,41**	0,26**	1,33	0,65*	0,71*	2,00	1,81*	1,07*	1,95	-0,01*
ADAMTS18	-0,15*	0,76**	-3,81**	-2,76**	0,79**	0,22**	-0,01*	-0,37*	-0,59*	-0,87*	2,16*	-0,80*	-0,80*	-0,77*	-0,64*	0,12*	0,07*	-0,09*
ADAMTS19	1,55**	0,65**	-0,92**	-2,06**	-0,16**	0,20*	1,85*	0,80*	-0,49**	-0,50**	-0,32	0,96**	1,04*	1,07*	0,66*	1,16**	0,15**	-1,33*
ADAMTS20	0,35**	-0,03*	-1,12*	-1,96*	-0,49*	0,27*	-0,29*	-0,67*	1,04*	0,85**	-0,43	0,10**	0,29*	1,36	0,65*	0,46*	-0,39*	0,74*
ZINC41638695																		
ZINC42430101																		
ZINC45076864																		
ZINC45076865																		
ZINC48847183																		
ZINC48847273																		
ZINC48858643																		
ZINC48859135																		
ZINC49571688																		
ZINC49571689																		
ZINC49587694																		
ZINC49974870																		
ZINC52011379																		
ZINC52011629																		
ZINC52012042																		
ZINC52012220																		
ZINC52012291																		
ZINC52329369																		
ZINC59116063																		

Tabelle A.9 (Fortsetzung): Normalisierte Docking-Scores der 79 Verbindungen aus dem inversen Screening. Die normalisierten Docking-Scores für ADAMTS7 resultieren aus dem Virtual Screening und sind fett markiert. Docking-Posen sind mit ** markiert, wenn die Zink-Koordinierung und die Wasserstoffbrücke zum Gly358-O-Analogon vorlag, mit *, wenn nur die Zink-Koordinierung vorlag und ohne weitere Markierung, wenn keines der beiden Kriterien zutrif.

ADAMTS7 (VS)	-0,60**	0,29**	-0,23**	-0,82**	0,41**	-0,43**	-0,07**	-1,13**	0,70**	0,03**	0,11**	0,00**	0,42**	-0,15**	0,19**	0,30**	0,10**	-1,42**	-0,40**
ADAMTS1	-2,59*	-2,60*	-2,93**	-2,34**	-2,43**	-2,77*	-2,20*	-2,59*	-3,15**	-2,16**	-2,05*	-2,64**	-1,90*	-2,79**	-6,44**	-1,96*	-0,61*	-1,64*	1,58
ADAMTS2	0,73**	0,72*	-0,07*	0,15*	0,04*	0,57**	0,77*	0,54**	0,18**	0,88**	0,76**	0,24*	-0,31**	-1,18**	0,55*	1,04*	1,40**	0,05**	2,60
ADAMTS3	-0,09**	0,47*	0,36*	0,42*	-0,04*	-0,42*	0,53*	0,14*	-3,96*	0,68*	-0,14*	-0,22*	0,04*	-2,28**	-1,70*	0,10*	-0,01**	0,37**	-0,54*
ADAMTS4	-0,30**	-0,49*	-0,86*	-1,27*	-1,52*	-1,06*	-0,40*	-0,87*	-3,43**	1,47**	-0,90**	-0,89**	-1,07**	-6,18*	-6,63*	-0,31**	-0,77**	-0,51*	-0,53**
ADAMTS5	0,18**	-1,88*	-0,69	-0,28	-1,64*	-0,88**	-1,44**	-1,49*	-9,67**	-0,63*	-0,11*	0,47*	-1,56*	0,75	3,38	-2,27*	0,40	0,69**	0,02
ADAMTS6	1,01**	0,38*	-1,30*	-1,54*	-0,38*	-0,37**	-0,19**	-0,44*	-5,87**	0,14**	-1,71**	-1,29**	-0,56*	-4,07*	-2,56*	-1,41*	-0,86**	-0,54**	0,08*
ADAMTS8	0,53**	0,32*	-0,34*	-0,04*	-0,25*	-0,02*	-1,02**	-0,03*	0,18*	-1,14*	0,16*	-0,11*	-1,31*	-1,87*	-2,66*	0,12*	0,20*	0,02**	-0,68**
ADAMTS9	1,95	0,26**	0,81*	0,50*	0,82*	0,53**	0,40**	0,59**	-1,40**	0,92*	1,09*	0,74*	0,71*	2,86*	1,43*	0,53*	1,56	0,66*	1,44*
ADAMTS10	-0,55**	-0,19**	0,75*	0,82*	0,18**	1,23**	0,19**	-0,04**	0,97*	1,15**	0,58**	1,32	1,50	2,60	3,29	0,77**	1,68	0,10**	-0,52**
ADAMTS12	0,51**	-0,25**	0,68*	1,04*	1,37**	1,13**	0,47	-0,16	-1,53*	1,39*	0,89**	0,63**	0,48*	1,76*	1,08*	0,46**	-0,26**	-0,20**	-0,84**
ADAMTS13	1,32*	-0,03*	-0,36*	-0,12*	-0,41*	0,48*	-0,40**	0,08*	-4,62**	-0,19*	-0,08*	0,15*	-0,74*	-0,91*	-1,77*	1,01*	0,45*	-0,45**	-0,69**
ADAMTS14	-0,99*	1,54*	1,03*	1,28*	1,21*	0,60*	0,41**	1,38**	-2,55*	-0,21**	-0,63**	-0,09*	0,92*	3,76**	1,10*	0,53**	0,17*	-0,01*	-0,70*
ADAMTS15	-0,09**	-0,53*	-0,44*	-0,50*	-0,36**	-0,92*	-0,17**	0,14*	-2,41**	-0,11**	0,07**	-0,18**	-0,35*	-1,35**	0,44*	-0,27**	-1,18**	-0,52**	-0,36*
ADAMTS16	-0,66**	1,16*	1,06**	-0,53*	0,67*	0,30*	1,00**	0,79*	-2,71**	-1,83*	-1,43**	-1,38**	0,72*	-2,61**	-4,08**	-1,18**	-1,99**	-1,19**	-1,17**
ADAMTS17	0,05**	1,39**	1,10*	0,93*	0,64**	0,64*	1,95**	1,34**	2,47**	0,09**	0,57**	1,35*	0,80**	0,29**	2,46*	0,86*	0,06**	1,39**	0,88**
ADAMTS18	-1,04**	0,22*	0,79*	0,57*	-0,01*	-0,56*	-0,73*	0,22*	-5,18**	-0,39*	0,21**	0,02**	-0,14*	-4,67**	-3,70**	0,90**	-1,52**	2,77**	-0,91*
ADAMTS19	0,00**	-0,36**	0,36	1,48*	0,97**	1,44*	1,56	1,32**	2,92**	0,46*	0,61**	0,79**	1,40**	-0,99**	2,81*	0,41**	0,94*	0,38**	0,65**
ADAMTS20	0,62**	-0,42**	0,29*	0,23*	0,73**	0,51*	-0,66**	0,21**	-2,56*	-0,54**	1,98	1,06*	0,94	0,31*	1,04*	0,38*	0,22*	0,05*	0,09*
ZINC61612585																			
ZINC61627384																			
ZINC62424078																			
ZINC62424080																			
ZINC62458532																			
ZINC62468257																			
ZINC62722318																			
ZINC70254898																			
ZINC71186797																			
ZINC71541252																			
ZINC71689288																			
ZINC71692517																			
ZINC71710083																			
ZINC71788171																			
ZINC71788172																			
ZINC74459502																			
ZINC75821893																			
ZINC75953566																			
ZINC76136191																			

Tabelle A.9 (Fortsetzung): Normalisierte Docking-Scores der 79 Verbindungen aus dem inversen Screening. Die normalisierten Docking-Scores für ADAMTS7 resultieren aus dem Virtual Screening und sind fett markiert. Docking-Posen sind mit ** markiert, wenn die Zink-Koordinierung und die Wasserstoffbrücke zum Gly358-O-Analogon vorlag, mit *, wenn nur die Zink-Koordinierung vorlag und ohne weitere Markierung, wenn keines der beiden Kriterien zutrif.

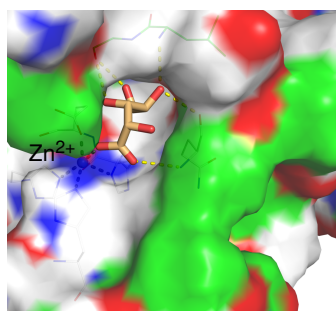
ZINC76136192	ADAMTS7 (VS)	-0,65**	-0,75**	-0,72**	0,72**	-0,03**	-0,81**	-0,77**	-0,63**	ZINC94569992
ZINC80583404	ADAMTS1	1,21	-2,18*	-2,03*	-2,33*	-4,06**	0,88*	-2,17**	-1,70*	ZINC94569989
ZINC80629149	ADAMTS2	1,88**	0,40**	0,71**	1,42**	2,25**	2,01**	27,35**	1,99**	ZINC94569517
ZINC86853423	ADAMTS3	-0,97**	1,19*	0,37*	-0,77*	1,68*	-0,09*	-0,45*	0,82*	ZINC94569514
ZINC94569514	ADAMTS4	-0,45**	-0,04*	-0,57*	0,80**	0,81**	1,89**	0,61**	0,33*	ZINC86853423
ZINC94569514	ADAMTS5	1,27	-0,58*	-1,36**	1,65*	-0,57**	1,63*	-0,38*	-0,17*	ZINC80629149
ZINC94569514	ADAMTS6	0,97*	0,10**	-0,20*	2,05*	-1,58*	1,54*	1,62**	0,85*	ZINC80583404
ZINC94569514	ADAMTS8	-0,70**	0,05*	0,16*	0,92*	-0,72**	1,66*	-2,33*	0,99**	ZINC76136192
ZINC94569514	ADAMTS9	1,16*	0,89*	0,96*	-1,27*	2,25*	2,22*	1,07**	-0,26**	ZINC80583404
ZINC94569514	ADAMTS10	-0,87**	1,04*	0,74**	2,83**	4,87*	7,60	6,37	5,37	ZINC80629149
ZINC94569514	ADAMTS12	-0,47**	-0,07*	0,21**	2,64**	2,14**	3,68**	2,17*	3,38*	ZINC86853423
ZINC94569514	ADAMTS13	-0,52*	0,28*	-0,78**	0,16*	-0,28*	0,89*	-1,25*	0,12*	ZINC94569514
ZINC94569514	ADAMTS14	-0,22*	-1,64*	0,08*	0,15*	-0,21*	3,00*	-0,88*	3,41*	ZINC80583404
ZINC94569514	ADAMTS15	-1,28**	-0,25*	0,56*	0,98**	1,86*	0,91*	0,99**	3,50*	ZINC80629149
ZINC94569514	ADAMTS16	-0,65**	0,70*	0,50*	-2,77*	2,88**	-0,14*	-0,83**	3,47*	ZINC86853423
ZINC94569514	ADAMTS17	-0,79**	1,23*	1,27*	1,57**	-0,50**	2,13**	1,05**	0,90**	ZINC94569514
ZINC94569514	ADAMTS18	-0,77*	-1,55*	-1,85*	1,72**	-1,46*	2,05**	-0,49*	0,93*	ZINC80583404
ZINC94569514	ADAMTS19	1,53	1,28*	1,61	1,11**	0,36*	2,00**	4,80**	3,88**	ZINC80629149
ZINC94569514	ADAMTS20	0,31**	-0,12**	0,35*	1,88*	2,44*	1,71**	1,82*	3,52*	ZINC94569514

