

Strukturelle und thermodynamische Charakterisierung der Fragment und Liganden Bindung an Thrombin

Dissertation

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

dem
Fachbereich Pharmazie
der Philipps-Universität Marburg
vorgelegt

von
Eggert Henning Rühmann

aus
Kiel

Marburg/Lahn 2015

Dem Fachbereich Pharmazie

der Philipps-Universität Marburg als Dissertation eingereicht am: 03.11.2015

Tag der mündlichen Prüfung am: 16.12.2015

Erstgutachter: Prof. Dr. Gerhard Klebe

Zweitgutachter: Prof. Dr. Andreas Heine

Hochschulkennziffer: 1180

Die Untersuchungen zum vorliegenden Thema wurden auf Anregung von Prof. Dr. Gerhard Klebe am Institut für Pharmazeutische Chemie des Fachbereichs Pharmazie der Philipps-Universität Marburg in der Zeit von Juni 2011 bis Mai 2015 durchgeführt.

Inhaltsverzeichnis/Table of Contents

Inhaltsverzeichnis/Table of Contents	I-IV
Aus dieser Dissertation hervorgegangene Publikationen	V
1 Einleitung	1
1.1 Strategien in der Wirkstoffentwicklung	1
1.2 Thrombin als Zielprotein.....	2
1.3 Isothermale Titrationskalorimetrie.....	8
1.4 Ziel der Arbeit und Fragestellungen	11
1.5 Referenzen	14
2 Thermodynamic Signatures of Fragment Binding: Validation of Direct versus Displacement ITC Titrations.....	16
2.1 Einleitende Bemerkungen/Introductory Remarks.....	16
2.2 Abstract.....	16
2.3 Introduction	17
2.4 Results.....	19
2.4.1 Data set of fragments and reference ligands.....	19
2.4.2 Importance of the c-value for the shape of the titration curve.....	24
2.4.3 Estimation of fragment affinity	26
2.4.4 Results obtained by the displacement titrations	27
2.4.5 Results obtained by direct low c-value titrations.....	29
2.5 Discussion.....	32
2.6 Conclusions	37
2.7 Experimental Section / Materials & Methods	39
2.7.1 ITC.....	39
2.7.2 Bioassay.....	42
2.8 References.....	43

3	Fragment Binding Can Be Either More Enthalpy-Driven or Entropy-Driven: Crystal Structures and Residual Hydration Patterns	
	Suggest Why.	46
3.1	Einleitende Bemerkung/Introductory remarks	46
3.2	Abstract	46
3.3	Introduction	47
3.4	Results	49
3.4.1	Crystal structures of the reference ligands	49
3.4.2	Crystal structures of the fragments	52
3.4.3	Influence of bound hirudin peptide on the thermodynamic signature	56
3.4.4	Analysis of the conformational flexibility of Glu192	57
3.5	Discussion	60
3.5.1	Methoxy fragment 1	62
3.5.2	Amidine fragments 3 and 4	63
3.5.3	Chlorine-substituted fragments 2, 5 and 6	64
3.5.4	Deviating thermodynamic profiles obtained using different reference ligands	65
3.6	Conclusion	67
3.7	Experimental Section / Materials & Methods	69
3.7.1	ITC measurements	69
3.7.2	X-ray Data	69
3.8	References	74
4	Boosting affinity by correct ligand preorganization for the S2 pocket of thrombin: A study by ITC, MD and high resolution crystal structures	77
4.1	Einleitende Bemerkungen/Introductory Remarks	77
4.2	Abstract	77
4.3	Introduction	78
4.4	Results	80
4.4.1	Crystal structures	80
4.4.2	Surface area and volume analysis	84
4.4.3	ITC results	85
4.4.4	MD simulation	88

4.5	Discussion.....	91
4.6	Conclusions	96
4.7	Experimental Section / Materials & Methods	97
4.7.1	Synthesis.....	97
4.7.2	ITC.....	97
4.7.3	X-ray Data	98
4.7.4	MD Simulation.....	102
4.7.5	Volume Calculation	103
4.8	Biochemical Assays.....	103
4.9	Accession codes.....	103
4.10	Acknowledgement.....	103
4.11	References.....	105
5	Kristallographische und thermodynamische Untersuchungen von MI067 im Vergleich zu MI001 und MI002.....	108
5.1	Einleitung	108
5.2	Kristallstrukturen	110
5.3	Thermodynamik und Affinitätsdaten.....	115
5.4	Diskussion und Zusammenfassung der Daten	118
5.5	Methoden	119
5.5.1	Assays	119
5.5.2	ITC.....	119
5.5.3	Röntgenstruktur	120
5.6	Referenzen	124
6	Röntgenstruktur des Inhibitors MI432 in Thrombin zur Modellierung eines Matriptase-Komplexes	125
6.1	Einleitende Bemerkungen/Introductory Remarks.....	125
6.2	Einleitung	125
6.3	Röntgenstruktur	127
6.3.1	Interaktion mit Thrombin.....	128
6.3.2	Interaktion mit Matriptase.....	130

6.4	Zusammenfassung	132
6.5	Material & Methoden	133
6.5.1	Assay.....	133
6.5.2	Kristallisation	133
6.5.3	Datensammlung und Prozessierung.....	134
6.5.4	Strukturbestimmung und Verfeinerung.....	134
6.6	Referenzen.....	137
7	Zusammenfassung/Summary	139
7.1	Deutsche Fassung	139
7.2	English Version.....	144
8	Anhang/Annex	149
8.1	Abkürzungen/Abbreviations	149
8.2	Beiträge auf Konferenzen/Conference contributions	151
8.3	Danksagung/Acknowledgments.....	152
8.4	Lebenslauf/Curriculum Vitae	154

Aus dieser Dissertation hervorgegangene Publikationen

Teile dieser Arbeit sind bereits in wissenschaftlichen Journalen publiziert worden oder für die Einreichung vorbereitet. Jedem dieser Kapitel ist eine einleitende Bemerkung vorangestellt, die die Beiträge des Autors dieser Dissertation zu diesen Publikationen hervorhebt.

Thermodynamic signatures of fragment binding: Validation of direct versus displacement ITC titrations.

Rühmann E., Betz M., Fricke M., Heine A., Schäfer M., Klebe G.

Biochim Biophys Acta. 2015 Apr.; 1850 (4):647-656.

doi: 10.1016/j.bbagen.2014.12.007.

Fragment Binding Can Be Either More Enthalpy-Driven or Entropy-Driven: Crystal Structures and Residual Hydration Patterns Suggest Why.

Rühmann E., Betz M., Heine A., Klebe G.

J Med Chem. 2015 Sep. 10; 58(17):6960-6971.

doi: 10.1021/acs.jmedchem.5b00812.

Boosting affinity by correct ligand preorganization for the S2 pocket of thrombin: A study by ITC, MD and high resolution crystal structures

Rühmann E., Rupp M., Betz M., Heine A., Klebe G.

Manuskript in Vorbereitung

New 3-amidinophenylalanine-derived inhibitors of matriptase

Hammami M., Rühmann E., Maurer E., Heine A., Gütschow M., Klebe G, Steinmetzer T.

Med. Chem. Commun., 2012, 3, 807-813.

doi: 10.1039/C2MD20074K

1 Einleitung

1.1 Strategien in der Wirkstoffentwicklung

Zunächst soll ein kurzer Überblick über zwei verschiedene populäre Ansätze in der industriellen Wirkstoffsuche und Entwicklung gegeben werden.

Die Entdeckung neuer Moleküle zur Modulierung biologischer Funktionen ist ein Hauptaspekt in der Wirkstoffforschung. [1] In den vergangenen zwei Jahrzehnten hat sich die pharmazeutische Industrie sehr auf Hochdurchsatz-*Screening* (HTS) [2] zum Durchsuchen großer Bibliotheken von chemischen Verbindungen (bis zu 10^6) konzentriert, die über die Zeit erstellt worden sind. Die Aufrechterhaltung der Vielfalt und Qualität dieser Bibliotheken ist allerdings sehr aufwändig und erfordert eine ständige Pflege.

Viele Moleküle in diesen Bibliotheken sind nicht sehr „wirkstoffartig“, was bedeutet, dass sie oft verhältnismäßig lipophil sind und somit unter physiologischen Bedingungen eine schlechte Löslichkeit und Bioverfügbarkeit besitzen. Als Folge treten auch oft falsch positive *Hits* auf, die sich nicht zur Weiterentwicklung zu einem Arzneistoff eignen.

Es ist auch unvermeidlich, dass diese HTS-Bibliotheken nur einen Bruchteil der möglichen Verbindungen des „chemischen Raums“ (ca. 10^{60}) abdecken und somit sich nur eine sehr begrenzte Anzahl an Molekülen als Ausgangspunkt für die weitere Entwicklung eignet.

Ein weiterer Nachteil ist schließlich, dass auch wenn die Struktur eines HTS-*hit* an das Zielprotein bindet, es noch nicht klar ist, welche Teile des Moleküls am meisten zur Bindungsenergie beitragen. Dies ist aber unerlässlich um später die Potenz des *Hits* gezielt steigern zu können.

Ein weiterer Ansatz der inzwischen oft verfolgt wird ist das Fragment-basierte Wirkstoffdesign. Hierfür werden schwach bindende kleine Moleküle mit Molekülmassen von in der Regel <300 Da verwendet, welche sich als alternativen Ansatzpunkt im frühen Stadium der Medikamentenentwicklung eignen.

Der Ansatz ist hierbei grundlegend unterschiedlich zum HTS, da sich die Bibliotheksgröße, die Methoden der Screening Verfahren und die Abhängigkeit von Strukturmethoden (NMR und Kristallstrukturanalyse) stark unterscheiden.

Das sogenannte *Fragment Based Drug Discovery* (FBDD), setzt auf kleine Bibliotheken und sensitive *Screening*-Methoden wie NMR (*Nuclear Magnetic Resonance*) [3], TSA (*Thermal Shift Assay*) [4], Röntgenkristallographie, SPR (*Surface Plasmon Resonance*) [5,6] und ITC (*Isothermal Titration Calorimetry*) [7]. Die Verwendung von Röntgenkristallographie war und ist, eine der wichtigsten und sensitivsten Screening-Plattformen, wenn die Kristallisierbarkeit des Zielproteins diesen Ansatz erlaubt. [8]

Auch schwach bindende Fragmente, bilden hochwertige Wechselwirkungen mit dem Zielprotein aus und können schließlich zu größeren hochwirksamen Kandidaten optimiert werden, in dem weitere Reste und funktionelle Gruppen angefügt werden. [9] Zudem ist es auch möglich mehrere Fragmente mit niedrigen molekularen Massen, zu größeren potenteren Liganden zu verbinden. [10]

Um nach der Detektion von Fragment-Kandidaten einen größeren potenten Wirkstoff herzustellen, ist somit besonders ein gutes Zusammenspiel der Strukturbiologie und der Synthesechemie entscheidend. Bedingt durch die begrenzte Bibliotheksgröße eignet sich FBDD auch für Forschungsprojekte im universitären Maßstab.

1.2 Thrombin als Zielprotein

Für viele Krankheitsbilder, beispielsweise nach Herzinfarkten, zur Schlaganfallprävention, tiefen Venenthrombosen oder auch nach einer Operation, kann es notwendig sein, die Fähigkeit des Bluts zur Gerinnung zu reduzieren. Aus diesem Grund besteht großes Interesse an der Entwicklung selektiver und vor allem oral verfügbarer Hemmstoffe der Blutgerinnungskaskade. Die Protease Thrombin spielt hierbei eine zentrale Rolle bei der Kontrolle der Blutgerinnung.

Thrombin, auch Blutgerinnungsfaktor IIa genannt, ist eine Serinprotease, die beim Menschen durch das F2 -Gen kodiert wird. [11]

Thrombin wird aus Prothrombin (Gerinnungsfaktor II) proteolytisch durch aktivierten Faktor X (Xa) erzeugt, um aktives Thrombin (Gerinnungsfaktor IIa) in der Blutgerinnungskaskade zu bilden. Prothrombin wird zuvor in der Leber hergestellt und posttranslational in einer Vitamin K-abhängigen Reaktion modifiziert. Mangel an Vitamin K oder Vitamin-K-Antagonisten (z.B. Phenprocoumon) führen zu einer Verlangsamung der Aktivierung der Gerinnungskaskade.

Faktor X kann sowohl durch den intrinsischen Pfad (reduzierter Blutfluss oder pathogen veränderte Gefäßwände), als auch den extrinsischen Pfad (Verletzung des Gefäßsystems) aktiviert werden. Die Aktivität von Faktor Xa wird hierbei stark durch die Bindung an den aktivierten Faktor V (Va) verbessert. Dieses Zusammenspiel wird als Prothrombinase-Komplex bezeichnet.

Thrombin katalysiert anschließend die Umsetzung von löslichem Fibrinogen zu unlöslichen Fibrinfäden, sowie andere blutgerinnungsbedingte Reaktionen. Diese Kaskade führt schließlich zu einer Verringerung eines Blutverlustes. Zusammen mit aggregierten Blutplättchen bildet sich ein Polymergeflecht, in dem sich verschiedene Blutkörperchen ansammeln. Es entsteht ein Thrombus aus Fibrin, der anschließend durch Faktor XIII (Transglutaminase) weiter vernetzt und stabilisiert wird. Mit diesem physiologisch sehr wichtigen Mechanismus des Körpers wird ein Wundverschluss gewährleistet.

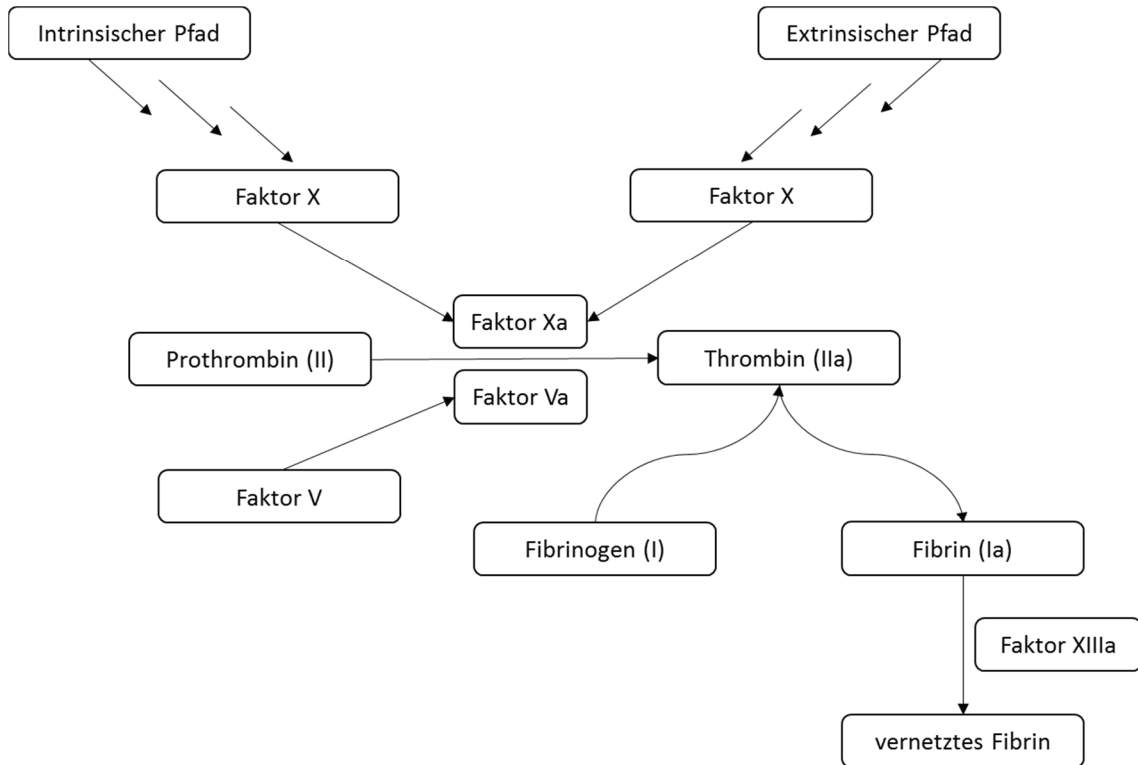


Abbildung 1.1 Gezeigt ist ein vereinfachtes Schema der Blutgerinnungskaskade zur Veranschaulichung der zentralen Rolle von Thrombin in diesem physiologisch wichtigen Prozess.

Strukturell zeichnen sich Serinproteasen wie Thrombin durch gut definierte Spezifitätstaschen aus, die um die katalytische Triade herum zur Substratanlagerung lokalisiert sind.

Die so genannte katalytische Triade wird von den Aminosäuren Serin, Histidin und Aspartat gebildet. (**Abbildung 1.2**)

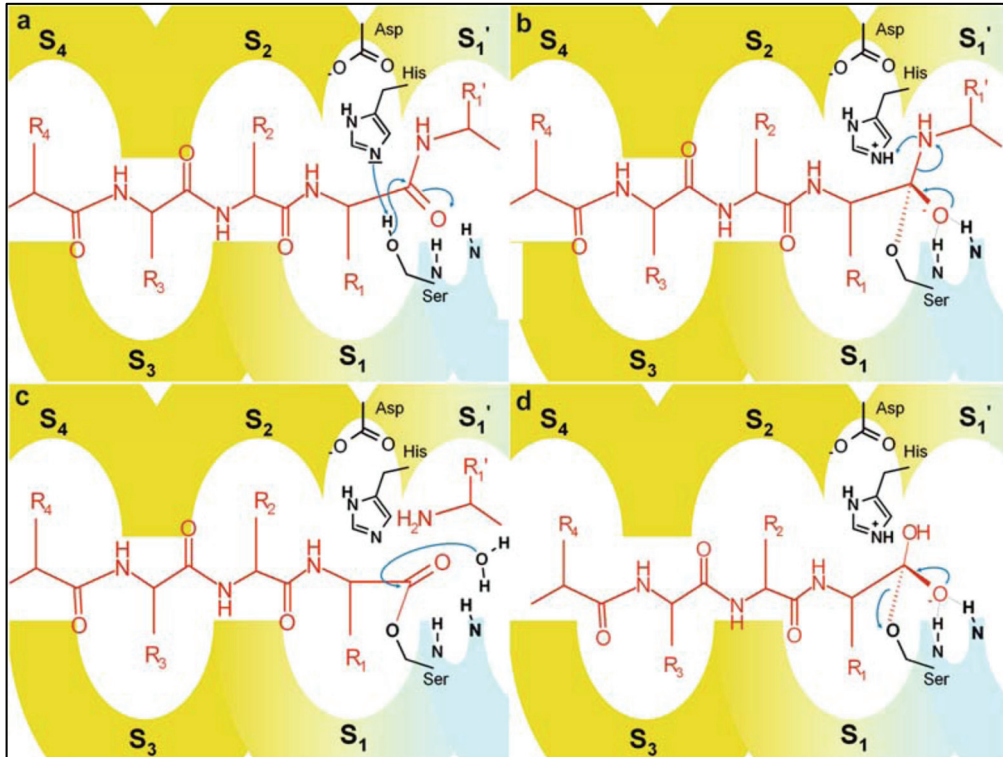


Abbildung 1.2 Mechanismus von Serinproteasen. [12]

Die durch Histidin 57 aktivierte (deprotonierte) Hydroxylgruppe des katalytischen Serin an der Position 195 führt einen nukleophilen Angriff auf die zu spaltende Amidbindung durch. Das zum Histidin 57 benachbarte Aspartat kann ein Proton des Histidinimidazolrings aufnehmen und später wieder abgeben. Es kompensiert somit die positive Ladung, die auf dem Histidinrest entsteht.

Eine weitere Stabilisierung des Übergangszustands findet in Serinproteasen noch über ein zusätzliches charakteristisches Strukturmotiv, das so genannte *oxyanion hole* statt.

Der gebildete Übergangszustand zerfällt unter Freisetzung des C-terminalen Spaltprodukts, das eine freie Aminogruppe trägt. Das N-terminale Spaltprodukt (zunächst noch mit der Protease als kovalent verknüpftes Acylenzym-Zwischenprodukt verbunden) wird in einem nachfolgenden Schritt durch den nucleophilen Angriff eines Wassermoleküls abgespalten. Anschließend steht das Enzym dann für die nächste Umsetzung bereit.

Ein wichtiges Werkzeug um strukturelle Informationen zu Protein-Ligand-Wechselwirkungen zu erhalten ist die Proteinkristallographie.

Mit Proteinkristallographie ist es möglich, die molekulare Struktur von Proteine im Detail zu analysieren, wenn es möglich ist sie in einen einkristallinen Zustand zu überführen.

Die Kenntnis der räumlichen Struktur eines Proteinmoleküls erlaubt die Entwicklung von Modellen zu Struktur-Wirkungsbeziehungen von Proteinen und Liganden. Die Proteinstrukturanalyse kann in drei Schritte eingeteilt werden: Kristallisation von Proteinen, Streuexperiment am Kristall, Phasenbestimmung und Verfeinerung des Modells der Proteinstruktur.

Ein auf diese Weise erstelltes Modell ist in **Abbildung 1.3** gezeigt, um die räumliche Anordnung der Thrombinbindetaschen zu verdeutlichen.

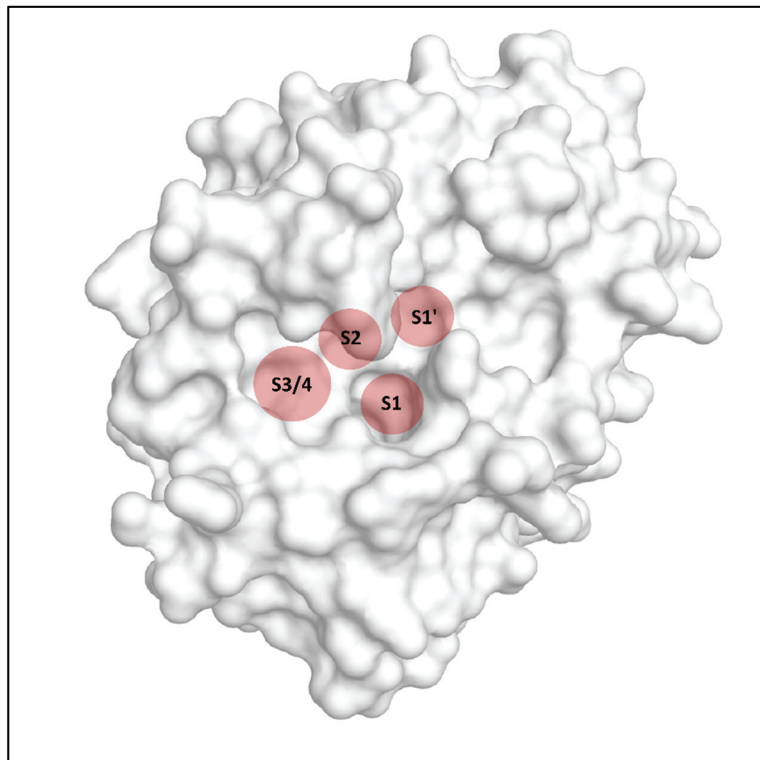


Abbildung 1.3 Thrombin mit seiner Lösungsmittel zugänglichen Oberfläche ist in hellgrau dargestellt. Die Spezifitätstaschen sind durch rote Kreise hervorgehoben.

Die S1-Spezifitätstasche besitzt negative elektrostatische Eigenschaften, hervorgerufen durch die Aminosäure Aspartat 189 am Boden der Tasche, für die viele bekannte Ankerfragmente (Amidine, Guanidine) beschrieben sind.

Die hydrophobe S2 Tasche wird nach oben hin durch den 60iger-*loop* wie ein Deckel abgeschirmt und sorgt für die nötige Selektivität gegenüber anderen Substraten mit größeren Aminosäuren in P2 Position. Diese Tasche wird optimal mit kleinen hydrophoben Aminosäuren wie Valin und Prolin adressiert.

In der weiten und offenen hydrophoben S3/S4-Tasche bindet sowohl beim natürlichen Substrat Fibrinogen als auch häufig bei Thrombininhibitoren die Aminosäure Phenylalanin. Das Erkennungsmuster des natürlichen Substrates (Val-Gly-Arg) wird häufig von peptidomimetischen Inhibitoren nachgeahmt.

Aus den zuvor genannten Gründen ist Thrombin sowohl ein relevantes Ziel für Arzneistoffe als auch in vielerlei Hinsicht ein sehr gut geeignetes Modellprotein für die systematische Untersuchung der Grundlagen der Protein-Ligandbindung.

Die Bindetaschen von Thrombin sind sehr gut definiert und das Protein ist insgesamt sehr rigide, sodass kaum Konformationsänderungen in Folge von Inhibitorbindungen zu erwarten sind. Darüber hinaus erfolgt die Bindung von Liganden weitgehend unabhängig von Ionen oder Kofaktoren.

Eine sehr große Zahl von Inhibitoren ist inzwischen bekannt, so dass relativ einfach systematische Variationen in den Inhibitorgrundgerüsten möglich sind.

Die Methoden sowohl für die kinetisch-photometrische, als auch Fluoreszenz basierte *Assays*, die Kristallisation und Messprotokolle für die ITC sind in der Arbeitsgruppe gut etabliert. Zudem lassen sich durch die gute Kristallisierbarkeit des Proteins und anschließend optimierten Kristallisationsbedingungen Thrombinkristallstrukturen inzwischen mit sehr hoher Auflösung bestimmen, was ebenfalls für das Verständnis der Protein-Ligandbindung und die Interpretation thermodynamischer Daten und Solvatationszustände eine große Rolle spielt.

1.3 Isothermale Titrationskalorimetrie

Die Berücksichtigung von thermodynamischen Bindungsprofilen ist sehr relevant in der *Hit*- und *Lead*-Optimierung von Wirkstoffkandidaten und somit ein wertvolles analytisches Werkzeug.

Eine entscheidende Rolle kommt der ITC bei der thermodynamischen Charakterisierung von Protein-Ligand-Interaktionen zu.

Bei der ITC wird tropfenweise der gelöste Ligand in eine Lösung des Proteins gegeben. Dabei befinden sich üblicherweise der Ligand in einer Spritze und die Proteinlösung in der thermisch isolierten Messzelle. Neben dieser Messzelle besitzen ITC-Geräte eine Referenzzelle mit der kontinuierlich die freiwerdende oder konsumierte Wärmeenergie in der Messzelle verglichen wird. (**Abbildung 1.4**)

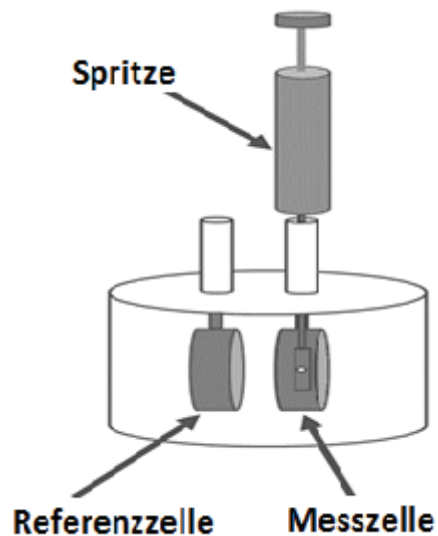


Abbildung 1.4 Schematische Anordnung von Mess-, Referenzzelle und Injektionsspritze eines ITC-Systems.

Die Bindung des Liganden an das Protein führt zu einer exothermen bzw. endothermen Reaktion. Sollte dies nicht der Fall sein, etwa bei einer entropischen Bindung, ist es mit dieser Messmethode nicht möglich Messdaten zu erhalten.

Mit jedem eingebrachten Tropfen deutet sich die geflossene Wärmemenge als Fläche unter den einzelnen *Peaks* an (**Abbildung 1.5**).

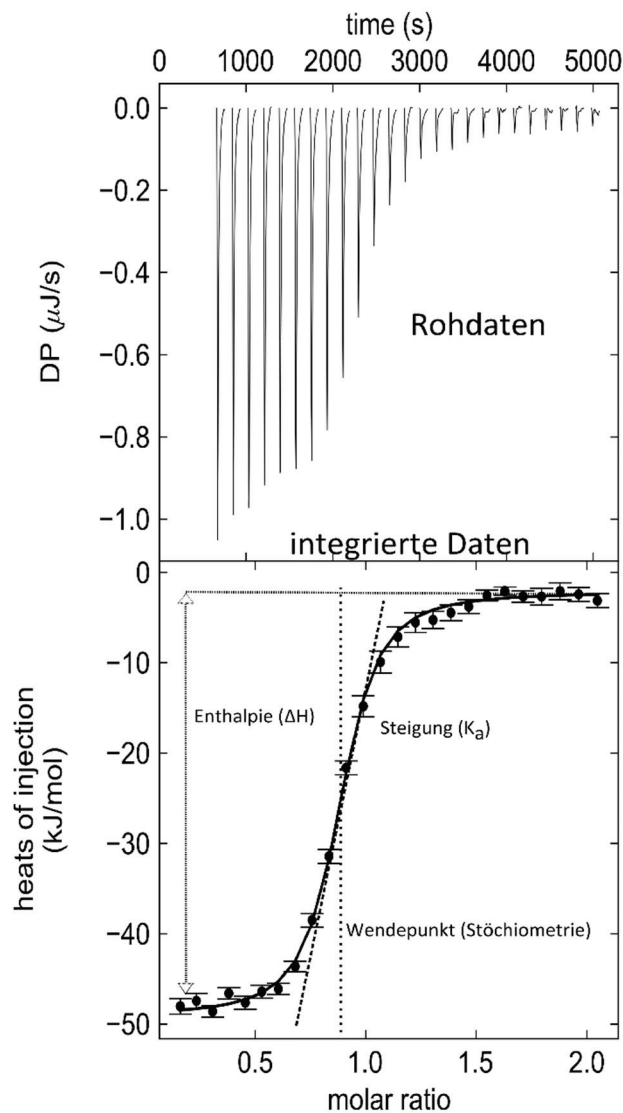


Abbildung 1.5 Gezeigt ist ein typisches ITC Thermogram. Im oberen Teil sind die Rohdaten einer exothermen ITC Messung zu sehen, wobei die abgegebene Wärmeenergie gegen die Zeit aufgetragen ist. Im unteren Teil sind die integrierten Wärmemengen der Injektionen gezeigt, die gegen das molare Verhältnis der Protein-Ligand- Bindung aufgetragen sind.

Die Summe der Wärmemengen der Reaktion, welche durch Integration der Peakflächen erhalten wird, ergibt die bei der Bindung übertragene Bindungsenthalpie ΔH . Mit zunehmender Zugabe des Liganden wird das Protein gesättigt, so dass die Wärmesignale in ihrer Intensität abnehmen. Die Bindungskonstante (Assoziationskonstante) kann aus der Steigung des Kurvenverlaufs errechnet werden. Über die Gleichung (1.1) kann die Freie Bindungsenthalpie ΔG der Reaktion ermittelt werden:

$$\Delta G^0 = -RT \ln K_a \quad (1.1)$$

Ein weiterer Wert der erhalten werden kann, ist das molare Verhältnis der Reaktion, das auch als Stöchiometrie der Umsetzung bezeichnet wird und sich aus dem Wendepunkt der Isotherme ergibt.

Die Entropie der Bindung kann schließlich aus der Gleichung **(1.2)** errechnet werden:

$$\Delta G^0 = \Delta H^0 - T\Delta S^0 \quad (1.2)$$

Die ITC stellt somit das wichtigste Werkzeug zur direkten Messung von thermodynamischen Parametern für Protein-Ligandinteraktionen dar.

Trotz dieser Vorteile besitzt die Methode auch einige Nachteile. So ist der Bedarf an Protein und Zeitaufwand für Messungen im Vergleich zu anderen Messmethoden hoch. Einen weiteren Nachteil stellt der verhältnismäßig schmale Messbereich (Affinitäten von niedrig zweistellig nM bis etwa zweistellig μ M bei der Vermessung von Proteinen) der Methode dar. Aus diesen Gründen besteht Optimierungsbedarf, um insbesondere schwach bindende Liganden wie Fragmente zuverlässig thermodynamisch untersuchen zu können.

1.4 Ziel der Arbeit und Fragestellungen

1.

Eine detaillierte Charakterisierung der thermodynamischen Signatur von schwach bindenden Fragmenten an Proteine ist wichtig, um den Entscheidungsprozess zu unterstützen, welche Fragmente im Zuge der *Hit-to-Lead*-Optimierung weiterverwendet werden sollen.

ITC ist das Verfahren der Wahl, um thermodynamische Daten aufzuzeichnen. Jedoch erfordern schwach bindende Liganden, wie Fragmente, die Entwicklung sinnvoller und sicherer Messprotokolle, da es auf Grund der begrenzten Löslichkeit kaum möglich ist sigmoidale Titrationskurven aufzeichnen.

Fragmente können entweder direkt unter niedrigen c-Wert Bedingungen (kein sigmoidaler Kurvenverlauf) oder indirekt durch Verwendung eines starken Bindungsliganden, welcher das vorinkubierte schwache Fragment während der Titration von dem Protein verdrängt, vermessen werden. Die Bestimmung der Gibbs-Energie ist zuverlässig und unabhängig vom angewandten Titrationsprotokoll.

Für diese Studie wird ein Satz von vier Fragmenten mit einstellig milli bis dreistellig mikromolaren Affinitäten zur Serin-Protease Thrombin ausgewählt, um anschließend direkte ITC-Titrations gegen Verdrängungsexperimente zu validieren.

Für die erforderlichen Verdrängungsexperimente werden drei starke Binder (A-C) mit dreistellig nanomolarer Affinität als Referenzliganden ausgewählt. Diese zeigen stark abweichende strukturelle und thermodynamische Profile, um den Einfluss dieser Unterschiede auf die Ergebnisse der Verdrängungstitrationen zu untersuchen.

In der Studie soll gezeigt werden, ob die mit unterschiedlichen Titrationsprotokollen gemessenen thermodynamischen Bindungsprofile für die verschiedenen Fragmente Unterschiede zeigen und ob sie unabhängig von den gewählten Verdrängungsliganden sind.

2.