A.4.2 Wechselwirkungen zwischen ADAMTS7-Bindetasche und potenziell (nicht) selektiven Inhibitoren

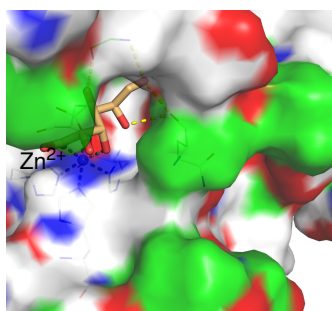
Analyse des Zusammenhangs zwischen Länge, Volumen und Selektivität von potenziellen ADAMTS7-Inhibitoren

Folgend ist das genutzte R-Skript (Paket e1071 [179]) zur Klassifizierung bzw. Bestimmung einer Klassifikationsgeraden aufgeführt. Durch dieser Klassifizierung sollen Liganden als selektive oder nicht selektive ADAMTS7-Inhibitoren anhand von Länge und Volumen in der Docking-Pose gruppiert werden. Die Trainingsdaten basieren auf den 18 als selektiv und 7 als nicht selektiv charakterisierten ADAMTS7-Inhibitoren:

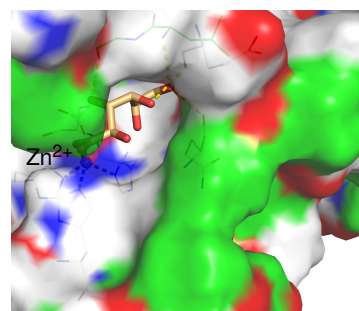
```
1 Laenge <- c(15.0 12.3 9.4 11.9 10.5 8.6 12.0 7.7 9.1 10.2 10.4 9.8 9.5 10.7 12.1 11.5 10.3
2 10.1 5.2 7.8 7.1 5.4 6.8 6.5 6.6)
3 Volumen <- c(242 245 184 221 228 188 221 178 181 220 204 221 197 213 223 222 181 188 115 140 143
4 114 141 156 191)
5 Gruppe <- c(1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 -1 -1 -1 -1 -1 -1)
6 my.data <- data.frame(x1=Laenge, x2=Volumen, type=as.factor(Gruppe))
7
8 library(e1071)
9 svm.model <- svm(type ~ ., data=my.data, type='C-classification', kernel='linear', scale=FALSE)
10
11 # Parameter der Hyperebene
12 w <- t(svm.model$coefs) %*% svm.model$SV
13 b <- -svm.model$rho
14 # in diesem Fall ist die Hyperebene die Gerade  $w[1,1]*x1 + w[1,2]*x2 + b = 0$ 
```



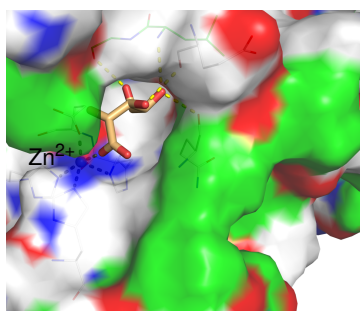
(a) ZINC00895294
Docking-Score: -8,77



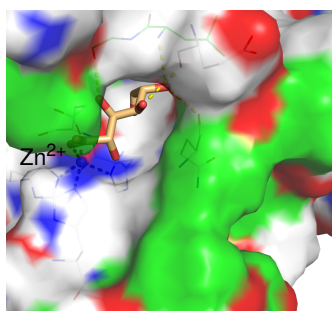
(b) ZINC03869786
Docking-Score: -9,74



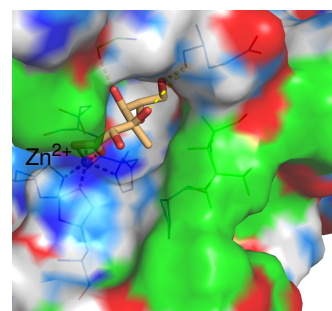
(c) ZINC05131990
Docking-Score: -9,44



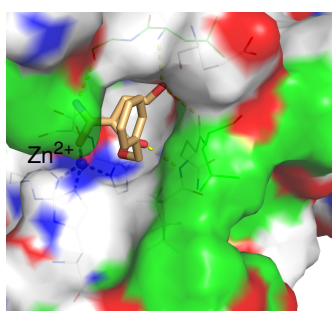
(d) ZINC08551175
Docking-Score: -8,52



(e) ZINC18141652
Docking-Score: -8,68



(f) ZINC45076864
Docking-Score: -8,82



(g) ZINC71186797
Docking-Score: -8,41

Abbildung A.7: Docking-Posen der sieben potenziell nicht selektiven Hits in ADAMTS7 mit Bindetasche in Oberflächendarstellung (grün: hydrophober Bereich; rot: Akzeptorbereich; blau: Donorbereich). Die Bilder sind analog zu denen in Abbildung 6.5 um 15° nach oben gedreht. Die Zink-koordinierenden Histidine sind als Linien dargestellt, genau wie die jeweils mit dem Liganden wechselwirkenden Aminosäuren. Die Liganden werden durch orange Stäbchen dargestellt und die Zink-Ionen sind als graue Kugeln zu erkennen. Zink-koordinierende Bindungen sind durch grau gestrichelte Linien dargestellt, Wasserstoffbrücken durch gelb gestrichelte Linien.

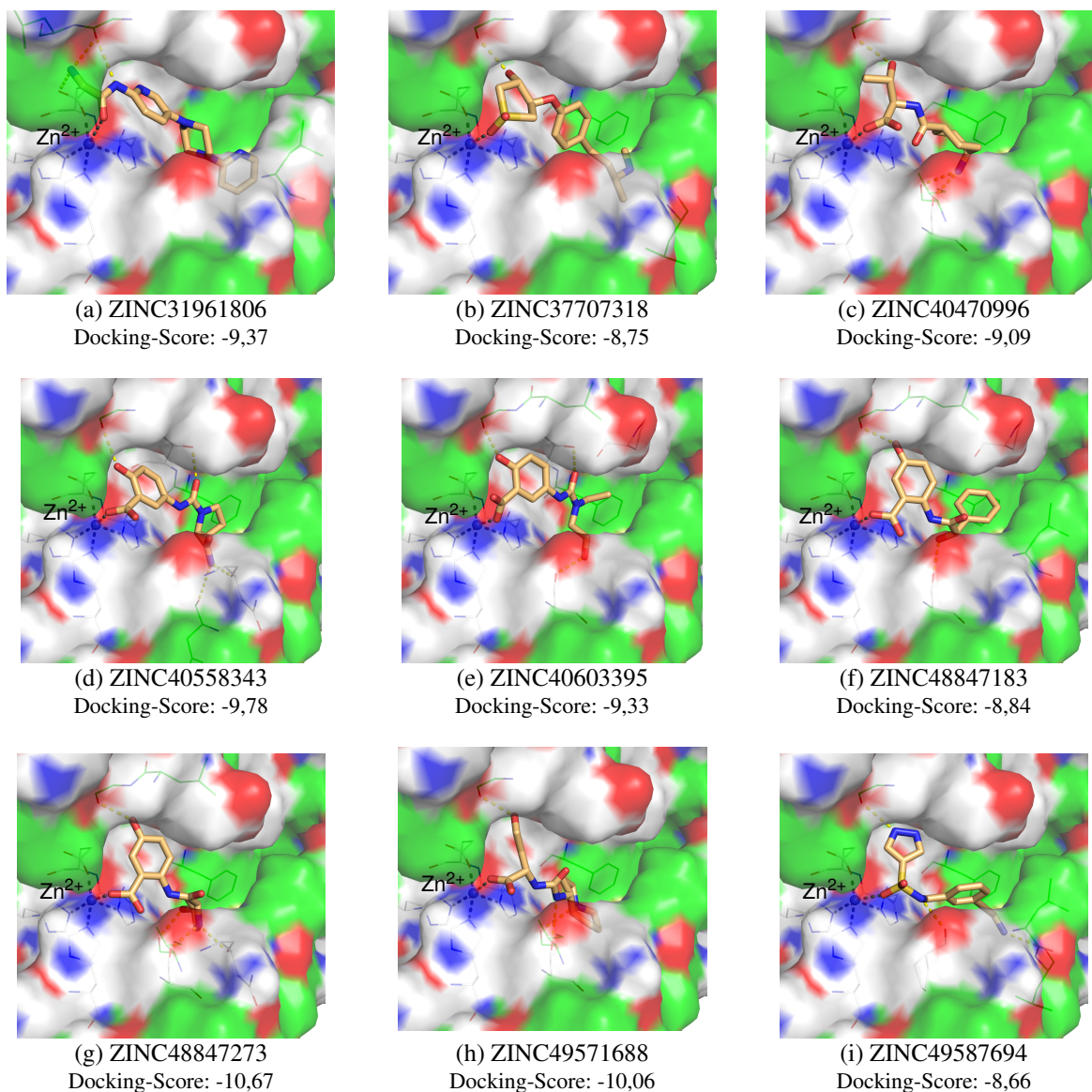
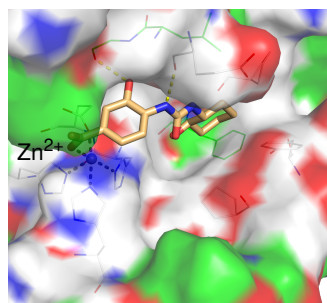
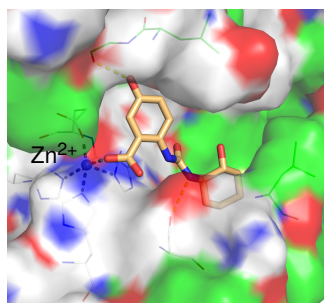


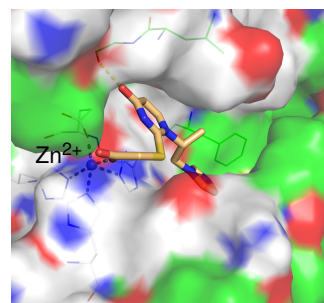
Abbildung A.8: Docking-Posen der 18 möglicherweise selektiven Hits in ADAMTS7 mit Bindetasche in Oberflächendarstellung (grün: hydrophober Bereich; rot: Akzeptorbereich; blau: Donorbereich). Die Bilder sind analog zu den Abbildungen 6.6-6.10 um 15° nach oben gedreht. Die Zink-koordinierenden Histidine sind als Linien dargestellt, genau wie die jeweils mit dem Liganden wechselwirkenden Aminosäuren. Die Liganden werden durch orange Stäbchen dargestellt und die Zink-Ionen sind als graue Kugeln zu erkennen. Zink-koordinierende Bindungen sind durch grau gestrichelte Linien dargestellt, Wasserstoffbrücken durch gelb gestrichelte Linien.



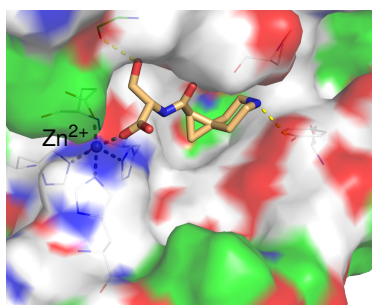
(j) ZINC52012042
Docking-Score: -9,03



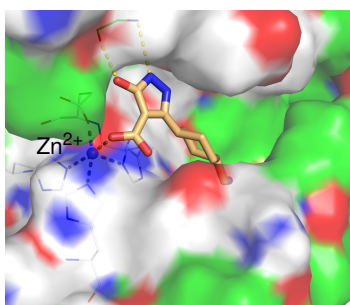
(k) ZINC52012291
Docking-Score: -8,65



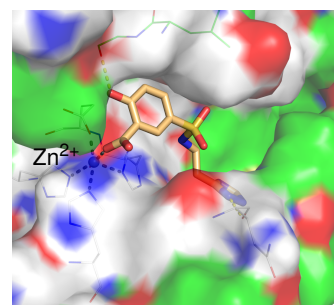
(l) ZINC52329369
Docking-Score: -10,13



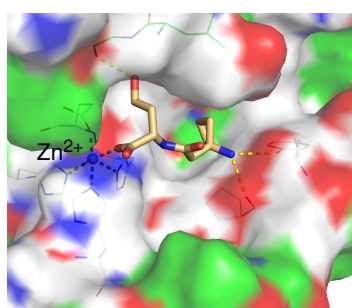
(m) ZINC70254898
Docking-Score: -8,98



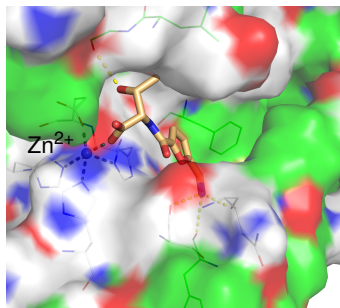
(n) ZINC75953566
Docking-Score: -9,57



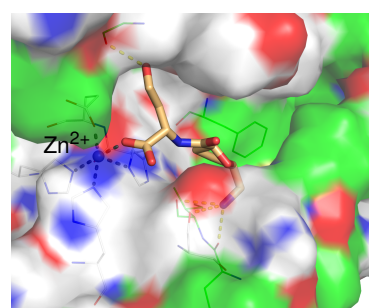
(o) ZINC80583404
Docking-Score: -8,51



(p) ZINC86853423
Docking-Score: -8,40



(q) ZINC94569517
Docking-Score: -9,33



(r) ZINC94569992
Docking-Score: -9,22

Abbildung A.8 (Fortsetzung): Docking-Posen der 18 möglicherweise selektiven Hits in ADAMTS7 mit Bindetasche in Oberflächendarstellung (grün: hydrophober Bereich; rot: Akzeptorbereich; blau: Donorbereich). Die Bilder sind analog zu den Abbildungen 6.6-6.10 um 15° nach oben gedreht. Die Zink-kordinierenden Histidine sind als Linien dargestellt, genau wie die jeweils mit dem Liganden wechselwirkenden Aminosäuren. Die Liganden werden durch orange Stäbchen dargestellt und die Zink-Ionen sind als graue Kugeln zu erkennen. Zink-kordinierende Bindungen sind durch grau gestrichelte Linien dargestellt, Wasserstoffbrücken durch gelb gestrichelte Linien.

Literaturverzeichnis

- [1] Rehmann, W. A.; Greve, K.: Arzneimittelgesetz: AMG. In: *Variations*, 1999, 29(1): 23–25.
- [2] Klebe, G.: *Wirkstoffdesign: Entwurf und Wirkung von Arzneistoffen*. 2. Auflage. Spektrum Akademischer Verlag, 2009.
- [3] Paul, S. M.; Mytelka, D. S.; Dunwiddie, C. T.; Persinger, C. C.; Munos, B. H.; Lindborg, S. R.; Schacht, A. L.: How to improve R&D productivity: the pharmaceutical industry's grand challenge. In: *Nat Rev Drug Discov*, 2010, 9: 203–214.
- [4] *research - Das Bayer-Forschungsmagazin*. 2014.
- [5] Sliwoski, G.; Kothiwale, S.; Meiler, J.; Lowe, E. W.: Computational methods in drug discovery. In: *Pharmacol Rev*, 2014, 66: 334–395.
- [6] Kapetanovic, I. M.: Computer-aided drug discovery and development (CADD): in silico-chemico-biological approach. In: *Chem Biol Interact*, 2008, 171: 165–176.
- [7] Blausen.com staff: *Blausen gallery 2014*. 1 (2), 2014.
- [8] Kessler, T.; Vilne, B.; Schunkert, H.: The impact of genome-wide association studies on the pathophysiology and therapy of cardiovascular disease. In: *EMBO Mol Med*, 2016, 8(7): 688–701.
- [9] Wagner, A. K.; Schunkert, H.; Erdmann, J.: Genetik des Herzinfarktes. Der lange Weg von der positiven Familienanamnese zum Gen. In: *Chemie in unserer Zeit*, 2009, 43(5): 288–295.
- [10] Ozaki, K.; Tanaka, T.: Molecular genetics of coronary artery disease. In: *J Hum Genet*, 2016, 61(1): 71–77.
- [11] Roberts, R.; Stewart, A. F. R.: Genes and coronary artery disease: where are we? In: *J Am Coll Cardiol*, 2012, 60(18): 1715–1721.
- [12] Thompson, R. C.; Allam, A. H.; Lombardi, G. P.; Wann, L. S.; Sutherland, M. L.; Sutherland, J. D.; Soliman, M. A.; Frohlich, B.; et al.: Atherosclerosis across 4000 years of human history: the Horus study of four ancient populations. In: *Lancet*, 2013, 381: 1211–1222.
- [13] Samani, N. J.; Erdmann, J.; Hall, A. S.; Hengstenberg, C.; Mangino, M.; Mayer, B.; Dixon, R. J.; Meitinger, T.; et al.: Genomewide association analysis of coronary artery disease. In: *N Engl J Med*, 2007, 357(5): 443–453.
- [14] Helgadottir, A.; Thorleifsson, G.; Manolescu, A.; Gretarsdottir, S.; Blondal, T.; Jonasdottir, A.; Jonasdottir, A.; Sigurdsson, A.; et al.: A common variant on chromosome 9p21 affects the risk of myocardial infarction. In: *Science*, 2007, 316(5830): 1491–1493.
- [15] McPherson, R.; Pertsemlidis, A.; Kavaslar, N.; Stewart, A.; Roberts, R.; Cox, D. R.; Hinds, D. A.; Pennacchio, L. A.; et al.: A common allele on chromosome 9 associated with coronary heart disease. In: *Science*, 2007, 316(5830): 1488–1491.
- [16] Erdmann, J.; Grosshennig, A.; Braund, P. S.; König, I. R.; Hengstenberg, C.; Hall, A. S.; Linsel-Nitschke, P.; Kathiresan, S.; et al.: New susceptibility locus for coronary artery disease on chromosome 3q22.3. In: *Nat Genet*, 2009, 41(3): 280–282.
- [17] Schunkert, H.; König, I. R.; Kathiresan, S.; Reilly, M. P.; Assimes, T. L.; Holm, H.; Preuss, M.; Stewart, A. F. R.; et al.: Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. In: *Nat Genet*, 2011, 43(4): 333–338.
- [18] Coronary Artery Disease (C4D) Genetics Consortium: A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. In: *Nat Genet*, 2011, 43(4): 339–344.
- [19] CARDIoGRAMplusC4D Consortium; Deloukas, P.; Kanoni, S.; Willenborg, C.; Farrall, M.; Assimes, T. L.; Thompson, J. R.; Ingelsson, E.; et al.: Large-scale association analysis identifies new risk loci for coronary artery disease. In: *Nat Genet*, 2013, 45(1): 25–33.