Von den für den ersten Teil der Arbeit untersuchten Liganden und Fragmenten sollen möglichst hochaufgelöste Röntgenstrukturen erstellt werden, um einen detaillierten Einblick in den Bindungsmodus der Fragmente und deren Solvatzustand im gebundenen Zustand zu erhalten. Zusätzlich sollen zwei weitere Kristallstrukturen von den strukturell verwandten Fragmenten 5 und 6 erstellt werden, die für die thermodynamischen Untersuchungen eine unzureichende Affinität aufwiesen. Die im ersten Teil der Arbeit gewonnenen thermodynamischen Daten sollen anschließend verwendet werden, um in Verbindung mit den Röntgenstrukturdaten Rückschlüsse auf die Unterschiede in den thermodynamischen Profilen der Fragmente zuzulassen und für ein besseres Verständnis der Protein-Fragment-Interaktionen zu sorgen.

Insgesamt werden hochaufgelöste Kristallstrukturen von sechs Fragmenten (1-6) mit Bindung an die S1-Tasche der Serinprotease Thrombin und zwei der drei Verdrängungsliganden (A und C) bestimmt und in Bezug auf ihr thermodynamisches Bindungsprofil analysiert.

Die Seitenkette von Glutamat 192 oberhalb der S1-Tasche spielt bei der Bindung der Fragmente und der entstehenden Wassernetzwerke eine wichtige Rolle und wird in diesem Kapitel ebenfalls untersucht. Hierzu wird eine Konformationsanalyse dieser Seitenkette über alle in der Protein Data Bank (PDB) deponierten hochaufgelösten ($< 2 \text{ \AA}$) Thrombinstrukturen durchgeführt.

Ein weiterer Aspekt der in der Studie untersucht wird ist der Einfluss von einem an das Protein gebundenen Hirudinfragment auf die Thermodynamik von in die Bindetasche bindenden Inhibitoren. Dies ist relevant, da das Hirudinfragment bisher ausschließlich nur für die Kristallisation von Thrombin verwendet wurde, um eine Kristallpackung zu erhalten, die eine für *Soaking*-Experimente zugängliche Proteinbindetasche bietet. Bei der Interpretation von thermodynamischen Daten mit Hilfe von Röntgenstrukturen wurde dieser experimentelle Unterschied bislang nicht berücksichtigt.

3.

Das Ziel einer dritten Studie ist es, den Einfluss auf die thermodynamischen Eigenschaften von Inhibitoren mit abweichender Rigidität ihres chemischen Grundgerüsts in Bezug auf die Bindung an Thrombin zu untersuchen.

Hierfür wählten wir eine Reihe von fünf potenten peptidomimetischen Thrombininhibitoren aus, die unterschiedliche sterische Einschränkungen in der P2 Position aufweisen. Wir verwenden für diese Studie ITC und Röntgenkristallographie kombiniert mit MD-Simulationen. Die Inhibitoren variieren systematisch an der Position P2 durch die Verwendung eines Glycin-, Alanin-, N-Methyl-Glycin-, N-Methyl-*homo*-Alanin- und Prolin- Restes und adressieren die hydrophobe S2-Tasche des *apo*-Enzyms.

In Bezug auf die verbleibenden Taschen der Protease weisen alle Inhibitoren die gleichen Substituenten auf und zeigen praktisch die gleichen Bindungsmodi.

Die untersuchten Inhibitoren weisen dabei eine stark abweichende Affinität in ihrer Bindungsaffinität auf (von dreistellig pM bis dreistellig nM).

Insbesondere Wasserstrukturen, Restmobilität und Besetzung der S2-Tasche durch die gebundenen Liganden sind wichtig, um die thermodynamischen Daten, die wir durch ITC erhalten, zu interpretieren.

Für die fünf untersuchten Liganden werden daher hochaufgelöste Röntgenstrukturen erstellt.

Da die ITC Messungen des verwendeten Inhibitorgrundgerüsts eine Pufferabhängigkeit aufweisen, ist es nötig die ITC Messungen in mindestens drei Puffern mit unterschiedlicher Ionisationsenthalpie durchzuführen, um die erhaltenen thermodynamischen Daten der Liganden hinsichtlich der Ionisationsenthalpie zu korrigieren.

Mit dieser Analyse wollen wir der Antwort auf die Frage näherkommen: Was sind die Bindungsbeiträge eines starren Inhibitor Moleküls im Vergleich zu einem strukturell eng verwandten Molekül mit einem weniger starren Gerüst?

1.5 Referenzen

- [1] S.B. Shuker, P.J. Hajduk, R.P. Meadows, S.W. Fesik, Discovering high-affinity ligands for proteins: SAR by NMR, *Science* 274 (1996) 1531–1534.
- [2] J.R. Broach, J. Thorner, High-throughput screening for drug discovery, *Nature* 384 (1996) 14–16.
- [3] C.W. Murray, T.L. Blundell, Structural biology in fragment-based drug design, *Curr. Opin. Struct. Biol.* 20 (2010) 497–507.
- [4] J.K. Kranz, C. Schalk-Hihi, Protein thermal shifts to identify low molecular weight fragments, *Meth. Enzymol.* 493 (2011) 277–298.
- [5] U.H. Danielson, Fragment library screening and lead characterization using SPR biosensors, *Curr Top. Med. Chem.* 9 (2009) 1725–1735.
- [6] T. Neumann, H.-D. Junker, K. Schmidt, R. Sekul, SPR-based fragment screening: advantages and applications, *Curr. Top. Med. Chem.* 7 (2007) 1630–1642.
- [7] N. Drinkwater, H. Vu, K.M. Lovell, K.R. Criscione, B.M. Collins, T.E. Prisinzano, S.-A. Poulsen, M.J. McLeish, G.L. Grunewald, J.L. Martin, Fragment-based screening by X-ray crystallography, MS and isothermal titration calorimetry to identify PNMT (phenylethanolamine N-methyltransferase) inhibitors, *Biochem. J.* 431 (2010) 51–61.
- [8] M. Congreve, C.W. Murray, T.L. Blundell, Structural biology and drug discovery, *Drug Discov. Today* 10 (2005) 895–907.
- [9] D.E. Scott, A.G. Coyne, S.A. Hudson, C. Abell, Fragment-based approaches in drug discovery and chemical biology, *Biochemistry* 51 (2012) 4990–5003.
- [10] N. Howard, C. Abell, W. Blakemore, G. Chessari, M. Congreve, S. Howard, H. Jhoti, C.W. Murray, L.C.A. Seavers, van Montfort, Rob L M, Application of fragment screening and fragment linking to the discovery of novel thrombin inhibitors, *J. Med. Chem.* 49 (2006) 1346–1355.
- [11] N.J. Royle, D.M. Irwin, M.L. Koschinsky, R.T. MacGillivray, J.L. Hamerton, Human genes encoding prothrombin and ceruloplasmin map to 11p11-q12 and 3q21-24, respectively, *Somat. Cell Mol. Genet.* 13 (1987) 285–292.
- [12] G. Klebe, *Wirkstoffdesign: Entwurf und Wirkung von Arzneistoffen*, Spektrum Akademischer Verlag, Heidelberg, 2009. 354

2 Thermodynamic Signatures of Fragment Binding: Validation of Direct versus Displacement ITC Titrations

2.1 Einleitende Bemerkungen/Introductory Remarks

Das folgende Kapitel wurde in *Biochimica et Biophysica Acta- General Subjects* publiziert. Die untersuchten Fragmente 1 und 2 wurden von Dr. Martina Schäfer und Marie Fricke von Bayer Health Care als Thrombin Binder identifiziert. Erste thermodynamische Daten für diese zwei Fragmente mittels Verdrängungstitrationen mit Referenzligand A wurden in einer Kollaboration in der Arbeitsgruppe von Prof. Dr. Klebe ermittelt.

Die Auswahl der weiteren Fragmente und Liganden, die Durchführung der weiteren thermodynamischen Messungen und Enzymassays der Fragmente und Liganden erfolgten durch den Autor dieser Dissertation.

2.2 Abstract

Detailed characterization of the thermodynamic signature of weak binding fragments to proteins is essential to support the decision making process which fragments to take further for the hit-to-lead optimization.

Isothermal titration calorimetry (ITC) is the method of choice to record thermodynamic data, however, weak binding ligands such as fragments require the development of meaningful and reliable measuring protocols as usually sigmoidal titration curves are hardly possible to record due to limited solubility.

Fragments can be titrated either directly under low c -value conditions (no sigmoidal curve) or indirectly by use of a strong binding ligand displacing the pre-incubated weak fragment from the protein. The determination of Gibbs free energy is reliable and rather independent of the applied titration protocol. Even though the displacement method achieves higher accuracy, the obtained enthalpy-entropy profile depends on the properties of the used displacement ligand. The relative enthalpy differences across

different displacement experiments reveal a constant signature and can serve as a thermodynamic fingerprint for fragments. Low c -value titrations are only reliable if the final concentration of the fragment in the sample cell exceeds 2-10 fold its K_D value. Limited solubility often prevents this strategy.

The present study suggests an applicable protocol to characterize the thermodynamic signature of protein-fragment binding. It shows however, that such measurements are limited by protein and fragment solubility. Deviating profiles obtained by use of different displacement ligands indicate that the residual solvation pattern takes most likely a determinant impact on the resulting binding signature.

2.3 Introduction

Consideration of thermodynamic binding profiles beyond sole affinity appreciations in drug optimization is increasingly regarded as a valuable analytical tool. [1, 2, 3] Analyses of thermodynamic signatures of ligand binding have shown that on late-stage optimization improved binding is frequently achieved by enhancing the entropic component to the Gibbs free energy of binding. [4, 5] This results either from an appropriate rigidification of the lead scaffold in the protein-bound conformation, or from the attachment of lipophilic groups of growing size to optimally fill remaining unoccupied pockets in the binding site. Usually these strategies make the lead candidates more complex and overall more 'greasy', provoking concomitant problems such as unsatisfactory bioavailability or growing risk of undesired toxicity.[6, 7, 8] Accordingly, the hypothesis has been proposed to start lead optimization with small enthalpically advantaged binders as the entropic component will be added inevitably to the binding signature during late-stage optimization.[9, 10] Therefore, methods giving reliable access to thermodynamic signature analysis, particularly of weak binding initial hits, are essential to support the decision making process which compounds to take further for lead optimization. To this end, more meaningful measuring protocols to record accurate isothermal titration calorimetry (ITC) data of two-digit micromolar to even millimolar binders are required.

Anticipating a sufficiently large enthalpic binding component, the scope of ITC measurements to record protein-ligand complex formation ranges approximately from low-micromolar to two-digit nanomolar binders. In consequence weak binding ligands such as fragments are hardly detectable by this method. [11-13] To some degrees this difficulty can be overcome by titrating huge amounts of protein and fragment at very high concentrations. Nonetheless, this strategy of direct titrations is limited as usually protein supply is crucial and low solubility of either protein, and/or ligand or increasing protein degradation at high concentration impedes direct ITC titrations. Furthermore, titration curves recorded under these conditions are difficult to analyze as they usually lack sigmoidal shape and their inflection point can hardly be defined. The enthalpic binding component can only be extracted if the binding stoichiometry is arbitrarily adjusted to, e.g., 1:1 during the data fitting. However, frequently higher binding ratios are given for small binders such as fragments. In contemporary drug discovery fragment-based approaches play an increasingly important role and ITC would be an ideal complementary asset to identify the most promising candidates with high enthalpic efficacy. [1, 8] At present fragments are usually discovered by other screening methods such as surface plasmon resonance (SPR), saturation transfer difference NMR (STD-NMR) and thermal shift analysis (TSA).

To develop ITC protocols for weak binders we used a set of fragments with milli to three-digit micromolar affinities inhibiting the serine protease thrombin, to validate direct ITC titrations versus alternative titration protocols such as displacement experiments suggested by Zhang et al. [12]

To perform the required displacement experiments we used several strong binding reference ligands of deviating structural and thermodynamic profiles to evaluate the binding properties of the selected fragments. In the present study we will show that, even though the individual thermodynamic binding profiles recorded for the various fragments can be distinct, the relative differences among the fragment profiles seem to be independent of the actually selected displacement ligand. Furthermore, the approach

allows estimating experimental errors and helps to validate the reliability and validity of fragment binding profiles obtained by direct ITC titrations.

2.4 Results

2.4.1 *Data set of fragments and reference ligands*

For the anticipated analysis we assembled a dataset of fragments binding to the serine protease thrombin, an important drug target in the blood coagulation cascade for which in recent time drugs have been launched to the market. [13] With respect to the structural and thermodynamic characterization, this protein is well established in our laboratory and from this experience we selected a sample set of appropriate reference ligands for displacement titrations. We have chosen the fragments **1 - 4 (Table 2.1)** previously described as micro to milli-molar binders for thrombin. [14-16]

All selected fragments bind into the S1 pocket of thrombin, as could be confirmed by X-ray crystallography. The adopted binding modes are depicted in Figure 1 together with the structures of the larger and more potent reference ligands. The corresponding structures will be reported in a subsequent study.

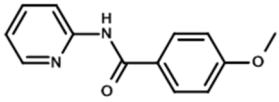
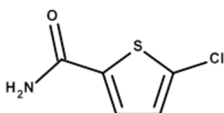
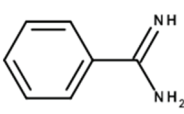
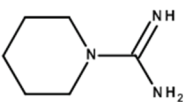
<i>Fragment</i>	<i>K_i Assay</i>	<i>K_D displacement ITC</i>	<i>K_D direct ITC</i>
<p>1</p> 	2430±309	1582±441	1003±224
<p>2</p> 	431±27	507±139	475±110
<p>3</p> 	258±1	455±109	355±29
<p>4</p> 	111±2	197±74	107±20

Table 2.1 Chemical formulas of the studied fragments with their estimated binding constants in μM .

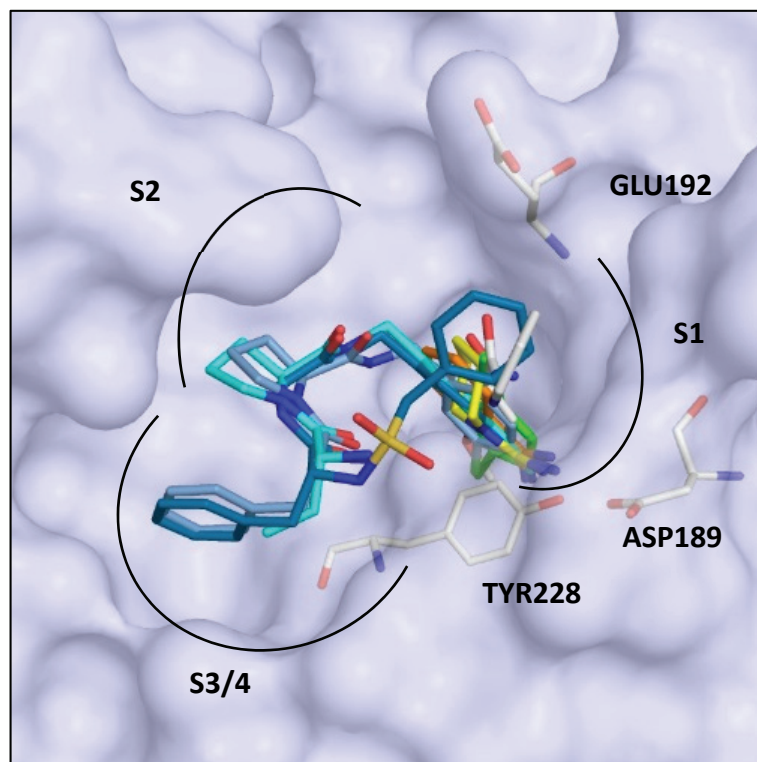


Figure 2.1 Fragments **1** (white), **2** (green), **3** (orange), **4** (yellow) and reference ligands **A** (light blue), **B** (turquoise) and **C** (dark blue) bound to the active site of thrombin. The protein is indicated by its solvent-accessible surface (grey) and the three residues Asp 189, Glu192 and Ty228. Oxygen atoms red, nitrogen atoms blue, sulfur atoms yellow are indicated. The specificity pockets S1, S2, S3/4 of the protease are schematically indicated.

To perform displacement titrations in order to characterize fragment binding (see below), appropriate reference ligands are required that exhibit an affinity in the three-digit nanomolar range, as they allow direct titrations with optimal sigmoidal shape of the titration curve. We selected three different reference ligands **A - C** (**Figure 2.2**) with distinct thermodynamic profiles and chemical composition, to investigate the influence of the reference ligand onto the thermodynamic profiles of the displaced fragments.