- [20] Nikpay, M.; Goel, A.; Won, H.; Hall, L. M.; Willenborg, C.; Kanoni, S.; Saleheen, D.; Kyriakou, T.; et al.: A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. In: *Nat Genet*, 2015, 47(10): 1121–1130.
- [21] Krzywinski, M.; Schein, J.; Birol, I.; Connors, J.; Gascoyne, R.; Horsman, D.; Jones, S. J.; Marra, M. A.: Circos: an information aesthetic for comparative genomics. In: *Genome Res*, 2009, 19(9): 1639–1645.
- [22] Webb, T. R.; Erdmann, J.; Stirrups, K. E.; Stitzel, N. O.; Masca, N. G. D.; Jansen, H.; Kanoni, S.; Nelson, C. P.; et al.: Systematic Evaluation of Pleiotropy Identifies 6 Further Loci Associated With Coronary Artery Disease. In: *J Am Coll Cardiol*, 2017, 69(7): 823–836.
- [23] Howson, J. M. M.; Zhao, W.; Barnes, D. R.; Ho, W.; Young, R.; Paul, D. S.; Waite, L. L.; Freitag, D. F.; et al.: Fifteen new risk loci for coronary artery disease highlight arterial-wall-specific mechanisms. In: *Nat Genet*, 2017, 49(7): 1113–1119.
- [24] Cohen, J. C.; Boerwinkle, E.; Mosley, T. H. Jr; Hobbs, H. H.: Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. In: *N Engl J Med*, 2006, 354(12): 1264–1272.
- [25] Myocardial Infarction Genetics Consortium; Kathiresan, S.; Voight, B. F.; Purcell, S.; Musunuru, K.; Ardissino, D.; Mannucci, P. M.; Anand, S.; et al.: Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. In: *Nat Genet*, 2009, 41(3): 334–341.
- [26] IBC 50K CAD Consortium: Large-scale gene-centric analysis identifies novel variants for coronary artery disease. In: *PLoS Genet*, 2011, 7(9): e1002260.
- [27] Trégouët, D.; König, I. R.; Erdmann, J.; Munteanu, A.; Braund, P. S.; Hall, A. S.; Grosshennig, A.; Linsel-Nitschke, P.; et al.: Genome-wide haplotype association study identifies the SLC22A3-LPAL2-LPA gene cluster as a risk locus for coronary artery disease. In: *Nat Genet*, 2009, 41(3): 283–285.
- [28] Reilly, M. P.; Li, M.; He, J.; Ferguson, J. F.; Stylianou, I. M.; Mehta, N. N.; Burnett, M. S.; Devaney, J. M.; et al.: Identification of ADAMTS7 as a novel locus for coronary atherosclerosis and association of ABO with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies. In: *Lancet*, 2011, 377(9763): 383–392.
- [29] Teslovich, T. M.; Musunuru, K.; Smith, A. V.; Edmondson, A. C.; Stylianou, I. M.; Koseki, M.; Pirruccello, J. P.; Ripatti, S.; et al.: Biological, clinical and population relevance of 95 loci for blood lipids. In: *Nature*, 2010, 466(7307): 707–713.
- [30] Patel, R. S.; Ye, S.: ADAMTS7: a promising new therapeutic target in coronary heart disease. In: *Expert Opin Ther Targets*, 2013, 17(8): 863–867.
- [31] Wang, L.; Zheng, J.; Bai, X.; Liu, B.; Liu, C.; Xu, Q.; Zhu, Y.; Wang, N.; Kong, W.; Wang, X.: ADAMTS-7 mediates vascular smooth muscle cell migration and neointima formation in balloon-injured rat arteries. In: *Circ Res*, 2009, 104(5): 688–698.
- [32] Kessler, T.; Zhang, L.; Liu, Z.; Yin, X.; Huang, Y.; Wang, Y.; Fu, Y.; Mayr, M.; et al.: ADAMTS-7 inhibits re-endothelialization of injured arteries and promotes vascular remodeling through cleavage of thrombospondin-1. In: *Circulation*, 2015, 131(13): 1191–1201.
- [33] Liu, C.; Kong, W.; Ilalov, S. K. and Y. K. and Yu; Xu, K.; Prazak, L.; Fajardo, M.; Sehgal, B.; Di Cesare, P. E.: ADAMTS-7: a metalloproteinase that directly binds to and degrades cartilage oligomeric matrix protein. In: *FASEB J*, 2006, 20(7): 988–990.
- [34] Bauer, R. C.; Tohyama, J.; Cui, J.; Cheng, L.; Yang, J.; Zhang, X.; Ou, K.; Paschos, G. K.; et al.: Knockout of Adamts7, A Novel CAD Locus in Humans, Reduces Atherosclerosis in Mice. In: *Circulation*, 2015, 131: 1202–1213.
- [35] Takeda, S.: ADAM and ADAMTS Family Proteins and Snake Venom Metalloproteinases: A Structural Overview. In: *Toxins*, 2016, 8(5): 155.
- [36] Kelwick, R.; Desanlis, I.; Wheeler, G. N.; Edwards, D. R.: The ADAMTS (A Disintegrin and Metalloproteinase with Thrombospondin motifs) family. In: *Genome Biol*, 2015, 16(1): 113.
- [37] Apte, S. S.: A disintegrin-like and metalloprotease (reprolysin type) with thrombospondin type 1 motifs: the ADAMTS family. In: *Int J Biochem Cell Biol*, 2004, 36(6): 981–985.

- [38] Tortorella, M. D.; Malfait, F.; Barve, R. A.; Shieh, H.-S.; Malfait, A.-M.: A review of the ADAMTS family, pharmaceutical targets of the future. In: *Curr Pharm Des*, 2009, 15(20): 2359–2374.
- [39] Huxley-Jones, J.; Clarke, T.; Beck, C.; Toubaris, G.; Robertson, D. L.; Boot-Handford, R. P.: The evolution of the vertebrate metzincins; insights from *Ciona intestinalis* and *Danio rerio*. In: *BMC Evol Biol*, 2007, 7: 63.
- [40] Gomis-Rüth, F. X.: Catalytic domain architecture of metzincin metalloproteases. In: *J Biol Chem*, 2009, 284(23): 15353–15357.
- [41] Porter, S.; Clark, I. M.; Kevorkian, L.; Edwards, D. R.: The ADAMTS metalloproteinases. In: *Biochem J*, 2005, 386(Pt 1): 15–27.
- [42] Edwards, D. R.; Handsley, M. M.; Pennington, C. J.: The ADAM metalloproteinases. In: *Mol Aspects Med*, 2008, 29(5): 258–289.
- [43] Bourboulia, D.; Stetler-Stevenson, W. G.: Matrix metalloproteinases (MMPs) and tissue inhibitors of metalloproteinases (TIMPs): Positive and negative regulators in tumor cell adhesion. In: *Semin Cancer Biol*, 2010, 20(3): 161–168.
- [44] Jones, G. C.; Riley, G. P.: ADAMTS proteinases: a multi-domain, multi-functional family with roles in extracellular matrix turnover and arthritis. In: *Arthritis Res Ther*, 2005, 7(4): 160–169.
- [45] Liu, C.: The role of ADAMTS-7 and ADAMTS-12 in the pathogenesis of arthritis. In: *Nat Clin Pract Rheumatol*, 2009, 5(1): 38–45.
- [46] Brocker, C. N.; Vasiliou, V.; Nebert, D. W.: Evolutionary divergence and functions of the ADAM and ADAMTS gene families. In: *Hum Genomics*, 2009, 4(1): 43–55.
- [47] Mosyak, L.; Georgiadis, K.; Shane, T.; Svenson, K.; Hebert, T.; McDonagh, T.; Mackie, S.; Olland, S.; et al.: Crystal structures of the two major aggrecan degrading enzymes, ADAMTS4 and ADAMTS5. In: *Protein Sci*, 2008, 17(1): 16–21.
- [48] Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E.: The Protein Data Bank. In: *Nucleic Acids Res*, 2000, 28(1): 235–242.
- [49] Gerhardt, S.; Hassall, G.; Hawtin, P.; McCall, E.; Flavell, L.; Minshull, C.; Hargreaves, D.; Ting, A.; et al.: Crystal structures of human ADAMTS-1 reveal a conserved catalytic domain and a disintegrin-like domain with a fold homologous to cysteine-rich domains. In: *J Mol Biol*, 2007, 373(4): 891–902.
- [50] Tortorella, M. D.; Tomasselli, A. G.; Mathis, K. J.; Schnute, M. E.; Woodard, S. S.; Munie, G.; Williams, J. M.; Caspers, N.; et al.: Structural and inhibition analysis reveals the mechanism of selectivity of a series of aggrecanase inhibitors. In: *J Biol Chem*, 2009, 284(36): 24185–24191.
- [51] Akiyama, M.; Takeda, S.; Kokame, K.; Takagi, J.; Miyata, T.: Crystal structures of the noncatalytic domains of ADAMTS13 reveal multiple discontinuous exosites for von Willebrand factor. In: *Proc Natl Acad Sci USA*, 2009, 106(46): 19274–19279.
- [52] Hu, J.; Van den Steen, P. E.; Sang, Q. A.; Opdenakker, G.: Matrix metalloproteinase inhibitors as therapy for inflammatory and vascular diseases. In: *Nat Rev Drug Discov*, 2007, 6(6): 480–498.
- [53] Liu, C.; Kong, W.; Xu, K.; Luan, Y.; Ilalov, K.; Sehgal, B.; Yu, S.; Howell, R. D.; Di Cesare, P. E.: ADAMTS-12 associates with and degrades cartilage oligomeric matrix protein. In: *J Biol Chem*, 2006, 281: 15800–15808.
- [54] Pi, L.; Jorgensen, M.; Oh, S.; Protopapadakis, Y.; Gjymishka, A.; Brown, A.; Robinson, P.; Liu, C.; et al.: A disintegrin and metalloprotease with thrombospondin type I motif 7: a new protease for connective tissue growth factor in hepatic progenitor/oval cell niche. In: *Am J Pathol*, 2015, 185(6): 1552–1563.
- [55] Guo, F.; Lai, Y.; Tian, Q.; Lin, E. A.; Kong, L.; Liu, C.: Granulin-epithelin precursor binds directly to ADAMTS-7 and ADAMTS-12 and inhibits their degradation of cartilage oligomeric matrix protein. In: *Arthritis Rheum*, 2010, 62(7): 2023–2036.
- [56] Luan, Y.; Kong, L.; Howell, D. R.; Ilalov, K.; Fajardo, M.; Bai, X.-H.; Di Cesare, P. E.; Goldring, M. B.; Abramson, S. B.; Liu, C.-J.: Inhibition of ADAMTS-7 and ADAMTS-12 degradation of cartilage oligomeric matrix protein by alpha-2-macroglobulin. In: *Osteoarthritis Cartilage*, 2008, 16(11): 1413–1420.

- [57] Levy, G. G.; Nichols, W. C.; Lian, E. C.; Foroud, T.; McClintick, J. N.; McGee, B. M.; Yang, A. Y.; Siemieniak, D. R.; et al.: Mutations in a member of the ADAMTS gene family cause thrombotic thrombocytopenic purpura. In: *Nature*, 2001, 413(6855): 488–494.
- [58] *The Human Protein Atlas - ADAMTS7*. (Stand: 27.07.2017).. – <http://www.proteinatlas.org/ENSG00000136378-ADAMTS7/tissue>
- [59] Uhlén, M.; Fagerberg, L.; Hallström, B. M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; Sivertsson, Å.; Kampf, C.; et al.: Proteomics. Tissue-based map of the human proteome. In: *Science*, 2015, 347: 1260419.
- [60] Hurskainen, T. L.; Hirohata, S.; Seldin, M. F.; Apte, S. S.: ADAM-TS5, ADAM-TS6, and ADAM-TS7, novel members of a new family of zinc metalloproteases. General features and genomic distribution of the ADAM-TS family. In: *J Biol Chem*, 1999, 274(36): 25555–25563.
- [61] Somerville, R. P. T.; Longpré, J.-M.; Apel, E. D.; Lewis, R. M.; Wang, L. W.; Sanes, J. R.; Leduc, R.; Apte, S. S.: ADAMTS7B, the full-length product of the ADAMTS7 gene, is a chondroitin sulfate proteoglycan containing a mucin domain. In: *J Biol Chem*, 2004, 279(34): 35159–35175.
- [62] Bai, X.; Wang, D.; Kong, L.; Zhang, Y.; Luan, Y.; Kobayashi, T.; Kronenberg, H. M.; Yu, X.; Liu, C.: ADAMTS-7, a direct target of PTHrP, adversely regulates endochondral bone growth by associating with and inactivating GEP growth factor. In: *Mol Cell Biol*, 2009, 29(15): 4201–4219.
- [63] Pu, X.; Xiao, Q.; Kiechl, S.; Chan, K.; Ng, F. L.; Gor, S.; Poston, R. N.; Fang, C.; et al.: ADAMTS7 cleavage and vascular smooth muscle cell migration is affected by a coronary-artery-disease-associated variant. In: *Am J Hum Genet*, 2013, 92(3): 366–374.
- [64] Krieger, E.; Nabuurs, S. B.; Vriend, G.: Homology modeling. In: *Methods Biochem Anal*, 2003, 44: 509–524.
- [65] Pei, J.; Kim, B.; Grishin, N. V.: PROMALS3D: a tool for multiple protein sequence and structure alignments. In: *Nucleic Acids Res*, 2008, 36(7): 2295–2300.
- [66] Pertsevlidis, A.; Fondon, J. W.: Having a BLAST with bioinformatics (and avoiding BLASTphemy). In: *Genome Biol*, 2001, 2(10): reviews2002.1–10.
- [67] Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J.: Basic local alignment search tool. In: *J Mol Biol*, 1990, 215(3): 403–410.
- [68] Needleman, S. B.; Wunsch, C. D.: A general method applicable to the search for similarities in the amino acid sequence of two proteins. In: *J Mol Biol*, 1970, 48(3): 443–453.
- [69] Smith, T. F.; Waterman, M. S.: Identification of common molecular subsequences. In: *J Mol Biol*, 1981, 147(1): 195–197.
- [70] Martins, W. S.; Del Cuvillo, J. B.; Useche, F. J.; Theobald, K. B.; Gao, G. R.: A multithreaded parallel implementation of a dynamic programming algorithm for sequence comparison. In: *Pac Symp Biocomput*, 2001,
- [71] Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J.: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. In: *Nucleic Acids Res*, 1997, 25(17): 3389–3402.
- [72] Pearson, W. R.: Rapid and sensitive sequence comparison with FASTP and FASTA. In: *Methods Enzymol*, 1990, 183: 63–98.
- [73] Gromiha, M. M.: *Protein Bioinformatics: From Sequence to Function*. Academic Press, 2010
- [74] Jones, D. T.: Protein secondary structure prediction based on position-specific scoring matrices. In: *J Mol Biol*, 1999, 292(2): 195–202.
- [75] Rost, B.; Sander, C.: Improved prediction of protein secondary structure by use of sequence profiles and neural networks. In: *Proc Natl Acad Sci USA*, 1993, 90: 7558–7562.
- [76] Meiler, J.; Baker, D.: Coupled prediction of protein secondary and tertiary structure. In: *Proc Natl Acad Sci USA*, 2003, 100: 12105–12110.
- [77] Pirovano, W.; Heringa, J.: Protein Secondary Structure Prediction. Version: 2009. In: *Methods in Molecular Biology*. Springer Nature, 2009, 327–348.
- [78] Cole, C.; Barber, J. D.; Barton, G. J.: The Jpred 3 secondary structure prediction server. In: *Nucleic Acids Res*, 2008, 36(suppl_2): W197–W201.

- [79] Drozdetskiy, A.; Cole, C.; Procter, J.; Barton, G. J.: JPred4: a protein secondary structure prediction server. In: *Nucleic Acids Res*, 2015, 43(W1): W389–W394.
- [80] Laskowski, R. A.; Rullmann, J. A.; MacArthur, M. W.; Kaptein, R.; Thornton, J. M.: AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. In: *J Biomol NMR*, 1996, 8(4): 477–486.
- [81] Schwede, T.: Protein Modeling: What Happened to the “Protein Structure Gap”? In: *Structure*, 2013, 21(9): 1531–1540.
- [82] Anfinsen, C. B.: Principles that govern the protein folding chains. In: *Science*, 1973, 181: 233–230.
- [83] Chothia, C.; Lesk, A. M.: The relation between the divergence of sequence and structure in proteins. In: *EMBO J*, 1986, 5(4): 823–826.
- [84] Mihasan, M.: Basic protein structure prediction for the biologist: A review. In: *Arch Biol Sci*, 2010, 62(4): 857–871.
- [85] Venselaar, H.; Joosten, R. P.; Vroling, B.; Baakman, C. A. B.; Hekkelman, M. L.; Krieger, E.; Vriend, G.: Homology modelling and spectroscopy, a never-ending love story. In: *Eur Biophys J*, 2009, 39(4): 551–563.
- [86] Forster, M. J.: Molecular modelling in structural biology. In: *Micron*, 2002, 33(4): 365–384.
- [87] Rost, B.; Schneider, R.; Sander, C.: Protein fold recognition by prediction-based threading. In: *J Mol Biol*, 1997, 270(3): 471–480.
- [88] Kleywegt, G. J.; Jones, T. A.: Model building and refinement practice. Version: 1997. In: *Methods in Enzymology*. Elsevier, 1997, 208–230.
- [89] Nayeem, A.: A comparative study of available software for high-accuracy homology modeling: From sequence alignments to structural models. In: *Protein Sci*, 2006, 15(4): 808–824.
- [90] Kapitel 18. In: Dolan, M. A.; Noah, J. W.; Hurt, D.: *Comparison of Common Homology Modeling Algorithms: Application of User-Defined Alignments*. Totowa, NJ : Humana Press, 2012, 399–414
- [91] Krieger, E.; Joo, K.; Lee, J.; Lee, J.; Raman, S.; Thompson, J.; Tyka, M.; Baker, D.; Karplus, K.: Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. In: *Proteins*, 2009, 77(Suppl 9): 114–122.
- [92] Oshiro, C.; Bradley, E. K.; Eksterowicz, J.; Evensen, E.; Lamb, M. L.; Lanctot, J. K.; Putta, S.; Stanton, R.; Grootenhuis, P. D. J.: Performance of 3D-Database Molecular Docking Studies into Homology Models. In: *J Med Chem*, 2004, 47(3): 764–767.
- [93] Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T.: The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. In: *Bioinformatics*, 2006, 22(2): 195–201.
- [94] Guex, N.; Peitsch, M. C.; Schwede, T.: Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. In: *Electrophoresis*, 2009, 30(S1): S162–S173.
- [95] Krivov, G. G.; Shapovalov, M. V.; Dunbrack, R. L.: Improved prediction of protein side-chain conformations with SCWRL4. In: *Proteins*, 2009, 77(4): 778–795.
- [96] Biasini, M.; Bienert, S.; Waterhouse, A.; Arnold, K.; Studer, G.; Schmidt, T.; Kiefer, F.; Cassarino, T. G.; et al.: SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. In: *Nucleic Acids Res*, 2014, 42(W1): W252–W258.
- [97] Benkert, P.; Künzli, M.; Schwede, T.: QMEAN server for protein model quality estimation. In: *Nucleic Acids Res*, 2009, 37(suppl_2): W510–W514.
- [98] Benkert, P.; Biasini, M.; Schwede, T.: Toward the estimation of the absolute quality of individual protein structure models. In: *Bioinformatics*, 2010, 27(3): 343–350.
- [99] MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; et al.: All-atom empirical potential for molecular modeling and dynamics studies of proteins. In: *J Phys Chem B*, 1998, 102: 3586–3616.
- [100] MacKerell, A. D.; Feig, M.; Brooks, C. L.: Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. In: *J Comput Chem*, 2004, 25: 1400–1415.