As a first step, we characterized the reference ligands by direct titration experiments resulting in thermograms similar to that obtained for ligand **A** shown in **Figure 2.3 a**). The synthesis of the reference compounds has been described previously. [17, 18, 19]

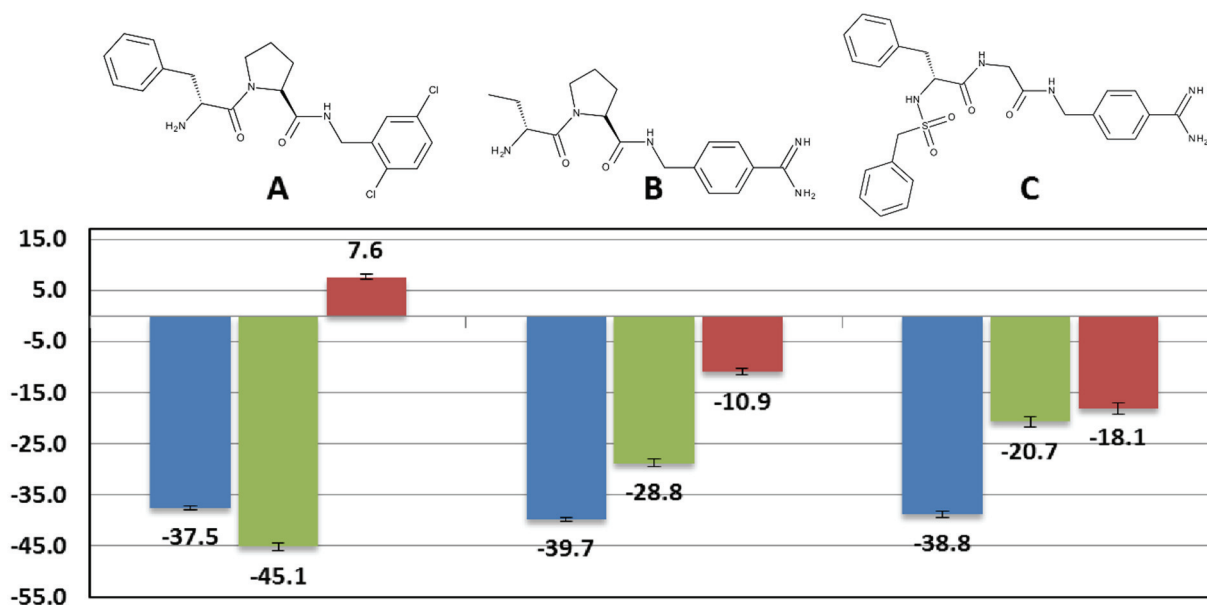


Figure 2.2 Thermodynamic parameters (ΔG° in blue, ΔH° in green, $-T\Delta S^\circ$ in red in kJ/mol) determined by direct ITC titrations for the strong binding reference ligands **A**, **B** and **C**. The estimated standard deviations are indicated by the black error bars.

In the following we will discuss only the ΔG° and ΔH° values, as these are the properties actually determined in an ITC experiment. The entropic component, usually specified as $-T\Delta S^\circ$, is calculated as the numerical difference of Gibbs free energy and enthalpic binding contribution according to $\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ$. All three reference ligands have been assessed with respect to a possibly superimposed protonation change which can be discovered by performing the titration experiments from different buffer conditions.

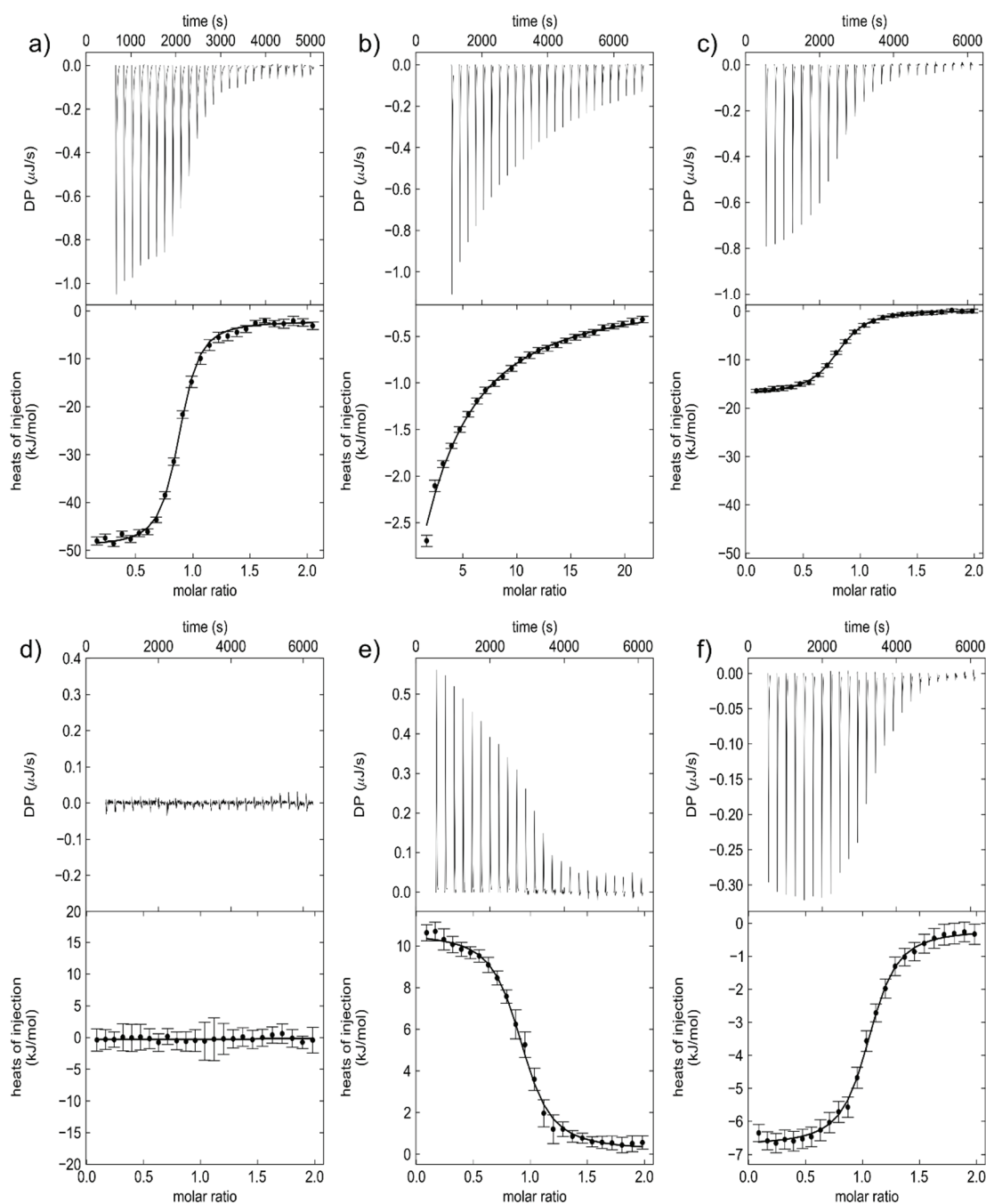


Figure 2.3 Images of representative thermograms obtained by different titration protocols applied in this study. The individual titration curves show: **a)** reference ligand **A** in direct titration; **b)** fragment **2** in direct titration; **c)** **2** is displaced by **A**; **d)** **C** displaced **4** but no difference in binding enthalpy could be recorded; **e)** **2** displaced by **C**; **f)** **3** displaced by **C**. In all cases the heat signals (in $\mu\text{J/s}$) as a response by the release of the ligand into the protein solution is shown over the course of the experiment along with the integrated heat signals of the injections (kJ/mol).

The compound classes of ligands **A** and **B** have been studied in great detail in the past and a simultaneous compensatory proton entrapment and release has been recognized suggesting a release of about 0.5 mol protons upon ligand binding from His57 which is partially protonated in the uncomplexed state. Virtually the same amount is picked up by the ligand upon protein binding. [19] For a scaffold closely related to that of ligand **C** previous studies indicated no compensatory protonation effect to be superimposed onto the binding event. [17, 20] Nonetheless, to minimize possibly sophisticating effects from proton release or pick-up of the surrounding buffer, all experiments were performed in pyrophosphate buffer which shows nearly negligible heat effects for changes in protonation states. Reference ligand **A** ($K_D=269\pm 50$ nM; $\Delta G^\circ=-37.5\pm 0.5$ kJ/mol) exhibits a predominantly enthalpic binding signature ($\Delta H^\circ=-45.1\pm 0.8$ kJ/mol) and addresses the S1 pocket of thrombin with a dichloro benzyl moiety, the S2 pocket with a proline and the S3 pocket with a D-Phe residue (**Figure 2.2**). Reference ligand **B** ($K_D=110\pm 16$ nM; $\Delta G^\circ=-39.7\pm 0.4$ kJ/mol) is a less enthalpic binder ($\Delta H^\circ=-28.8\pm 0.8$ kJ/mol) and accommodates the S1 pocket of thrombin with a benzamidine moiety, the S2 pocket with a proline and the S3 pocket with a *homo*-Ala residue. Reference ligand **C** ($K_D=162\pm 34$ nM; $\Delta G^\circ=-38.8\pm 0.6$ kJ/mol) shows balanced enthalpic ($\Delta H^\circ=-20.7\pm 0.9$ kJ/mol) and entropic ($-T\Delta S=-18.1\pm 1.1$ kJ/mol) binding contributions and is accommodated in the S1 pocket of thrombin using its benzamidine moiety, in the S2 pocket by an Ala residue, and the S3 pocket is occupied by a D-Phe residue and the S4 pocket by a benzylsulfonamide group.

2.4.2 Importance of the c-value for the shape of the titration curve

Crucial for the shape of a recorded ITC titration curve is the so-called c parameter, first introduced by Wiseman [21] and discussed in detail by Turnbull et al. [22] Detailed parameter studies have elucidated that a c-value of ≥ 10 reveals a proper sigmoidal curve but the value should remain below 1000 to avoid a vertical stepwise curve. However, if c falls below 10, the inflection point can hardly be extracted as the titration curve is no longer sigmoidal. In consequence, the binding stoichiometry is no longer available from the experiment without arbitrary assumptions. An example for such a titration curve is given in **Figure 2.3b**.

The following estimation of the concentrations required for an ITC experiment with fragments should make the practicability of such a determination clearer. To obtain a minimal c -value of 10 for a typical fragment with $K_D = 1$ mM, a protein concentration of 10 mM in the sample cell according to $c = n[\text{Protein}] K$ is required. At the end of the titration the final fragment concentration should be at least twice (better 10-times) as high as this value to achieve sufficient saturation.

If the fragment is released from the syringe of an ITC200 device with a volume of approximately 40 μl into the sample cell which has a five-time larger volume, the concentration of the fragment in the syringe has to be at least 100 mM. In many cases, such a high solubility is difficult to achieve. Titration experiments can principally be reversed, but as a similar amount of protein will be required, solubility and particularly protein stability will hardly be accomplishable to perform such titration experiments. This crude estimation makes obvious that protein and ligand solubility will be the predominant limiting factors to study low-affinity systems following the popular strategy to design an ITC experiment resulting in sigmoidal titration curves. If fragment titrations at c -values ≥ 10 are hardly feasible, the following alternatives can be suggested.

The first strategy focuses on displacement titrations. In this alternative, the protein must be saturated with the fragment prior to the titration experiment while preparing the protein solution. To achieve a reasonable saturation a crude estimate about the expected affinity of the fragment must be available (see below). Subsequently, the pre-incubated weak binding ligand is displaced in a titration experiment with a beforehand characterized reference ligand. By subtracting the curves of the directly titrated reference ligand and the curve of the displacement experiment, the thermodynamic profile of the weak binding fragment becomes available (**Figure 2.3 a)** and **c)**). Is this protocol of any advantage with respect to the required concentrations during the experiment? The fragment which is initially dissolved in DMSO is added to the protein solution in a way to match the finally desired DMSO concentration in the sample cell. Following this strategy, overall a lower concentration of the respective fragment will be needed usually not exceeding the solubility limits. The displacement titration strategy might suffer from the fact that

particularly in novel drug discovery projects potent and well-characterized reference ligands for the displacement strategy are not available.

In such cases, the only feasible alternative is to perform ITC experiments as low c -value titrations. As mentioned, they suffer from the fact that a sigmoidal shape is no longer given and the inflection point of the curve is poorly defined. Therefore, it is necessary to arbitrarily adjust the binding stoichiometry to a predefined value during curve fitting (Figure 2.3b) and it is crucial to control concentrations of ligand and protein very accurately. To achieve a sufficient saturation of the protein in during the titration, the concentration of the fragment added to the sample cell must be larger than its estimated K_D value. Furthermore, a reasonable binding stoichiometry, mostly set to 1:1, of the formed protein-fragment complex must be anticipated. Nonetheless, in protein crystallography often the binding of more than one fragment molecule to the target protein is observed, particularly at high concentrations, which makes the assumption of a 1:1 ratio questionable. To perform a direct low c -value titration of a fragment, a rather concentrated fragment solution is released from the syringe to the protein in the sample cell, however, this solution will be less concentrated than that in a scenario with a c -value ≥ 10 . Nonetheless, compared to the alternative displacement strategy, the required concentration even for the just described direct low c -value titration will be undoubtedly higher, which makes the displacement strategy better applicable under practical conditions. This aspect has also been discussed by Turnbull et al. [23, 24] To contrast the different strategy options, we will compare in the following the results obtained by displacement protocols with those recorded under direct low c -value conditions.

2.4.3 *Estimation of fragment affinity*

Both suggested ITC protocols require a crude estimation of the binding constant of the studied fragment to secure sufficient saturation of the protein. In a real-life scenario, K_i values of the studied fragments are not necessarily available. However, in the present case, they could be determined by an independent, highly sensitive fluorescence-based assay (see Materials & Methods).

2.4.4 Results obtained by the displacement titrations

The reference ligands appropriate for displacement titrations need to exhibit an affinity in three-digit nanomolar range (optimal 100 to 400 nM) to fall into the optimal sensitivity range of ITC. This allows thermodynamic characterization via direct titrations with good c -values of 20 - 100 and a low standard deviation for K_D and ΔH° . Broecker et al. suggested a c -value of 40 to be optimal. [25] In **Figure 2.4 a), b), c)**, the results for the displacement titrations of fragments **1 - 4** using the reference ligands **A – C** are compiled.

The fragment concentrations for the pre-incubation were adjusted in a way to achieve 2 mM concentration in the sample cell. The deviations of the Gibbs free energy determinations are rather small, for displacement titrations with the three reference ligands, suggesting a standard deviation of 0.6 kJ/mol. In contrast, the recorded enthalpies show mutual deviations facing the individual experiments based on different reference ligands against one another, which are much larger than the expected experimental error. Obviously different reference ligands do not reveal similar enthalpies to displace the considered pre-incubated fragments.

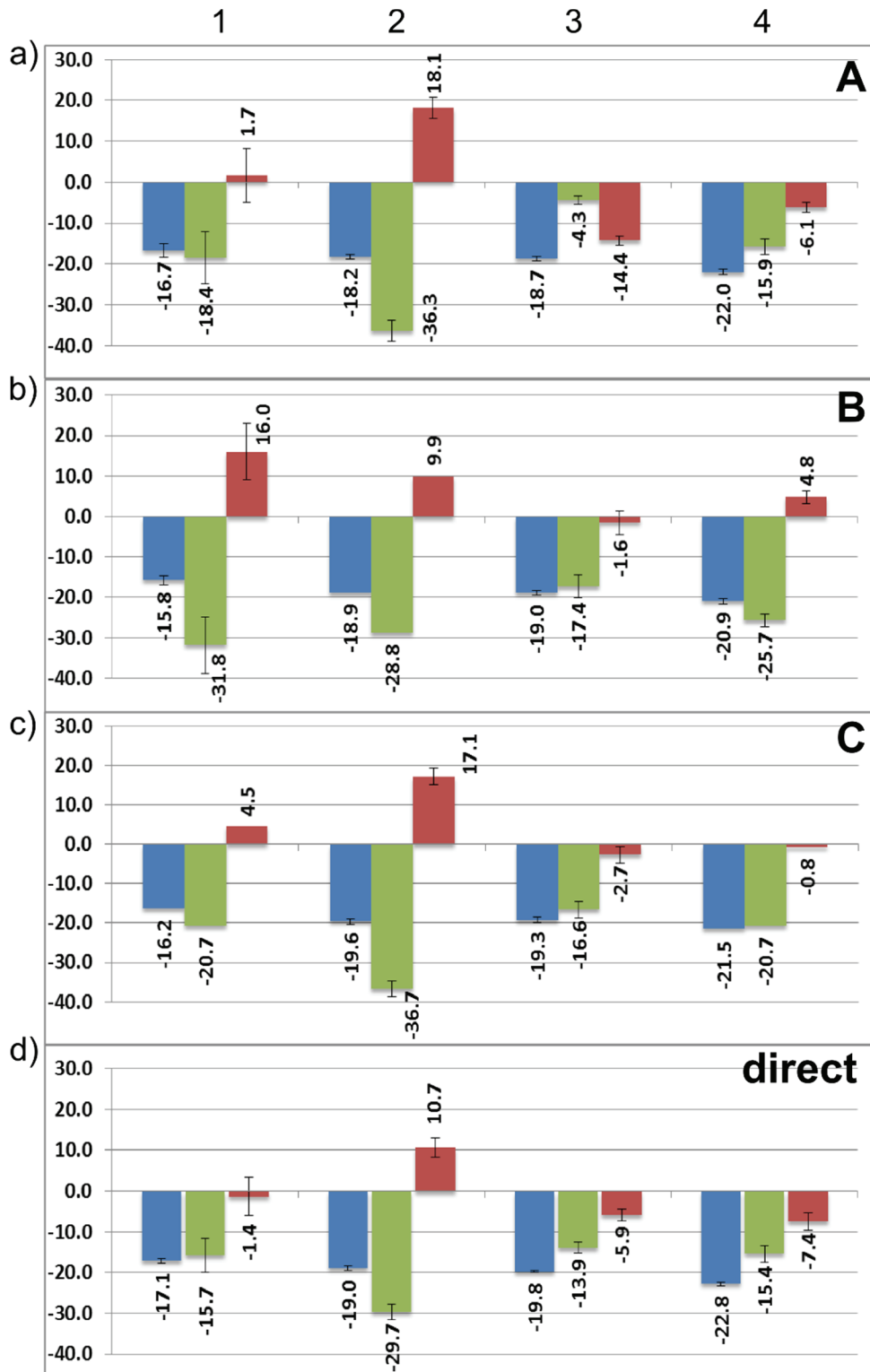


Figure 2.4 Detailed figure caption on next page.