- [101] Canutescu, A. A.; Shelenkov, A. A.; Dunbrack, R. L. Jr: A graph-theory algorithm for rapid protein side-chain prediction. In: *Protein Sci*, 2003, 12(9): 2001–2014.
- [102] Hooft, R. W. W.; Vriend, G.; Sander, C.; Abola, E. E.: Errors in protein structures. In: *Nature*, 1996, 381(6580): 272–272.
- [103] Jacobson, M. P.; Pincus, D. L.; Rapp, C. S.; Day, T. J. F.; Honig, B.; Shaw, D. E.; Friesner, R. A.: A hierarchical approach to all-atom protein loop prediction. In: *Proteins*, 2004, 55(2): 351–367.
- [104] Banks, J. L.; Beard, H. S.; Cao, Y.; Cho, A. E.; Damm, W.; Farid, R.; Felts, A. K.; Halgren, T. A.; et al.: Integrated Modeling Program, Applied Chemical Theory (IMPACT). In: *J Comput Chem*, 2005, 26(16): 1752–1780.
- [105] Carlson, H. A.: Protein flexibility is an important component of structure-based drug discovery. In: *Current pharmaceutical design*, 2002, 8: 1571–1578.
- [106] MacArthur, M. W.; Thornton, J. M.: Conformational analysis of protein structures derived from NMR data. In: *Proteins*, 1993, 17: 232–251.
- [107] Sippl, M. J.: Recognition of errors in three-dimensional structures of proteins. In: *Proteins*, 1993, 17(4): 355–362.
- [108] Wiederstein, M.; Sippl, M. J.: ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. In: *Nucleic Acids Res*, 2007, 35: W407–W410.
- [109] Pawlowski, M.; Gajda, M. J.; Matlak, R.; Bujnicki, J. M.: MetaMQAP: a meta-server for the quality assessment of protein models. In: *BMC Bioinformatics*, 2008, 9: 403.
- [110] Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M.: *PROCHECK*: a program to check the stereochemical quality of protein structures. In: *J Appl Cryst*, 1993, 26(2): 283–291.
- [111] Hospital, A.; Goñi, J. R.; Orozco, M.; Gelpí, J. L.: Molecular dynamics simulations: advances and applications. In: *Adv Appl Bioinform Chem*, 2015, 8: 37–47.
- [112] Guliaev, A. B.; Cheng, S.; Hang, B.: Protein dynamics via computational microscope. In: *World J Methodol*, 2012, 2: 42–49.
- [113] Monticelli, L.; Tieleman, D. P.: Force fields for classical molecular dynamics. In: *Biomolecular simulations: Methods and protocols*, 2013, 924: 197–213.
- [114] Leach, A. R.: *Molecular modelling : principles and applications*. 2. Auflage. Pearson Prentice Hall, 2001.
- [115] Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M.: CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. In: *J Comput Chem*, 1983, 4(2): 187–217.
- [116] Krieger, E.; Vriend, G.: YASARA View - molecular graphics for all devices - from smartphones to workstations. In: *Bioinformatics*, 2014, 30: 2981–2982.
- [117] Krieger, E.; Darden, T.; Nabuurs, S. B.; Finkelstein, A.; Vriend, G.: Making optimal use of empirical energy functions: force-field parameterization in crystal space. In: *Proteins*, 2004, 57: 678–683.
- [118] Guex, N.; Peitsch, M. C.: SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. In: *Electrophoresis*, 1997, 18(15): 2714–2723.
- [119] Karplus, M.; McCammon, J. A.: Molecular dynamics simulations of biomolecules. In: *Nat Struct Biol*, 2002, 9(9): 646–652.
- [120] Karplus, M.; Petsko, G. A.: Molecular dynamics simulations in biology. In: *Nature*, 1990, 347(6294): 631–639.
- [121] Bowers, K. J.; Chow, E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B.A.; Klepeis, J. L.; Kolossvary, I.; et al.: Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. In: *Proceedings of the ACM/IEEE Conference on Supercomputing*, 2006, 0: 43.
- [122] Walton, E. B.; VanVliet, K. J.: Equilibration of experimentally determined protein structures for molecular dynamics simulation. In: *Physical Review E*, 2006, 74(6): 061901.
- [123] Grossfield, A.; Zuckerman, D. M.: Quantifying uncertainty and sampling quality in biomolecular simulations. In: *Annual reports in computational chemistry*, 2009, 5: 23–48.
- [124] Martínez, L.: Automatic Identification of Mobile and Rigid Substructures in Molecular Dynamics Simulations and Fractional Structural Fluctuation Analysis. In: *PLoS One*, 2015, 10(3): e0119264.

- [125] Halgren, T.: New method for fast and accurate binding-site identification and analysis. In: *Chem Biol Drug Des*, 2007, 69(2): 146–148.
- [126] Halgren, T. A.: Identifying and characterizing binding sites and assessing druggability. In: *J Chem Inf Model*, 2009, 49(2): 377–389.
- [127] Laskowski, R. A.: SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions. In: *J Mol Graph*, 1995, 13(5): 323–330.
- [128] Van Drie, J. H.: Monty Kier and the Origin of the Pharmacophore Concept. In: *Internet Electron J Mol Des*, 2007, 6: 271–279.
- [129] Kaserer, T.; Beck, K. R.; Akram, M.; Odermatt, A.; Schuster, D.: Pharmacophore Models and Pharmacophore-Based Virtual Screening: Concepts and Applications Exemplified on Hydroxysteroid Dehydrogenases. In: *Molecules*, 2015, 20: 22799–22832.
- [130] Yang, S.-Y.: Pharmacophore modeling and applications in drug discovery: challenges and recent advances. In: *Drug Discov Today*, 2010, 15(11-12): 444–450.
- [131] Dixon, S. L.; Smondyrev, A. M.; Knoll, E. H.; Rao, S. N.; Shaw, D. E.; Friesner, R. A.: PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. In: *J Comput Aided Mol Des*, 2006, 20(10-11): 647–671.
- [132] Qing, X.; Lee, X. Y.; De Raeymaecker, J.; Tame, J. R. H.; Zhang, K. Y. J.; De Maeyer, M.; Voet, A.: Pharmacophore modeling: advances, limitations, and current utility in drug discovery. In: *J Receptor Ligand Channel Res*, 2014, 7: 81–92.
- [133] Salam, N. K.; Nuti, R.; Sherman, W.: Novel method for generating structure-based pharmacophores using energetic analysis. In: *J Chem Inf Model*, 2009, 49(10): 2356–2368.
- [134] Loving, K.; Salam, N. K.; Sherman, W.: Energetic analysis of fragment docking and application to structure-based pharmacophore hypothesis generation. In: *J Comput Aided Mol Des*, 2009, 23(8): 541–554.
- [135] *Glide Fragment Library*. (Stand: 27.07.2017). – <http://www.schrodinger.com/Glide/Fragment-Library>
- [136] Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; et al.: Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. In: *J Med Chem*, 2004, 47(7): 1739–1749.
- [137] Huang, S.; Zou, X.: Advances and challenges in protein-ligand docking. In: *Int J Mol Sci*, 2010, 11: 3016–3034.
- [138] Akhter, M.: Challenges in Docking: Mini Review. In: *JSM Chem*, 2016, 4(2): 1025.
- [139] Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J.: AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. In: *J Comput Chem*, 2009, 30(16): 2785–2791.
- [140] Trott, O.; Olson, A. J.: AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. In: *J Comput Chem*, 2010, 31(2): 455–461.
- [141] Korb, O.; Stützle, T.; Exner, T. E.: An ant colony optimization approach to flexible protein–ligand docking. In: *Swarm Intelligence*, 2007, 1(2): 115–134.
- [142] Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J.: Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. In: *J Comput Chem*, 1998, 19(14): 1639–1662.
- [143] Nelder, J. A.; Mead, R.: A simplex method for function minimization. In: *The Computer Journal*, 1965, 7(4): 308–313.
- [144] Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L.: Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. In: *J Med Chem*, 2004, 47(7): 1750–1759.
- [145] Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T.: Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes. In: *J Med Chem*, 2006, 49(21): 6177–6196.