Figure 2.4¹ Thermodynamic signatures (ΔG° in blue, ΔH° in green, $-\Delta S^\circ$ in red as kJ/mol) for fragments **1-4** determined by ITC displacement titrations using reference ligands **A (a)**, **B (b)**, **C (c)** and obtained by direct low-c ITC titrations (**d**) are shown. The estimated standard deviations of the different experiments are indicated by the black error bars. In case the enthalpies were taken as similar to the binding enthalpy of the reference ligand (B-2, C-1 and C-4) no error estimations are indicated.

2.4.5 Results obtained by direct low c-value titrations

In all experiments a concentrated fragment solution was titrated to the protein in the sample cell. For **1** and **2** an almost saturated solution of 5 mM has been applied. Owing to their better solubility, fragments **3** and **4** could be titrated from the syringe with concentrations of 10 mM. The obtained thermodynamic profiles are shown in Figure 4d. Remarkably, the binding free energies for the direct titrations match reasonably well with those obtained by the displacement titrations. The enthalpic signal of **1**, which shows large standard deviations, falls roughly into the range indicated by the three displacement titrations. For **2 – 4** the enthalpic signals could be determined with lower standard deviations. They match, interestingly enough, with the some of the experiments performed with deviating reference ligands. These results suggest furthermore, that the profiles obtained for the different ligands depend on the chosen reference ligand.

For low c-value titrations the final concentration of the fragment in the sample cell is crucial and has to exceed its binding constant. To study the influence of the attained fragment concentration at the end of the titration, we determined the thermodynamic parameters for **4** using different concentrations in the syringe. The results are listed in **Table 2.2** and the corresponding thermograms are shown in **Figure 2.5**. The stoichiometric equivalence, adjusted in the evaluation procedure, is passed, dependent on the concentration, after a deviating number of drops released from the syringe to the sample cell. By use of the lowest syringe concentration of 1mM, we observe a higher binding affinity for **4** compared to the other three concentrations (5 mM; 10mM and 15mM) which

¹ We identified a programming error in the routine used for the calculation of the reported binding enthalpies of the displacement titrations. Due to this error a corrigendum considering an explanation of the error and recalculated data has been send to the journal. Fortunately, the recalculated data do not deviate by more than twice the standard deviations from the originally reported data, thus the interpretation of our results is not affected. In the present chapter, however, the text is entirely based on the corrected data and all affected figures (**Figure 2.4 a-c**); **Figure 2.7**, **Figure 2.8**) have been replaced according to the corrected data.

show rather consistent values. Nonetheless with increasing concentrations in the syringe, larger standard deviations of ΔH° are recorded.

4 Syringe Concentration	ΔG°	ΔH°	$-T\Delta S^\circ$
1 mM	-25.2±0.2	-13.4±0.1	-11.7±0.3
5 mM	-22.8±0.2	-15.4±2.8	-7.4±3.0
10 mM	-22.6±0.2	-16.3±2.8	-6.3±2.5
15 mM	-22.4±0.5	-14.1±4.6	-8.2±4.1

Table 2.2 Thermodynamic parameters in kJ/mol determined for **4** in direct low-c titrations using different high concentrations in the syringe.

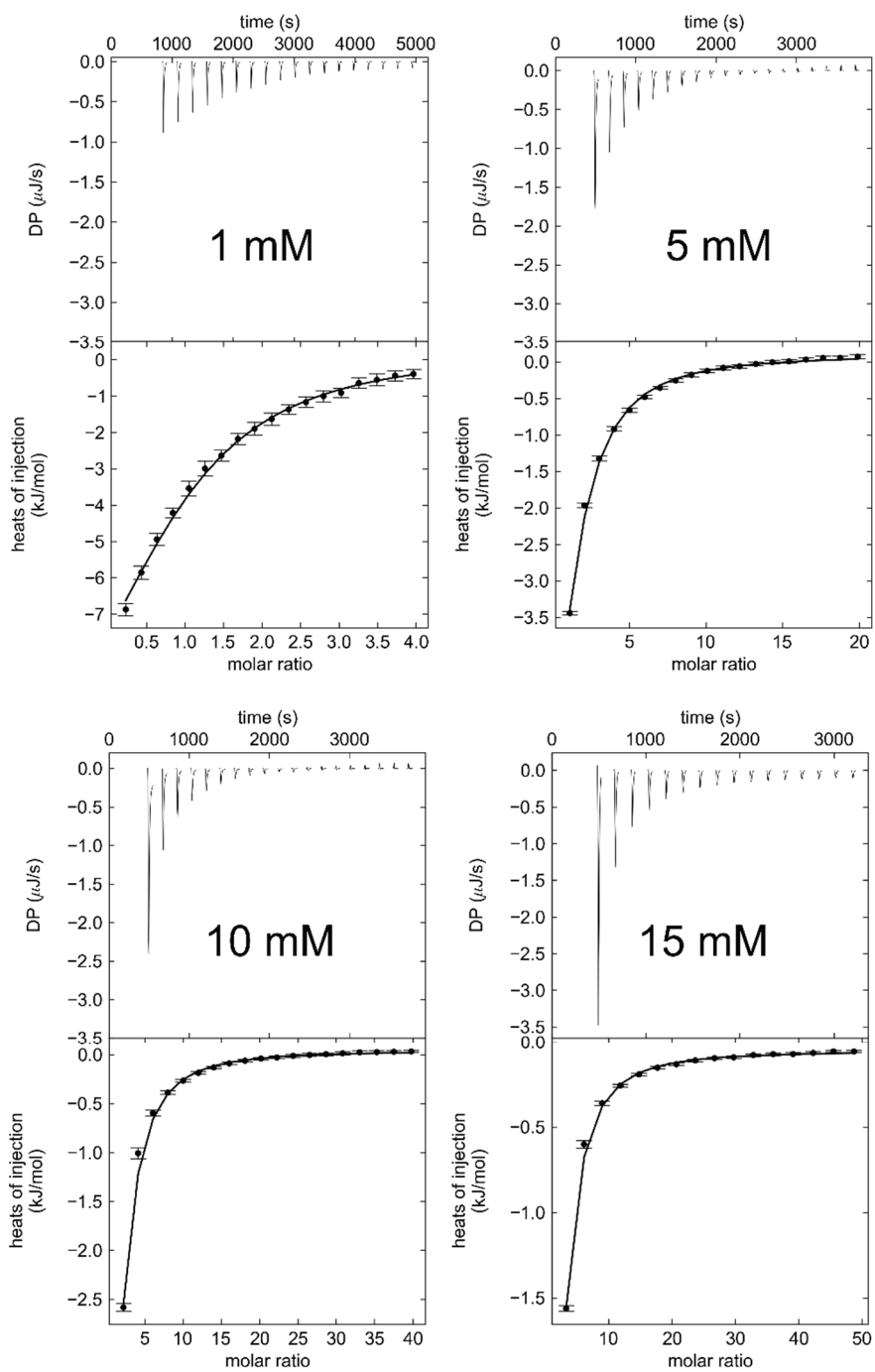


Figure 2.5 Thermograms of 4, applying different high fragment concentrations from the syringe.

2.5 Discussion

The different ITC experiments show that the determination of the binding constant and accordingly of the Gibbs free energy of binding is less dependent on the selected titration strategy than the enthalpy values. In **Figure 2.6**, the affinities based on the various K_D values determined from direct and displacement ITC experiments are plotted against the K_i values from the biochemical assay of fragments **1 - 4**.

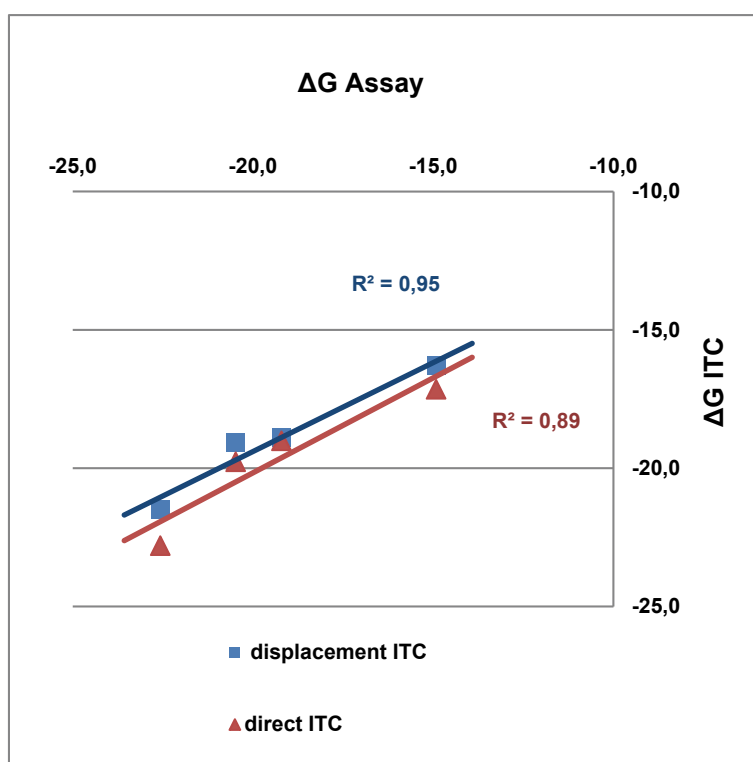


Figure 2.6 The binding affinities expressed by $\Delta G = -RT \ln K$, are plotted against each other in kJ/mol to assess their mutual correlation, where K reflects either the thermodynamic dissociation constant K_D (displacement ITC and direct ITC) or the enzyme kinetic binding constant K_i (biochemical assay) of fragments **1-4**.

A convincing correlation of $r^2=0.95$ is observed for the displacement ITC experiments whereas direct ITC titrations shows a slightly lower correlation of $r^2=0.89$.

We selected three reference ligands of which **A** is a strong enthalpy-dominated binder, whereas **B** and **C** show smaller enthalpic signals. This leads to the fact that e.g., the displacement of **2** by **A** (**Figure 2.3c**) and of **3** by **C** (**Figure 2.3f**) results in an exothermic displacement reaction whereas the displacement of **2** by **C** corresponds to an endothermic process (**Figure 2.3e**). For the titrations of **B** to displace **2**, **C** to displace **1** and **C** to displace **4** no significant enthalpy difference signal could be recorded (e.g. **Figure 2.3d**), simply as the signal for the release of the pre-incubated fragment from the protein is enthalpically compensated by the binding of the more potent reference ligand. For these fragments the binding enthalpies were supposed to match the binding enthalpy of the respective reference ligands (ΔH° of **C** for **C** displacing **1** and **4**; ΔH° of **B** for **B** displacing **2**) and we refrained from giving error estimations for these titrations (**Figure 2.4**). Instead of estimating roughly the accuracy we consider the errors found for the characterization of the reference ligands.

In case of the flat thermograms, we calculated the free energy of ΔG° of the fragment as the mean of the ΔG° values obtained by the titrations with the other two reference ligands which revealed a measurable enthalpic signal. In **Figure 2.7** we compiled the relative differences in the binding enthalpies of the various fragments and the reference ligands. This diagram clearly shows that, different to the ΔG° determinations, the recorded enthalpy signals strongly depend on the profile of the selected reference ligand.

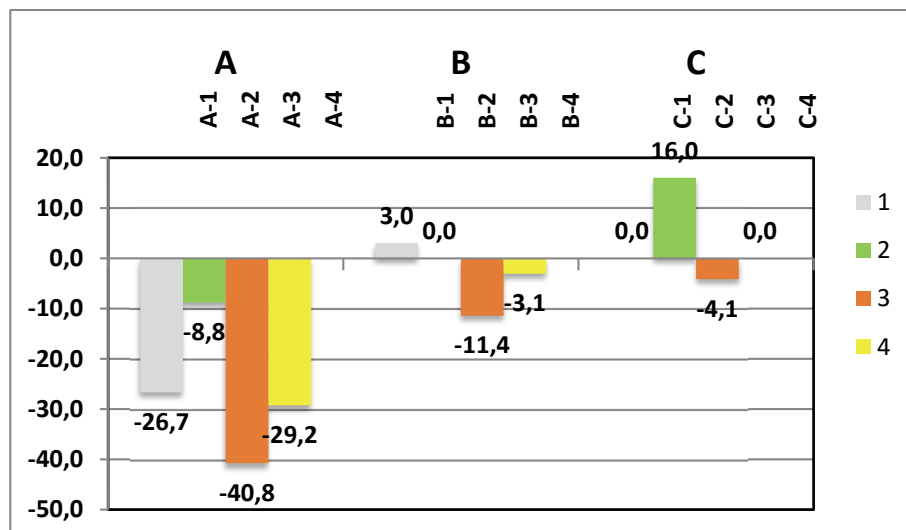


Figure 2.7² Binding enthalpy differences $\Delta\Delta H$ between reference ligands **A**, **B**, **C** and fragment **1-4** are plotted in kJ/mol. In three cases no measurable enthalpy difference could be detected as the heat signal produced by the displacement of the fragment is virtually compensated by the signal of the binding of the reference ligand.

To emphasize the dependence on the profiles of the applied reference ligands, we plotted the binding enthalpy determined for fragments in the individual displacement experiments and also from the direct low c titration experiment (**Figure 2.8**).

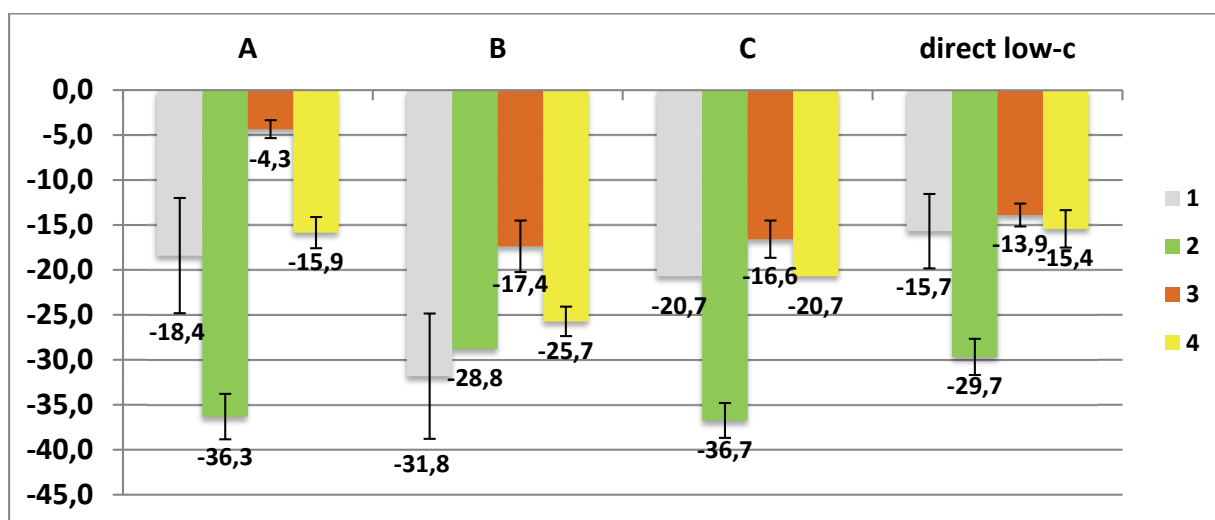


Figure 2.8³ Direct Comparison of the enthalpic data across the fragment series, **1** (white), **2** (green), **3** (orange) and **4** (yellow) plotted in kJ/mol for displacement titrations with reference ligands **A**, **B**, **C** and **direct low-c** titrations.

^{2,3} See footnote 1 on page 29.

Interestingly, this evaluation reveals a similar relative difference of each fragment independent of the actually applied displacement ligands. For example, fragment **1** binds much less enthalpic than fragment **2** or fragment **4** is a more enthalpic binder than **3**. According to this evaluation, fragment **2** is the most enthalpic binder of the series. Fragment **1** shows less enthalpic binding compared to **2** and **4**, however, a slightly more exothermic signature is apparent compared to **3**.

As an alternative scenario ITC experiments can be performed as low c-value titrations. Disadvantage of this strategy is the fact that the inflection point of the titration curve is poorly defined and the binding stoichiometry must be fixed arbitrarily. To achieve a sufficient saturation of the titrated protein, the concentration of the added fragment to the sample cell must be larger than its K_D value, simultaneously assuming a particular binding stoichiometry, e.g., of 1:1 for the formed protein-fragment complex. Turnbull et al. experienced comparable results for titrations where only 70, 80 or 90% protein saturation could be achieved which indicates some robustness of the titration settings. [22, 24]

In our direct low c-titrations, for fragment **1** we could only accomplish a final concentration in the sample cell of about 1 mM (which corresponds to approx. 30% saturation) due to its poor solubility in the applied buffer. In the biochemical assay we determined a K_i value of 2.4 mM and in the displacement titrations a K_D value of 1.6 mM, respectively. Which means that, as an end point of our titration, we achieved approximately a 30% saturation of the protein but the required excess concentration of the fragment could not be realized. As a matter of fact, this titration results in rather large errors of the binding enthalpy determination (**Figure 2.4 d**). In consequence, the ΔH° values obtained for the direct low c-value titration and those available from displacement titrations show strongly deviating values (**direct** $\Delta H^\circ = -15.7 \pm 4.1$ kJ/mol; **A** $\Delta H^\circ = -18.4 \pm 6.4$ kJ/mol; **B** $\Delta H^\circ = -31.8 \pm 7.0$ kJ/mol; **C** $\Delta H^\circ = -20.7 \pm 0.9$ kJ/mol). This observation is evidently associated with the experimental uncertainties and errors in the parameter determination.

By contrast the results obtained for the direct titration of fragment **2** agree reasonably well to the results derived from the three displacement experiments (**direct** $\Delta H^\circ = -29.7 \pm 2.0$ kJ/mol; **A** $\Delta H^\circ = -36.3 \pm 2.5$ kJ/mol; **B** $\Delta H^\circ = -28.8 \pm 0.8$ kJ/mol; **C** $\Delta H^\circ = -36.7 \pm 1.9$ kJ/mol) (**Figure 2.4**). Exhibiting a binding affinity of about 500 μ M fragment **2** can easily attain the required degree of saturation of the protein of more than 70%. Also the strong enthalpic signal of $\Delta H^\circ = -29.7 \pm 2.0$ kJ/mol in the direct titration leads to good signal to noise ratio allowing data integration associated with lower uncertainties.

Also the direct titrations of fragment **3** ($\Delta H^\circ = -13.9 \pm 1.2$ kJ/mol) match well with the results obtained by displacement titrations with reference ligands **B** ($\Delta H^\circ = -17.4 \pm 2.9$ kJ/mol) and **C** ($\Delta H^\circ = -16.6 \pm 2.1$ kJ/mol), however, not for the titration using ligand **A** ($\Delta H^\circ = -4.3 \pm 1.0$ kJ/mol) (**Figure 2.4**).

It is quite remarkable that fragment **4** which appears on first sight chemically related to fragment **3** and which should experience a similar binding mode in the S1 pocket of thrombin shows a reverse signature of the measured enthalpy values. The result of the direct titration ($\Delta H^\circ = -15.4 \pm 2.1$ kJ/mol) fits to the displacement experiment with reference ligand **A** ($\Delta H^\circ = -15.9 \pm 1.7$ kJ/mol) whereas the titrations using **B** ($\Delta H^\circ = -25.7 \pm 1.6$ kJ/mol) and **C** ($\Delta H^\circ = -20.7 \pm 0.9$ kJ/mol) as reference ligands result in quite deviating binding enthalpies. Obviously, features of the individual binding modes and accordingly of the binding characteristics of the reference ligand are responsible for the obtained thermodynamic profile differences. Presumably, this can be explained by considering in detail the residual solvation structure with respect to the bound water molecules present in the various complexes, an issue which will be discussed in a subsequent study.

To study the influence of different syringe concentrations which lead to deviating concentrations of the fragment in the sample cell at the end of the titration, we titrated **4** at four different concentrations (**Table 2.2, Figure 2.5**). At the lowest concentration of 1mM only a saturation of the protein of about 63% is attained. As a result, ΔG° is overestimated and a too small value is found for ΔH° by integrating over all peaks. At the higher concentrations much better saturation is achieved (5mM approx. 90%; 10mM

approx. 95% and 15mM approx. 97%). However, higher fragment concentrations in the syringe lead to an earlier saturation of the protein usually within only 1-2 injections, therefore extrapolation of the incomplete binding isotherm becomes crucial and rather uncertain at higher concentrations. Quite large standard deviations are obtained for ΔH° . To circumvent either incomplete or too fast protein saturation an amount of 70% to 90 % over-titration is advisable. As a rough estimate the shape of the thermogram should be analyzed. If stoichiometry is passed within one injection, the concentration is likely too high, if several injections (4-5 injections) are required to meet stoichiometry, the syringe concentration is too low and sufficient saturation will hardly be accomplished. Remarkably, the measurements at high saturation levels of over 90% show comparable results for ΔG° (**Table 2.2**) and they are in good accordance with the fluorescence-based assay results (**Table 2.1**). This finding suggests that the latter protocol gives more reliable ΔG° estimations for the fragments.

As important insight the present analysis shows, however, that the binding enthalpy of weak binding ligands measured by different displacement experiments may depend on the actual chemical properties of the reference ligand used for the titration. Important enough relative differences between their individual binding enthalpies recorded with respect to the deviating reference ligands are conserved.

In consequence, the relative data obtained by this strategy can be used to classify different fragments to exhibit a more enthalpic or entropic binding signature with respect to another fragment from the studied series.

2.6 Conclusions

Applying displacement titrations and direct titrations at low c-value, ITC can be used as a reliable technique to study low-affinity interactions as usually given for the binding of a fragment to a target protein.

Some important aspects can be concluded from this study. First of all, the fragment must bind to a part of the protein binding pocket that overlaps with the binding pose region of the applied reference ligand. Secondly, the selected reference ligand must have a significant higher or lower binding enthalpy so that a heat difference signal can be recorded for the fragment (**Figure 2.6**). Furthermore, it has to be regarded that, as for any displacement titration, all errors affecting the parameter determinations of the reference ligand will add to the accuracy of the parameters obtained for the fragment. Finally, if no suitable reference ligand is available and the fragment binds enthalpically enough, low c -value titrations can be applied as an alternative. In this case, however, some anticipated knowledge about the expected binding affinity of the fragment must be available to estimate the required excess concentration of the fragment at the end of the titration. In addition, the stoichiometry of the binding event must be arbitrarily fixed to an assumed value which may be in many cases questionable. If these conditions are regarded, displacement titrations as well as low c -value titrations are in convincing agreement to determine the binding affinities of fragments particularly if it is possible to achieve a concentration of the fragment in the sample cell at the end of the titration that exceeds the K_D value by a factor of 2-10 which corresponds to a saturation of 70 to 90 %.

To our experience solubility of the fragment and the protein are the most crucial issues in low c -value titrations because the required high concentrations of the fragment to be studied in the injection syringe must match or even exceed the final sample cell concentrations with respect to the fragment's K_D value. Only then data with good signal-to-noise level can be recorded. Furthermore, to estimate the required concentration of the fragment, a crude idea of the fragment's K_D must be available from an independent experiment.

Measured binding enthalpies should not be compared quantitatively across different experimental conditions, but only relative to each other by using one carefully selected measurement protocol. Remarkably, different reference ligands used for the displacement of pre-incubated fragments reveal deviating enthalpic signals. Most likely these difference can be traced back to differences in the water structure and thus in the residual solvation

pattern at the binding site of the different ligands used for titration. Nonetheless, considering the relative enthalpy differences across the different displacement experiments seems to reveal a consistent relative difference, which subsequently allows to characterize the studied fragments relative to one another. Furthermore, detailed interpretation of the enthalpic signature requires a comparison of the corresponding crystal structures which will be performed in a subsequent contribution. Usually this information is required to select fragments as superior enthalpy-dominated candidates for further development.

2.7 Experimental Section / Materials & Methods

2.7.1 ITC

ITC experiments were performed using an ITC200™ system from GE Healthcare, Northampton, MA, USA.

Thrombin was obtained from CSL Behring (Marburg, Germany) and purified from Beriplast® as one batch for the entire experimental series. Thrombin was extracted by dialysis using an experimental buffer of 50 mM TSPP, 100 mM NaCl, 0.1 % polyethylene glycol 8000, and pH 7.8. Subsequently, the protein could be used for all titration experiments. The protein concentration was measured by absorbance at 280 nm using a NanoDrop 2000c Spectrophotometer from Thermo Scientific.

All ITC experiments were started at 25 °C with a reference power of 5 kcal/s after a stable baseline had been achieved. The pre-titration delay was set to 300s. For direct titrations ligand injections of 0.3 µL (to prevent artefacts arising from small syringe leakages or air in the syringe) were followed by 19 to 27 injections of 1.5 – 2.0µL with at least 180 s interval between each injection.

For displacement experiments, fragments from 500 mM DMSO stock solutions were used directly to saturate the thrombin solution (42 μ M). Subsequently the DMSO concentration in all solutions was adjusted to 3 %. All titrations were performed at least in duplicate.

In case of fragment **1** only a saturation concentration close to its K_D for thrombin could be achieved due to poor solubility in the buffer. For fragments **2** to **4** a final concentration five to ten-times larger than the K_D for thrombin binding could be achieved and according to the **equation 2.1** corresponds to a 50 % to 91 % saturation (D_{sat}).

$$[fragment]_{cell} = \frac{D_{sat} [Protein] - [Protein] - K_D fragment}{1 - \frac{1}{D_{sat}}} \quad (2.1)$$

As supporting information, we provide an Excel sheet to calculate the degree of saturation, according to **equation 1** depicted below (**Figure 2.9**) for ligands with affinities from 0.1 mM to 2 mM the typical range found for fragments. Subsequently, the high-affinity ligands **A**, **B**, **C** (0.5 mM from 50 mM DMSO stock solution) were added to the thrombin-fragment complex using 22 to 27 injections of 1.3 to 1.5 μ L. Data were integrated using the program Nitpic 1.0.0 [26] which provides an automatic baseline determination and peak integration by peak-shape analysis, which is extremely valuable when dealing with data showing low signal-to-noise levels.

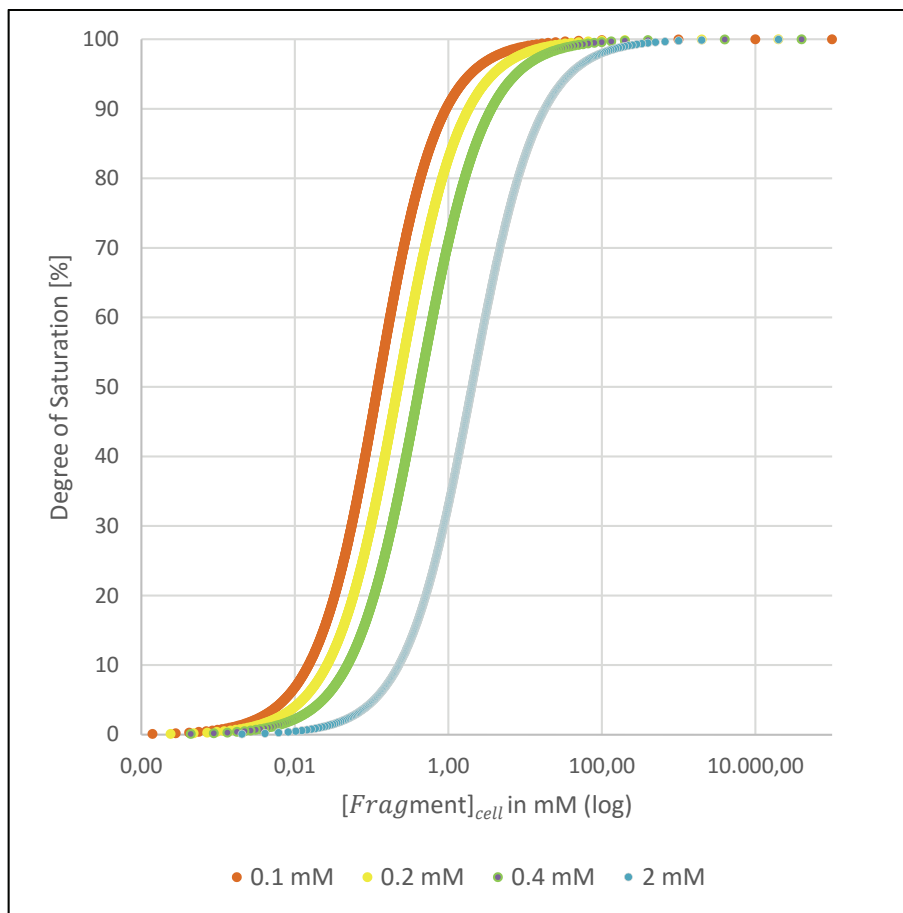


Figure 9 Degree of saturation (D_{sat}) is plotted against ligand concentration in the sample cell ($[\text{Fragment}]_{\text{cell}}$) on logarithmic scale for affinities from 0.1 mM to 2 mM.

Isotherm fitting was conducted with the program SEDPHAT 10.58d [27]. The first data points were excluded from data analysis. K_a and ΔH for the fragments have been calculated using **equations 2.2** and **2.3**, adapted from Zhang et al. [12]

$$K_{\text{fragment}} = \left(\frac{K_{\text{reference ligand}}}{K_{\text{displacement titration}}} - 1 \right) \frac{1}{[\text{Fragment}]_{\text{cell}}} \quad (2.2)$$

$$\Delta H^{\circ}_{\text{fragment}} = (\Delta H^{\circ}_{\text{reference ligand}} - \Delta H^{\circ}_{\text{displacement titration}}) \left(1 + \frac{1}{K_{\text{fragment}} [\text{Fragment}]_{\text{cell}}} \right) \quad (2.3)$$

2.7.2 Bioassay

Kinetic inhibition of human thrombin (from Beriplast®, CSL Behring, Marburg, Germany) was determined photometrically for the reference ligands at 405 nm using the chromogenic substrate Pefachrom tPa (LoxoGmbH, Dossenheim, Germany) as previously described [28] under the following conditions: 50 mM Tris-HCl, pH 7.4, 154 mM NaCl, 5 % DMSO, 0.1 % polyethylene glycol 8000 at 25°C using different concentrations of substrate and inhibitor. K_i values with $n \geq 3$ were determined as described by Dixon. [29]

Thrombin fragments were characterized by a more sensitive fluorogenic assay. Kinetic data were obtained using Tos-Gly-Pro-Arg-AMC (tosyl-Gly-Pro-Arg-aminomethylcoumarin) as the fluorogenic substrate [30] with a Safire II plate reader (Tecan, Switzerland, $ex = 380$ nm, $em = 460$ nm). The experimental buffer contained 50 mM Tris-HCl, 154 mM NaCl, 0.1 % polyethylene glycol 8000 and 5 % DMSO at pH 7.4. The K_m of the substrate ($5.7 \pm 0.4 \mu\text{M}$) was measured at eleven different substrate concentrations and the resulting curve was analyzed using GraFit 4 (Erithacus Software Limited, Staines, UK). Substrate cleavage was measured by monitoring the change in fluorescence across a dilution series of at least ten inhibitor concentrations (6.25 – 0.01 mM) at 1 μM substrate over 600 sec. The fluorescence signal was plotted versus time and by linear regression the reaction rates were calculated. The reaction rates were plotted against the respective inhibitor concentration and the resulting curve was fitted with ORIGIN software using Michaelis-Menten kinetics. All measurements were performed at least in triplicate.

2.8 References

- [1] J.E. Ladbury, G. Klebe, E. Freire, Adding calorimetric data to decision making in lead discovery: a hot tip, *Nat Rev Drug Discov* 9 (2010) 23–27.
- [2] E. Freire, A thermodynamic approach to the affinity optimization of drug candidates, *Chem Biol Drug Des* 74 (2009) 468–472.
- [3] J.B. Chaires, Calorimetry and thermodynamics in drug design, *Annu Rev Biophys* 37 (2008) 135–151.
- [4] T.S.G. Olsson, M.A. Williams, W.R. Pitt, J.E. Ladbury, The thermodynamics of protein-ligand interaction and solvation: insights for ligand design, *J. Mol. Biol.* 384 (2008) 1002–1017.
- [5] A. Glas, D. Bier, G. Hahne, C. Rademacher, C. Ottmann, T.N. Grossmann, Constrained peptides with target-adapted cross-links as inhibitors of a pathogenic protein-protein interaction, *Angew. Chem. Int. Ed. Engl.* 53 (2014) 2489–2493.
- [6] G.E. de Kloe, D. Bailey, R. Leurs, I.J.P. de Esch, Transforming fragments into candidates: small becomes big in medicinal chemistry, *Drug Discov. Today* 14 (2009) 630–646.
- [7] G.G. Ferenczy, G.M. Keserű, Thermodynamics of fragment binding, *J Chem Inf Model* 52 (2012) 1039–1045.
- [8] G.G. Ferenczy, G.M. Keserű, How are fragments optimized? A retrospective analysis of 145 fragment optimizations, *J. Med. Chem.* 56 (2013) 2478–2486.
- [9] V. Borsi, V. Calderone, M. Fragai, C. Luchinat, N. Sarti, Entropic contribution to the linking coefficient in fragment based drug design: a case study, *J. Med. Chem.* 53 (2010) 4285–4289.
- [10] S.F. Martin, J.H. Clements, Correlating structure and energetics in protein-ligand interactions: paradigms and paradoxes, *Annu. Rev. Biochem.* 82 (2013) 267–293.
- [11] D.C. Rees, M. Congreve, C.W. Murray, R. Carr, Fragment-based lead discovery, *Nat Rev Drug Discov* 3 (2004) 660–672.
- [12] Y.L. Zhang, Z.Y. Zhang, Low-affinity binding determined by titration calorimetry using a high-affinity coupling ligand: a thermodynamic study of ligand binding to protein tyrosine phosphatase 1B, *Anal. Biochem.* 261 (1998) 139–148.
- [13] W.G. Eisert, N. Huel, J. Stangier, W. Wienen, A. Clemens, J. van Ryn, Dabigatran: an oral novel potent reversible nonpeptide inhibitor of thrombin, *Arterioscler. Thromb. Vasc. Biol.* 30 (2010) 1885–1889.
- [14] M. Fricke, Diploma Thesis, FU Berlin, Germany, (2013)
- [15] K. Hilpert, J. Ackermann, D.W. Banner, A. Gast, K. Gubernator, P. Hadváry, L. Labler, K. Müller, G. Schmid, T.B. Tschopp, Design and synthesis of potent and highly selective thrombin inhibitors, *J. Med. Chem.* 37 (1994) 3889–3901.
- [16] T. Steinmetzer, J. Stürzebecher, Progress in the development of synthetic thrombin inhibitors as new orally active anticoagulants, *Curr. Med. Chem.* 11 (2004) 2297–2321.