- [146] Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P.: Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. In: *J Comput Aided Mol Des*, 1997, 11(5): 425–445.
- [147] Chen, Y. Z.; Zhi, D. G.: Ligand-protein inverse docking and its potential use in the computer search of protein targets of a small molecule. In: *Proteins*, 2001, 43(2): 217–226.
- [148] Grinter, S. Z.; Liang, Y.; Huang, S.; Hyder, S. M.; Zou, X.: An inverse docking approach for identifying new potential anti-cancer targets. In: *J Mol Graph Model*, 2011, 29(6): 795–799.
- [149] Wang, W.; Zhou, X.; He, W.; Fan, Y.; Chen, Y.; Chen, X.: The interprotein scoring noises in glide docking scores. In: *Proteins*, 2012, 80(1): 169–183.
- [150] Luo, Q.; Zhao, L.; Hu, J.; Jin, H.; Liu, Z.; Zhang, L.: The scoring bias in reverse docking and the score normalization strategy to improve success rate of target fishing. In: *PLoS One*, 2017, 12(2): e0171433.
- [151] Vigers, G. P. A.; Rizzi, J. P.: Multiple active site corrections for docking and virtual screening. In: *J Med Chem*, 2004, 47(1): 80–89.
- [152] Broccatelli, F.; Brown, N.: Best of both worlds: on the complementarity of ligand-based and structure-based virtual screening. In: *J Chem Inf Model*, 2014, 54(6): 1634–1641.
- [153] Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G.: ZINC: a free tool to discover chemistry for biology. In: *J Chem Inf Model*, 2012, 52(7): 1757–1768.
- [154] Sastry, M.; Lowrie, J. F.; Dixon, S. L.; Sherman, W.: Large-scale systematic analysis of 2D fingerprint methods and parameters to improve virtual screening enrichments. In: *J Chem Inf Model*, 2010, 50(5): 771–784.
- [155] Cheng, T.; Li, Q.; Zhou, Z.; Wang, Y.; Bryant, S. H.: Structure-Based Virtual Screening for Drug Discovery: a Problem-Centric Review. In: *AAPS J*, 2012, 14(1): 133–141.
- [156] Varnek, A. (Hrsg.); Tropsha, A. (Hrsg.): *Chemoinformatics Approaches to Virtual Screening*. RSC, 2008.
- [157] Lipinski, C. A.: Drug-like properties and the causes of poor solubility and poor permeability. In: *J Pharmacol Toxicol Methods*, 2000, 44(1): 235–249.
- [158] Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J.: Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. In: *Adv Drug Deliv Rev*, 2001, 46(1-3): 3–26.
- [159] Duan, J.; Dixon, S. L.; Lowrie, J. F.; Sherman, W.: Analysis and comparison of 2D fingerprints: insights into database screening performance using eight fingerprint methods. In: *J Mol Graph Model*, 2010, 29(2): 157–170.
- [160] Bender, A.; Mussa, H. Y.; Glen, R. C.; Reiling, S.: Similarity searching of chemical databases using atom environment descriptors (MOLPRINT 2D): evaluation of performance. In: *J Chem Inf Comput Sci*, 2004, 44(5): 1708–1718.
- [161] *SIAS web service*. (Stand: 27.07.2017). – <http://imed.med.ucm.es/Tools/sias.html>
- [162] Cavasotto, C. N.; Phatak, S. S.: Homology modeling in drug discovery: current trends and applications. In: *Drug Discov Today*, 2009, 14(13-14): 676–683.
- [163] De Savi, C.; Pape, A.; Cumming, J. G.; Ting, A.; Smith, P. D.; Burrows, J. N.; Mills, M.; Davies, C.; et al.: The design and synthesis of novel N-hydroxyformamide inhibitors of ADAM-TS4 for the treatment of osteoarthritis. In: *Bioorg Med Chem Lett*, 2011, 21(5): 1376–1381.
- [164] Shieh, H.; Mathis, K. J.; Williams, J. M.; Hills, R. L.; Wiese, J. F.; Benson, T. E.; Kiefer, J. R.; Marino, M. H.; et al.: High resolution crystal structure of the catalytic domain of ADAMTS-5 (aggrecanase-2). In: *J Biol Chem*, 2008, 283(3): 1501–1507.
- [165] Shieh, H.; Tomasselli, A. G.; Mathis, K. J.; Schnute, M. E.; Woodard, S. S.; Caspers, N.; Williams, J. M.; Kiefer, J. R.; et al.: Structure analysis reveals the flexibility of the ADAMTS-5 active site. In: *Protein Sci*, 2011, 20(4): 735–744.
- [166] Akiyama, M.; Nakayama, D.; Takeda, S.; Kokame, K.; Takagi, J.; Miyata, T.: Crystal structure and enzymatic activity of an ADAMTS-13 mutant with the East Asian-specific P475S polymorphism. In: *J Thromb Haemost*, 2013, 11(7): 1399–1406.

- [167] Harding, M. M.: Small revisions to predicted distances around metal sites in proteins. In: *Acta Crystallogr D Biol Crystallogr*, 2006, 62(Pt 6): 678–682.
- [168] Berendsen, H. J. C.; Postma, J. P. M.; Gunsteren, W. F.; Hermans, J.: Interaction models for water in relation to protein hydration. In: *Intermolecular forces*. Springer, 1981, 331–342.
- [169] UniProt Consortium: UniProt: a hub for protein information. In: *Nucleic Acids Res*, 2015, 43(Database issue): D204–D212.
- [170] Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W.: Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. In: *J Comput Aided Mol Des*, 2013, 27(3): 221–234.
- [171] Harding, M. M.: The geometry of metal-ligand interactions relevant to proteins. In: *Acta Crystallogr D Biol Crystallogr*, 1999, 55(Pt 8): 1432–1443.
- [172] Hu, X.; Shelver, W. H.: Docking studies of matrix metalloproteinase inhibitors: zinc parameter optimization to improve the binding free energy prediction. In: *J Mol Graph Model*, 2003, 22(2): 115–126.
- [173] Müller, M.; Kessler, T.; Schunkert, H.; Erdmann, J.; Tennstedt, S.: Classification of ADAMTS binding sites: The first step toward selective ADAMTS7 inhibitors. In: *Biochem Biophys Res Commun*, 2016, 471(3): 380–385.
- [174] Santiago, D. N.; Pevzner, Y.; Durand, A. A.; Tran, M.; Scheerer, R. R.; Daniel, K.; Sung, S.; Woodcock, H. L.; Guida, W. C.; Brooks, W. H.: Virtual target screening: validation using kinase inhibitors. In: *J Chem Inf Model*, 2012, 52(8): 2192–2203.
- [175] Cortes, C.; Vapnik, V.: Support-vector networks. In: *Machine learning*, 1995, 20(3): 273–297.
- [176] R Core Team: *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2017.
- [177] Durham, T. B.; Klimkowski, V. J.; Rito, C. J.; Marimuthu, J.; Toth, J. L.; Liu, C.; Durbin, J. D.; Stout, S. L.; et al.: Identification of potent and selective hydantoin inhibitors of aggrecanase-1 and aggrecanase-2 that are efficacious in both chemical and surgical models of osteoarthritis. In: *J Med Chem*, 2014, 57(24): 10476–10485.
- [178] Topliss, J.: *Quantitative structure-activity relationships of drugs*. Bd. 19. Elsevier, 2012.
- [179] Meyer, D.; Dimitriadou, E.; Hornik, K.; Weingessel, A.; Leisch, F.: *e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien*, 2017. – R package version 1.6-8.

Tabellenverzeichnis

Tabelle 1.1	Mit KHK assoziierte GWAS-Genloci	4
Tabelle 1.2	Humane ADAMTS-Gene/Proteine sowie deren bekannte Substrate und Inhibitoren .	13
Tabelle 3.1	Humane ADAMTS-Röntgenkristallstrukturen	40
Tabelle 3.2	Ergebnisse der Ramachandran-Plots für die ADAMTS5-Röntgenkristallstrukturen 2RJQ und 3HYG	41
Tabelle 3.3	Sequenz- und Strukturvergleich der Bindetaschen der ADAMTS7-Modelle	44
Tabelle 3.4	Modellvalidierung der ADAMTS7-Modelle	46
Tabelle 3.5	Modellvalidierung der ADAMTS7-Modelle II	46
Tabelle 3.6	Strukturvergleich ($RMSD_{C\alpha}$) der S1'-Konformationen von ADAMTS7	52
Tabelle 4.1	Ergebnisse des Ramachandran-Plots für die ADAMTS1-Röntgenkristallstruktur 2V4B	56
Tabelle 4.2	Ergebnisse des Ramachandran-Plots für die ADAMTS4-Röntgenkristallstruktur 2RJP	56
Tabelle 4.3	ADAMTS-Proteine und genutzte Templates für das Homology Modeling	57
Tabelle 4.4	Modellvalidierung der erstellten ADAMTS-Modelle	58
Tabelle 4.5	Modellvalidierung der erstellten ADAMTS-Modelle II	59
Tabelle 4.6	Paarweiser Vergleich der Sequenzidentitäten der ADAMTS-Proteine	60
Tabelle 4.7	Paarweiser Vergleich der Sequenzidentitäten der ADAMTS-Metalloproteinase-Domänen	61
Tabelle 4.8	Strukturvergleich der ADAMTS-Metalloproteinase-Domänen	61
Tabelle 4.9	Paarweiser Vergleich der Sequenzidentitäten der ADAMTS-Bindetaschen mit ADAMTS7	64
Tabelle 4.10	Strukturvergleich der Bindetaschenregionen mit ADAMTS7	65
Tabelle 5.1	Ergebnis der <i>e-Pharmacophore</i> -Methode für ADAMTS7	72
Tabelle 6.1	Wasserstoffbrückenbindungen zwischen den 79 potenziellen ADAMTS7-Inhibitoren und der ADAMTS7-Bindetasche	86
Tabelle 6.2	Hydrophobe Wechselwirkungen zwischen den 79 potenziellen ADAMTS7-Inhibitoren und der ADAMTS7-Bindetasche	87

Tabelle 6.3	Aromatische Wechselwirkungen zwischen den 79 potenziellen ADAMTS7-Inhibitoren und der ADAMTS7-Bindetasche	87
Tabelle 6.4	Ausgewählte ADAMTS-Röntgenkristallstrukturen für das inverse Screening	88
Tabelle 6.5	Ähnlichkeitsmatrix der sieben potenziell nicht selektiven ADAMTS7-Hits	93
Tabelle 6.6	Ähnlichkeitsmatrix der 18 potenziellen ADAMTS7-Hits	96
Tabelle 6.7	Wasserstoffbrücken in Simulationen der ADAMTS7-Docking-Posen	112
Tabelle A.1	Paarweiser Vergleich der Sequenzidentitäten des ADAMTS-Bereichs um S4	128
Tabelle A.2	Paarweiser Vergleich der Sequenzidentitäten des ADAMTS-Bereichs um H4	128
Tabelle A.3	Paarweiser Vergleich der Sequenzidentitäten des L1-Loops	129
Tabelle A.4	Strukturvergleich der Region um S4	129
Tabelle A.5	Strukturvergleich der Region um H4	130
Tabelle A.6	Strukturvergleich der Region um L1	130
Tabelle A.7	Redockingergebnisse für ADAMTS-Protein-Inhibitor-Komplexe (numerisch)	133
Tabelle A.8	Strukturformelverzeichnis der besten 79 Verbindungen des ADAMTS7-Virtual Screenings	134
Tabelle A.9	Normalisierte Docking-Scores des inversen Screenings	142

Abbildungsverzeichnis

Soweit nicht anders beschrieben ist in den 3D-Abbildungen Sauerstoff in rot, Stickstoff in blau, Zink in grau und Wasserstoff in weiß abgebildet.