- [17] M. T. Khayat, PhD Thesis, University at Buffalo, USA, (2013)
- [18] B. Baum, M. Mohamed, M. Zayed, C. Gerlach, A. Heine, D. Hangauer, G. Klebe, More than a simple lipophilic contact: a detailed thermodynamic analysis of nonbasic residues in the s1 pocket of thrombin, *J. Mol. Biol.* 390 (2009) 56–69.
- [19] B. Baum, L. Muley, A. Heine, M. Smolinski, D. Hangauer, G. Klebe, Think twice: understanding the high potency of bis(phenyl)methane inhibitors of thrombin, *J. Mol. Biol.* 391 (2009) 552–564.
- [20] A. Biela, F. Sielaff, F. Terwesten, A. Heine, T. Steinmetzer, G. Klebe, Ligand binding stepwise disrupts water network in thrombin: enthalpic and entropic changes reveal classical hydrophobic effect, *J. Med. Chem.* 55 (2012) 6094–6110.
- [21] T. Wiseman, S. Williston, J.F. Brandts, L.N. Lin, Rapid measurement of binding constants and heats of binding using a new titration calorimeter, *Anal. Biochem.* 179 (1989) 131–137.
- [22] W.B. Turnbull, A.H. Daranas, On the value of c : can low affinity systems be studied by isothermal titration calorimetry? *J. Am. Chem. Soc.* 125 (2003) 14859–14866.
- [23] W.B. Turnbull, B.L. Precious, S.W. Homans, Dissecting the cholera toxin-ganglioside GM1 interaction by isothermal titration calorimetry, *J. Am. Chem. Soc.* 126 (2004) 1047–1054.
- [24] Dr. W. Bruce Turnbull, Divided we fall? Studying low-affinity fragments of ligands by ITC, http://www.gelifesciences.com/gehcls_images/GELS/Related%20Content/Files/1314823637792/litdoc28987044_20140311181217.pdf (2011).
- [25] J. Broecker, C. Vargas, S. Keller, Revisiting the optimal c -value for isothermal titration calorimetry, *Analytical Biochemistry* 418 (2011) 307–309.
- [26] S. Keller, C. Vargas, H. Zhao, G. Piszczek, C.A. Brautigam, P. Schuck, High-Precision Isothermal Titration Calorimetry with Automated Peak-Shape Analysis, *Anal. Chem.* 84 (2012) 5066–5073.
- [27] J.C. Houtman, P.H. Brown, B. Bowden, H. Yamaguchi, E. Appella, L.E. Samelson, P. Schuck, Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: Application to adaptor protein complexes in cell signaling, *Protein Sci.* 16 (2007) 30–42.
- [28] J. Stürzebecher, U. Stürzebecher, H. Vieweg, G. Wagner, J. Hauptmann, F. Markwardt, Synthetic inhibitors of bovine factor Xa and thrombin comparison of their anticoagulant efficiency, *Thromb. Res.* 54 (1989) 245–252.
- [29] M. Dixon, The graphical determination of K_m and K_i , *Biochem. J.* 129 (1972) 197–202.
- [30] M.J. Bennett, S.I. Blaber, I.A. Scarisbrick, P. Dhanarajan, S.M. Thompson, M. Blaber, Crystal structure and biochemical characterization of human kallikrein 6 reveals that a trypsin-like kallikrein is expressed in the central nervous system, *J. Biol. Chem.* 277 (2002) 24562–24570.

3 Fragment Binding Can Be Either More Enthalpy-Driven or Entropy-Driven: Crystal Structures and Residual Hydration Patterns Suggest Why.

3.1 Einleitende Bemerkung/Introductory remarks

Das nachfolgende Kapitel wurde im Journal of Medicinal Chemistry publiziert. Die Clusteranalyse der Seitenkettenkonformationen von Glutamat 192 wurde von Michael Betz durchgeführt. Die Messung der thermodynamischen Daten und Erstellung der Röntgenstrukturen und des Manuskripts erfolgte durch den Autor dieser Dissertation.

3.2 Abstract

For lead optimization it has been suggested to start with small enthalpically advantaged binders as an entropic binding component will be added inevitably during late-stage optimization. Methods giving access to thermodynamic signatures of initially weak binding fragments are therefore essential to support the decision-making process which fragment to take to further optimization. High-resolution crystal structures of six fragments binding to the S1 pocket of the serine protease thrombin were determined and analyzed with respect to the thermodynamic binding profile observed for the bound fragments. The two most potent fragments exhibiting an amidine-type scaffold are not the most enthalpic binders, likely related to their unfavorable desolvation costs. They show very similar binding modes including the residual solvation pattern seen in the binding pocket. Their thermodynamic signatures also display a similar partitioning in enthalpy and entropy. Two chemically very similar chloroaromatic fragments differ strongly in their potency (430 μM vs. 10 mM); their binding modes are related, but the surrounding residual water network differs. The more potent and strongly enthalpic binding fragment recruits a water molecule and involves Glu192 in binding, thus succeeding to firmly cap the S1 pocket from above. Obviously, fragments exhibiting a rather perfect solvation pattern in the adopted binding pose also experience the highest potency. The residual solvation pattern takes impact on the observed thermodynamic signature. Displacement titrations, also performed to record the thermodynamic profiles of the fragments, suggest non-additivity in the displacement

process resulting from differences in the water displacement or entrapment in the S1 pocket.

3.3 Introduction

The consideration of thermodynamic binding profiles in hit and lead optimization is increasingly relevant as a valuable analytical tool in drug discovery. [1-3] Analyses of thermodynamic data of ligand binding have shown that during late-stage optimization improved binding is frequently achieved by enhancing the entropic component to the Gibbs free energy of binding. [4,5] In many cases this results either from an appropriate rigidification of the lead scaffold in the protein-bound conformation, or from the attachment of lipophilic groups of growing size to optimally fill remaining unoccupied pockets in the binding site. Usually these strategies make the lead candidates more complex, provoking concomitant problems such as unsatisfactory bioavailability or growing risk of undesired toxicity. [6-8] Accordingly, it has been recommended to start lead optimization with small enthalpically advantaged binders as the entropic component will be added inevitably to the binding signature during late-stage optimization. [9,10] Therefore, methods giving reliable access to thermodynamic signatures, particularly of weak binding initial hits such as fragments, are essential to support the decision-making process which compounds to take further into lead optimization. [11,12]

In this study we investigated the important drug target thrombin, a serine protease that converts soluble fibrinogen into insoluble strands of fibrin, as well as activating and catalyzing many other coagulation-related reactions. Its active site consists of three well-shaped pockets: The S1, S2 and S3/4 pocket. We investigated a series of fragments binding to the S1 pocket (**Figure 3.1**). For the characterization of their binding profiles we applied isothermal titration calorimetry and used a set of more potent displacement ligands to determine their thermodynamic signature. The reference ligands used for displacement titrations address all three pockets.

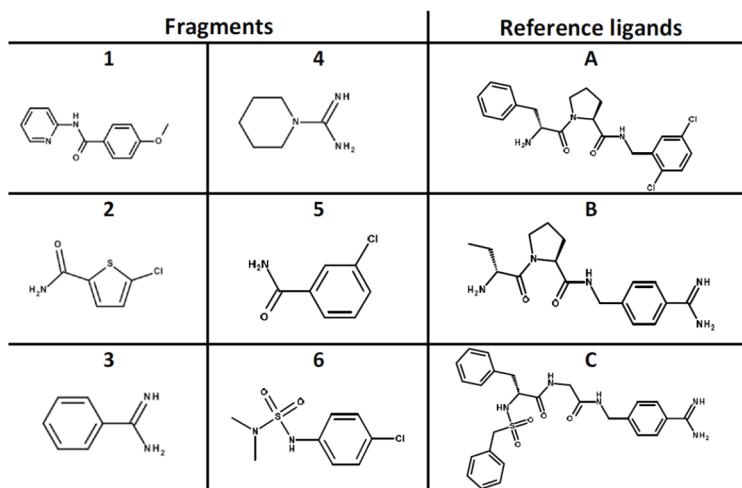


Figure 3.1 Chemical structures of the investigated fragments and reference ligands.

In our previous study [13] we showed that fragments binding to the S1 pocket of thrombin can be titrated either directly under so called low *c*-value conditions (no sigmoidal curve) [14–16] or indirectly by use of a strongly binding ligand displacing the low affinity fragment that had already been bound to the protein in a pre-incubation step. [17] We could show that the Gibbs free energy of binding is determined quite reliably and the obtained values are independent of the applied titration protocol, rendering the method valid for dissociation constant determinations. In contrast, the characterization of enthalpy and entropy is more difficult. Even though the displacement method achieves higher accuracy, the obtained enthalpy-entropy profile depends on the properties of the displacement ligand used in the second titration step. Although the absolute values deviate, the relative enthalpy differences determined across the experiments using the different displacement ligands revealed an unchanged signature and thus can serve as a kind of thermodynamic fingerprint for the fragments' binding signatures.

In the current contribution we will present the high-resolution crystal structures of thrombin in complex with the studied fragments and the reference ligands used for the displacement titrations. The uncovered structural details particular with respect to the residual solvation patterns, will be compared to the results obtained by the deviating thermodynamic signatures and the differences recorded by the displacement titrations.

3.4 Results

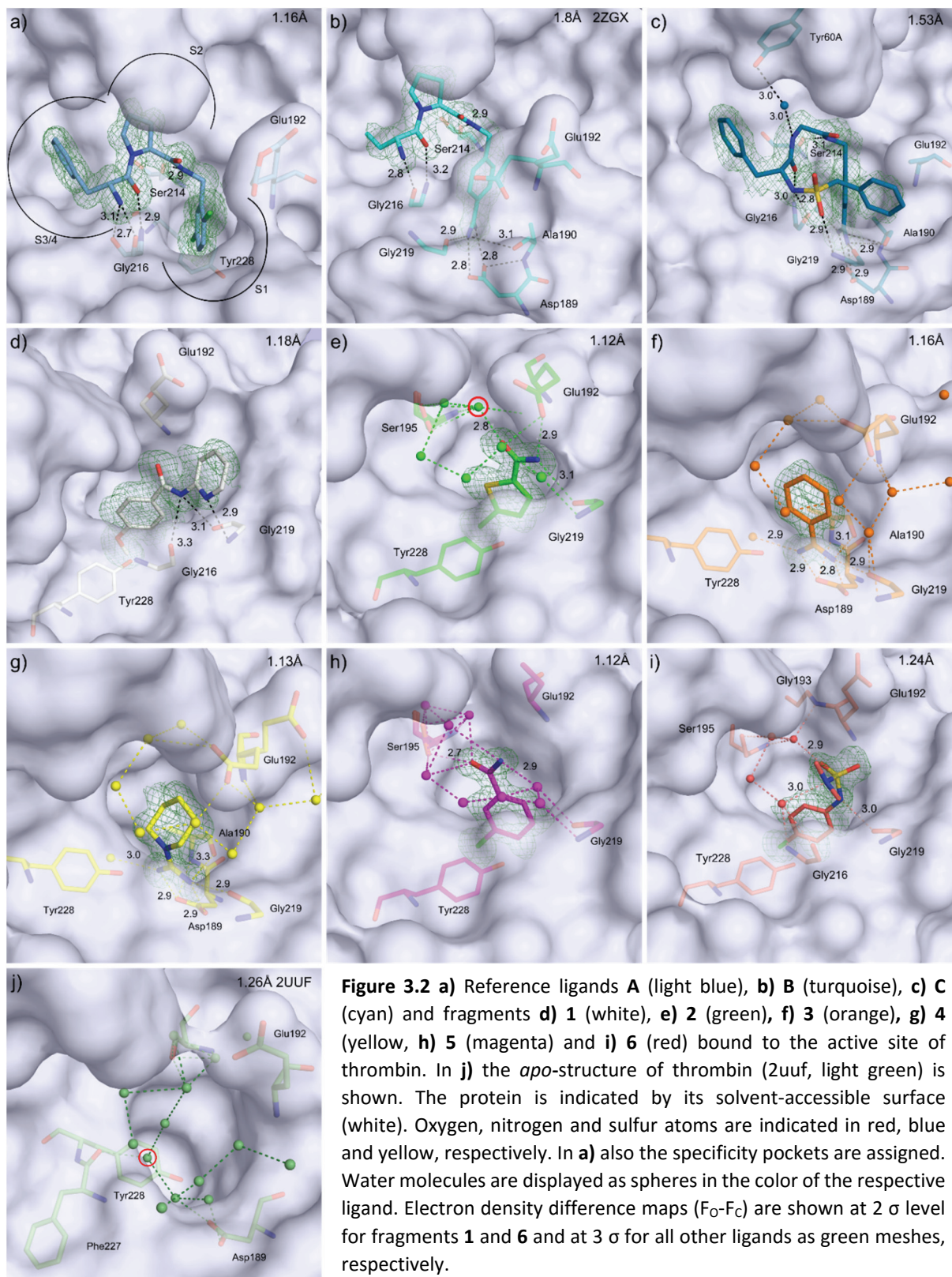
3.4.1 Crystal structures of the reference ligands

Reference ligand **A** exhibits a predominantly enthalpic binding signature and addresses the S1 pocket of thrombin with its dichlorobenzyl moiety, the S2 pocket with a proline and the S3/S4 pocket with a D-Phe portion (**Figure 3.2a**). At the P1 position the 2,5-dichlorobenzylamide portion forms a chlorine- π interaction with Tyr228. The 2-chlorine atom points along the C-Cl bond vector to the edge of the aromatic ring of Tyr228 at a distance of 4.9 Å. No additional water molecule can be detected in the S1 pocket. Furthermore, a hydrogen bond is formed from the NH group of this P1 group to the carbonyl oxygen of Ser214 (2.9 Å). The proline portion at P2 fills nicely the space beneath the 60-loop covering the S2 pocket. Accordingly, it prevents solvent exposure of the ligand's central amide NH which is hydrogen-bonded to Ser214CO. The adjacent carbonyl function (2.9 Å) and the N-terminal amino group (2.7 Å and 3.1 Å) of the D-Phe P3 portion form a dual β -ladder-type H-bonding motif to the backbone amide group of Gly216 which has been refined in two conformations. The Glu192 residue, spatially adjacent, also adopts two conformations with comparable occupancy pointing toward the solvent.

Reference ligand **B** is a less enthalpic binder and addresses the S1 pocket of thrombin with its benzamidine moiety, the S2 pocket with a proline and the S3 pocket with a *homo*-Ala residue (**Figure 3.2b**). The crystal structure of reference ligand **B** (PDB 2ZGX) has been discussed previously. [18] As for reference ligand **A**, Glu192 is present in two conformers with almost equal occupancy. Furthermore, a water molecule is found on top of the aromatic ring of Tyr228 in the S1 pocket.