Die Cartoon-Darstellungen von Proteinen sind, sofern nicht anders beschrieben, folgendermaßen koloriert: α -Helices sind rot eingefärbt, β -Faltblätter gelb und Loops grün.

Abbildung 1.1	Stufen der Medikamentenentwicklung	1
Abbildung 1.2	Darstellung der Verkalkung der Arterien als Auslöser für KHK	2
Abbildung 1.3	Circosplot mit bekannten KHK-Risikoloci	4
Abbildung 1.4	Phylogenetischer Baum der humanen ADAMTS-Proteine	7
Abbildung 1.5	Domänenstrukturen von ADAM-, ADAMTS- und MMP-Proteinen	8
Abbildung 1.6	Domänenstruktur der ADAMTS-Proteine	10
Abbildung 1.7	Beispielhafte Darstellung von ADAMTS-Röntgenkristallstrukturen	11
Abbildung 1.8	Beispielhafte Darstellung von ADAMTS-Röntgenkristallstrukturen II	12
Abbildung 1.9	Phylogenetischer Baum der humanen ADAMTS-Proteine (Ergänzung)	13
Abbildung 2.1	Schematische Darstellung der wesentlichen Bestandteile eines Kraftfeldes	24
Abbildung 3.1	Zink-Histidin-Bedingung	43
Abbildung 3.2	Charakterisierung und Darstellung der ADAMTS7-Bindetasche	45
Abbildung 3.3	ADAMTS7-Homologiemodell SP1	47
Abbildung 3.4	Darstellung der Sekundärstrukturen im ADAMTS7-Homologiemodell und Stabilität dieser in MD-Simulation	48
Abbildung 3.5	Vergleich des S2'-Loops anhand von ADAMTS4-Röntgenkristallstrukturen	50
Abbildung 3.6	Vergleich des S1'-Loops anhand von ADAMTS5-Röntgenkristallstrukturen	51
Abbildung 3.7	S1'-Loop und S2'-Loop im ADAMTS7-Homologiemodell	52
Abbildung 3.8	Sampling des S1'-Loops von ADAMTS7	53
Abbildung 4.1	Einteilung der ADAMTS7-Bindetasche sowie Sequenzalignment der ADAMTS-Bindetaschen	63

Abbildung 4.2	Bindemuster der bekannten ADAMTS-Röntgenkristallstrukturen mit Inhibitor . . .	67
Abbildung 5.1	Ergebnis der <i>e-Pharmacophore</i> -Methode für ADAMTS7	73
Abbildung 5.2	SiteMap von ADAMTS7	74
Abbildung 5.3	Finale ADAMTS7-Pharmacophore	76
Abbildung 5.4	Redockingergebnisse für ADAMTS-Protein-Inhibitor-Komplexe	77
Abbildung 5.5	Virtual Screening-Route für ADAMTS7	79
Abbildung 5.6	Docking-Scores aller Hits, die die Zink-Koordinierung und Wasserstoffbrücke zu Gly358-O in ADAMTS7 aufweisen	82
Abbildung 6.1	Auswertung des inversen Screenings für die ausgewählten 79 Verbindungen des ADAMTS7-Virtual Screenings	91
Abbildung 6.2	Strukturformeln der sieben potenziell nicht selektiven ADAMTS7-Inhibitoren . .	93
Abbildung 6.3	Strukturformeln der 18 potenziell selektiven ADAMTS7-Inhibitoren	95
Abbildung 6.4	Längen- und Volumenvergleich der potenziell nicht selektiven und selektiven Hits	97
Abbildung 6.5	Docking-Posen der sieben potenziell nicht selektiven ADAMTS7-Inhibitoren . .	98
Abbildung 6.6	Docking-Posen von ZINC37707318, ZINC52329369, ZINC75953566 und ZINC31961806	101
Abbildung 6.7	Docking-Posen von ZINC48847183, ZINC52012291, ZINC40603395, ZINC48847273 und ZINC52012042	103
Abbildung 6.8	Docking-Posen von ZINC40470996 und ZINC49571688	104
Abbildung 6.9	Docking-Posen von ZINC80583404, ZINC40558343, ZINC94569517 und ZINC94569992	106
Abbildung 6.10	Docking-Posen von ZINC49587694, ZINC70254898 und ZINC86853423	107
Abbildung 6.11	RMSD der Simulationen der ADAMTS7-Docking-Posen	109
Abbildung 7.1	Wechselwirkungen und Eigenschaften potenzieller ADAMTS7-Inhibitoren	120
Abbildung A.1	Multiples Alignment von der ADAMTS7-Sequenz und den bekannten ADAMTS-Röntgenkristallstrukturen	123
Abbildung A.2	Multiples Alignment von der ADAMTS7-Sequenz und den ADAMTS5-Template-Strukturen	125
Abbildung A.3	Multiples Sequenzalignment der Metalloproteinase- und Disintegrin-ähnlichen Domänen der ADAMTS-Proteine	125
Abbildung A.4	RMSD _{Cα} über MD-Simulation des ADAMTS7-Homologiemodells	126

Abbildung A.5	Multiples Sequenzalignment der Metalloproteinase-Domänen der ADAMTS-Proteine	127
Abbildung A.6	Oberflächen der Bindetaschen der bekannten ADAMTS-Röntgenkristallstrukturen mit Inhibitor	131
Abbildung A.7	Docking-Posen der sieben potenziell nicht selektiven ADAMTS7-Inhibitoren in der ADAMTS7-Bindetasche (Oberflächendarstellung)	148
Abbildung A.8	Docking-Posen der 18 potenziell selektiven ADAMTS7-Inhibitoren in der ADAMTS7-Bindetasche (Oberflächendarstellung)	149

Formelverzeichnis

Formel (2.1)	Potentielle Energie E_{pot} eines Moleküls im Beispielfeld	25
Formel (2.2)	Differentialgleichung (MD-Simulation)	27
Formel (2.3)	Mittlere quadratische Abweichung (RMSD)	27
Formel (2.4)	Mittlere quadratische Schwankung (RMSF)	28
Formel (2.5)	Scoring-Funktion von AutoDock4	32
Formel (2.6)	Scoring-Funktion von AutoDockVina	32
Formel (2.7)	Scoring-Funktion PLANTS _{PLP}	33
Formel (2.8)	Scoring-Funktion PLANTS _{CHEMPLP}	33
Formel (2.9)	Scoring-Funktion von Glide (SP- und HTVS-Docking)	34
Formel (2.10)	Scoring-Funktion von Glide (XP-Docking)	34
Formel (2.11)	Mittelwert der Docking-Scores für i -ten Liganden (μ_i)	35
Formel (2.12)	Varianz der Docking-Scores für i -ten Liganden (σ_i^2)	35
Formel (2.13)	Normalisierter Docking-Score nach der MACS-Methode ($Score'_{ij}$)	35
Formel (2.14)	Tanimoto-Koeffizient (τ)	37

Publikationen

In diesem Anhang sind meine veröffentlichten wissenschaftlichen Beiträge aufgelistet. Die Arbeiten, welche sich auf diese Thesis beziehen, sind fettgedruckt hervorgehoben.

Publikationen in wissenschaftlichen Journalen

Eva Koschmidder, Friederike Borngräber, Alexander Schmidt, Aleksandar Rakovic, Ingrid Brænne, Anne Weissbach, Leopold Größer, Susen Schaake, Michaela Müller, Humera Manzoor, Hans-Christian Jabusch, Daniel Alvarez-Fischer, Meike Kasten, Vladimir Kostic, Thomas Gasser, Kirsten Zeuner, Eckart Altenmüller, Christine Klein, Katja Lohmann: Functional characterization of rare RAB12 variants and their role in musician's dystonia. *Genes* (submitted).

Michaela Müller, Thorsten Kessler, Heribert Schunkert, Jeanette Erdmann, Stephanie Tennstedt: Classification of ADAMTS binding sites: The first step toward selective ADAMTS7 inhibitors. In: *Biochemical and biophysical research communications*, 2016, 471(3): 380–385.

Vorträge

Michaela Müller, Thorsten Kessler, Heribert Schunkert, Jeanette Erdmann, Stephanie Tennstedt: Classification of ADAMTS binding sites: The first step toward selective ADAMTS7 inhibitors. EFMC-ISMC 2016, 2016, Manchester, UK.

Michaela Müller: *In silico Targeting für potentielle Wirkstoffe*. btS Lübeck: Vortragsreihe „Exploring the Code – New Perspectives of Genetics“, 2014, Lübeck, Deutschland.

Poster

Michaela Müller, Stephanie Tennstedt, Björn Windshügel, Philip Gribbon, Simon von Ameln, Thorsten Kessler, Heribert Schunkert, Jeanette Erdmann: *Virtual Screening: Dem Arzneimittel auf der Spur*. Uni im Dialog - Doktorandentag, 2016, Lübeck, Deutschland.

Michaela Müller, Stephanie Tennstedt, Björn Windshügel, Philip Gribbon, Jeanette Erdmann: Virtual Screening for ADAMTS7 Inhibitors. Young-DZHK-Retreat, 2014, Braunlage, Deutschland.

Michaela Müller, Stephanie Tennstedt, Zouhair Aherrahrou, Ann-Kathrin Sowa, Frank Kaiser, Andreas Dendorfer, Jeanette Erdmann, Heribert Schunkert: Homology model of Abcc6 provides insight into the function of mutations causing cardiovascular phenotype. CECAM workshop „Long time dynamics from short time simulations“, 2014, Lugano, Schweiz.

Michaela Müller, Stephanie Tennstedt, Zouhair Aherrahrou, Ann-Kathrin Sowa, Frank Kaiser, Andreas Dendorfer, Jeanette Erdmann, Heribert Schunkert: 3D-Modellierung: Der Krankheit auf der Spur. Uni im Dialog - Doktorandentag, 2013, Lübeck, Deutschland.