Reference ligand **C** shows balanced enthalpic and entropic binding contributions. Similar to reference ligand **B**, it addresses the S1 pocket of thrombin using its benzamidine moiety (**Figure 3.2c**). It forms a salt bridge with distances (both 2.9 Å) to the deprotonated Asp189 at the bottom of this pocket. Furthermore, the water molecule observed in the complex with **B** is also present above Tyr228. In the S2 pocket, the NH of a glycine residue addresses

the hydroxyl moiety of Tyr60A via a water mediated H-bond, both with equal distances of 3.0 Å. The S3 pocket is occupied by the ligand's D-Phe residue and the S4 pocket is filled by a terminally attached benzylsulfonamide group. The sulfonyl moiety forms a hydrogen bond (2.9 Å) to the backbone NH of Gly219. Glu192 is disordered over several orientations leading to a poor electron density for this sidechain which has thus not been modeled into the electron density. Due to the benzylsulfonamide in P4 position which is extending toward the S1 pocket, Glu192 is most likely pointing toward the solvent as already observed in related structures. The binding mode of related inhibitors has also been described in a previous publication. [19]



3.4.2 Crystal structures of the fragments

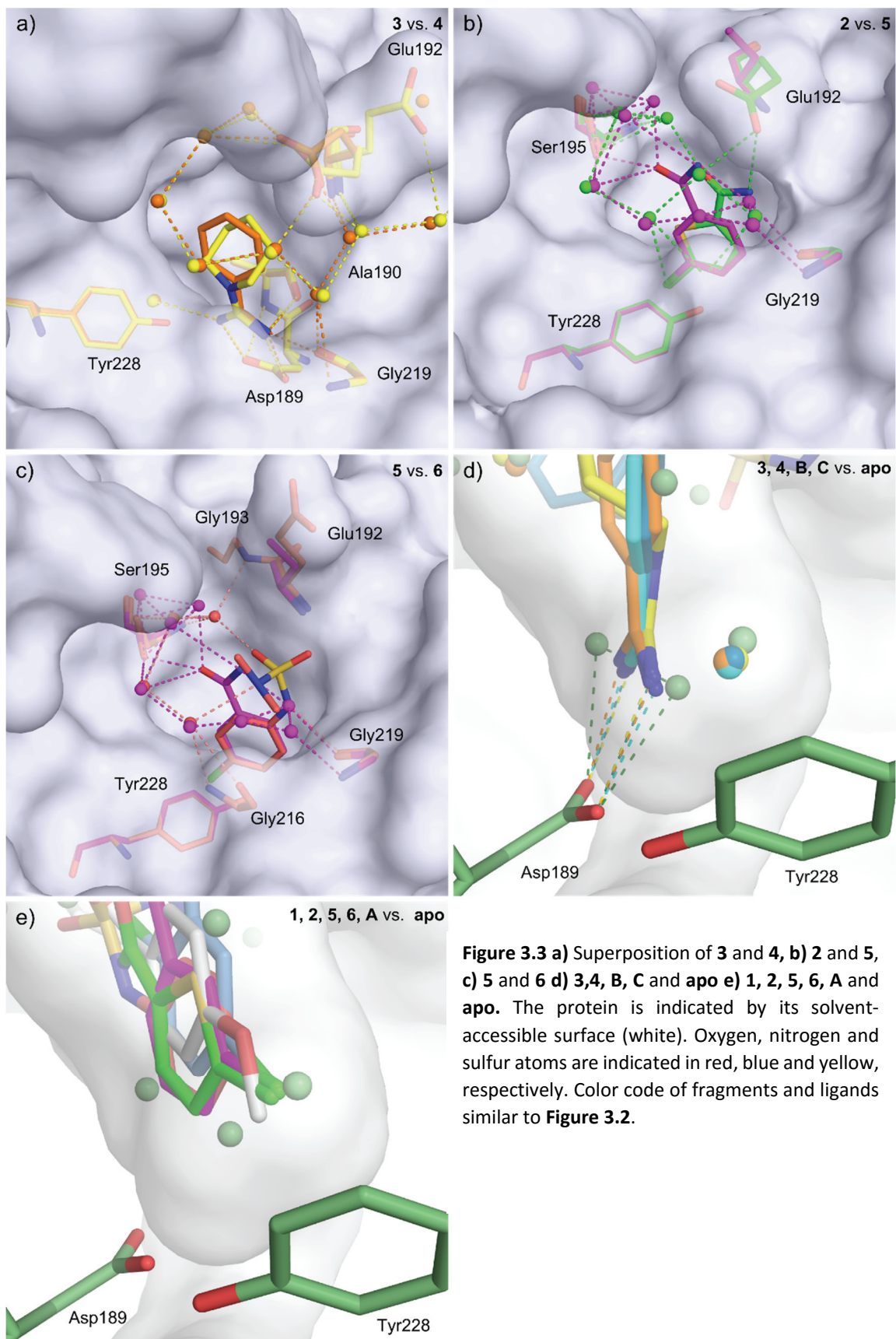
Fragment **1** (**Figure 3.2d**) establishes a bifurcated hydrogen bond via its amide nitrogen to the carbonyl oxygens of Gly216 (3.3 Å) and Gly219 (3.1 Å). Both hydrogen bonds are characterized by a rather large distance and angular relationships indicate strong deviations from ideal linearity ($\text{N-H}\cdots\text{O}=\text{C}$; with 128° and 130°). The pyridine nitrogen of **1** forms an additional polar contact to Gly219 (2.9 Å). The mutual arrangement of the involved groups does not suggest the formation of an energetically favorable H-bond ($\text{N-H}\cdots\text{N}$) as it deviates strongly from linearity (126°). In contrast to fragments **2-6** no further direct interactions with the protein in the deeply buried S1 pocket could be recognized. Fragment **1** could be refined with full occupancy at a resolution of 1.18 Å. Glu192 is observed in one single conformation and orients towards the solvent.

Fragment **2** forms a chlorine- π interaction with Tyr228 deep in the S1 pocket where the chlorine atom of the chlorothiophene points toward Tyr228 (**Figure 3.2e**). This chlorine- π interaction in protein-ligand complexes is characterized by an “edge on” geometry, as the carbon-chlorine bond vector points toward a C-C bond of the adjacent aromatic portion. [20] Furthermore, hydrogen bonds are formed between the carboxamide group of **2** and Glu192 (2.9 Å) as well as Gly219 (3.1 Å). One of the carboxylate oxygens of Glu192 is placed in a position suited to address the carboxamide nitrogen of the fragment allowing to establish a favorable H-bonding geometry (162° to Gly219 and 173° to Glu192). As a consequence, Glu192 adopts one definite conformation. This contact, along with the support from reasonably refined B-values, suggest the unambiguous assignment of O and N to the electron density of the carboxamide group. In addition, an interstitial water molecule (red circle, **Figure 3.2e**) mediating hydrogen-bonding contacts between the carboxamide oxygen of **2** and the backbone of amino acids Cys191, Gly193 (amino acids not shown in **Figure 2e**) and both alternative conformations (occupancy 0.42 and 0.58) of Ser195 is observed.

The amidino moiety of fragment **3** (**Figure 3.2f**) is accommodated in the S1 pocket of thrombin and forms a salt bridge with distances of 2.9 and 2.8 Å to the most likely

deprotonated Asp189 at the bottom of this pocket. Planes defined by the atoms of either the ligand's amidino and the carboxylate group of Asp189 are twisted by 14.5° from coplanarity, which is assumed to represent the optimum for this contact. This contrasts the binding geometry observed for **4** (see below). In anti-direction the amidino nitrogens of **3** form H-bonds to the backbone carbonyl oxygen of Ala190 (3.1 Å) and Gly219 (2.9 Å). A water molecule mediating an H-bond to the carbonyl oxygen of Phe227 (not shown in **Figure 3.2f**) can be detected and Glu192 is observed in one conformation which partly seals the S1 pocket from above. Beyond the phenyl ring of **3** toward the solvent environment an elaborate water network is detected, consisting of seven water molecules and involving the carboxylate group of Glu192.

Via its amidino group, the amidinopiperidine fragment **4** (**Figure 3.2g**) forms a salt bridge with comparable distances as in **3** (2.9 and 2.9 Å) to the deprotonated Asp189 at the bottom of the S1 pocket. Here the twist angle between planes through the amidino and carboxylate groups approaches more closely the optimal coplanarity compared to **3** (3.6°). Further H-bonds can be observed from the amidino nitrogens to the backbone oxygen atoms of Ala190 (3.3 Å) and Gly219 (2.9 Å). The H-bond to Ala190 is longer compared to **3** (3.1 Å). The ligand's piperidine ring adopts a chair conformation fitting well below one of the two observed Glu192 conformers which refine to almost equal occupancies of 49% and 51%. Whereas one conformer is integrated into the water network capping the S1 pocket, the second conformer is oriented toward the solvent. None of the conformers is involved in a direct contact to other residues of the protein. Remarkably, both show an orientation and accordingly torsion angles that deviate from the other structures. Similar to **3** the corresponding water-mediated H-bond to Phe227 could be detected and the water molecule is located also at the same position as in the complex with **3**. The generated water network exhibits similar features as the one observed for **3**, expect that Glu192 occurs in a second conformation and only the first one is involved in the network. The water network above the S1 pocket and **3** and **4** is largely conserved (**Figure 3.3a**).



Fragment **5** (resolution: 1.12 Å) refines to an occupancy of 0.80. It also forms a chlorine- π interaction to Tyr228 in the S1 pocket and the carbon-chlorine bond vector points towards a C-C bond of the aromatic ring of Tyr228 (**Figure 3.2h**). In this structure Glu192 is disordered and no interactions to **5** can be detected. A hydrogen bond is formed between the carbonyl group of the fragment's amide group and one conformer (occupancy 0.54) of the hydroxy methylene side chain of the catalytic Ser195 (2.7 Å). Here, the assignment of N and O to the density of the fragment's carboxamide groups is not unequivocally defined by the H-bonding topology, but the B-factor refinement supports our allocation. In addition, water-mediated hydrogen bonds are formed via the NH₂ group of the amide function of **5** and the carbonyl oxygens of the backbone amide groups of residues Gly216 and Gly219.

For fragment **6** (resolution: 1.24 Å) a chlorine- π interaction to Tyr228 is also observed in the S1 pocket since the chlorine atom of the *para*-chloro-benzyl portion points towards Tyr228 (**Figure 3.2i**). An additional hydrogen bond is formed via the NH group of the sulfamide unit to the carbonyl group of Gly219 (3.0 Å). Glu192 is oriented towards the solvent, as similarly observed for fragment **1**. In addition, an interstitial water molecule mediates a hydrogen bond between one oxygen of the sulfamide group of **6** and the backbone amide groups of Gly193 and Ser195. Furthermore, the terminal dimethyl amino group of the sulfamide is involved in a water-mediated contact to the peptide backbone of Gly216. Fragment **6** refines to an occupancy of 0.75. In the difference electron density weak density signals attributable to several water molecules at the position of the phenyl ring of **6** became visible suggesting partial occupancy of **6** and corresponding to the water structure of the *apo*-protein. Due to the low occupancy, these water molecules were not included in the refinement model.

One water molecule (O1279; *apo*-structure **2UUF** [21] red circle **Figure 3.2j**) is known to be important for the binding of different P1 head groups in the S1 binding pocket of thrombin. It is found in the benzamidine (**B**, **C**, **3**) and amidino piperidine (**4**) complexes (**Figure 3.3d**). [18] The location of this water molecule seems to be optimal in the *apo*-protein as it builds up a network with other water molecules to interact with Asp189.

Furthermore, it forms a short contact to water molecules placed on top of the aromatic ring of Tyr228. Interestingly, the chlorine substituents of **2**, **5**, **6**, **A** and also the methoxy moiety of **1** are located close to the latter water position above the aromatic ring of Tyr228 and displace it from the S1 pocket (**Figure 3.3e**). With the amidine-type ligands this water molecule remains in the S1 pocket but it is pushed into a slightly more distal position compared to the *apo*-protein.

3.4.3 Influence of bound hirudin peptide on the thermodynamic signature

All crystal structures presented in this study originate from crystals that were prepared in the presence of a 12 residue long peptide derived from the 66 amino acid natural product inhibitor hirudin in order to block the *exo*-site of thrombin. Even though the binding site of this peptide is rather remote from the active site where all fragments and the reference ligands bind, it cannot be excluded that the binding of the peptide takes some influence on their binding properties. In contrast to the crystallographic analysis all titration experiments discussed in our previous communication were performed in the absence of the peptide. Accordingly, to study the influence of the hirudin peptide used in the crystallization procedure, we titrated the peptide directly to thrombin and performed additional titrations with the three displacement ligands **A**, **B**, and **C** in the presence of the hirudin peptide in order to compare them with the obtained thermodynamic data in presence and absence of the *exo*-site binder.

A direct titration of the hirudin peptide to thrombin revealed no significant enthalpic signal that could be evaluated, likely due to a predominantly entropic binding profile. Subsequently, we titrated the three displacement ligands to the protein preincubated with the hirudin peptide. The measured binding enthalpy for ligand **A** (with hirudin $\Delta H^\circ = -47.0 \pm 1.1$ kJ/mol, without hirudin $\Delta H^\circ = -46.5 \pm 0.3$ kJ/mol), **B** (with hirudin $\Delta H^\circ = -27.6 \pm 0.7$ kJ/mol, without hirudin $\Delta H^\circ = -29.9 \pm 0.6$ kJ/mol) and **C** (with hirudin $\Delta H^\circ = -21.0 \pm 0.2$ kJ/mol, without hirudin $\Delta H^\circ = -22.4 \pm 0.9$ kJ/mol) show almost identical thermodynamic signatures within experimental accuracy (**Figure 3.4**). Hence, the presence of the hirudin peptide

seems not to take a measurable impact on the thermodynamic properties of the binding of our active-site inhibitors.

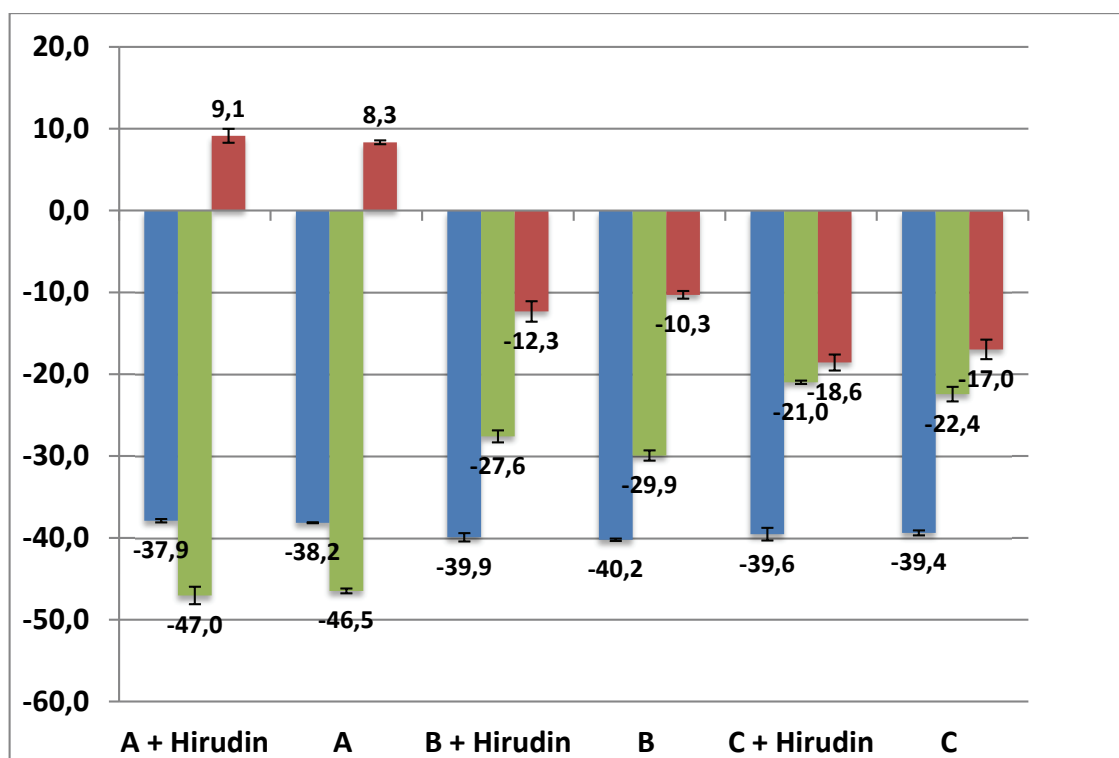


Figure 3.4 Thermodynamic parameters (ΔG° in blue, ΔH° in green, $-T\Delta S^\circ$ in red in kJ/mol) determined by direct ITC titrations for the ligands **A**, **B** and **C** in the presence of Hirudin and without. The estimated standard deviations were calculated from duplicate measurements (to save hirudin material) are indicated by the black error bars.

3.4.4 Analysis of the conformational flexibility of Glu192

A survey of a large number of high-resolution thrombin structures in the PDB shows that Glu192 is quite flexible and adopts multiple conformations in the various crystal structures.

We analyzed the different orientations of glutamic acid 192 (Glu192) in 174 thrombin crystal structures. All 174 thrombin structures with a resolution better than 2 Å were taken from the PDB, and extended by our eight new structures discussed in this paper. The complex with **B** has already been deposited and is comprised in the retrieved pdb set. A list of all considered pdb codes can be found in **Figure 3.5**.

In total, 13 structures were removed from the data set due to major deviations in the backbone, or cases missing Glu192 side chains: 1HAG, 2A0Q, 2AFQ, 2GP9, 2ZO3, 3BEI, 3DUX, 3HK3, 3JZ1, 3K65, 3SQE, 5AFY (structure **5**), 5AFZ (structure **C**). Only 12 structures of the sample set show a second alternative conformations of the Glu192 side chain; one had even a third orientation. In total our analysis covers 187 side chain conformers.

A distance matrix was calculated based on the rmsd of all side chain atoms. For visualization, a hierarchical cluster analysis using the average linkage method was performed. The distance criteria for the separation into distinct clusters was a dissimilarity of 1.7 Å (**Figure 3.5**) shows the dendrogram used for the hierarchical clustering of the 187 Glu192 side chains entries. **Figure 3.6** displays the distribution of the Glu192 side chain orientations in stick representation and colored according to their cluster affiliation.

The Glu192 side chain found in the complex (1A4W) with a thiazole-containing inhibitor probing the S1' pocket is structurally remote to all other entries and remains as a singleton (shown in orange). It appears puzzling that a rather large B value of 50.0 Å² has been assigned to this residue in the structure 1A4W. According to this analysis, Glu192 is quite flexible and can adopt, depending on the bound ligand, different conformations.

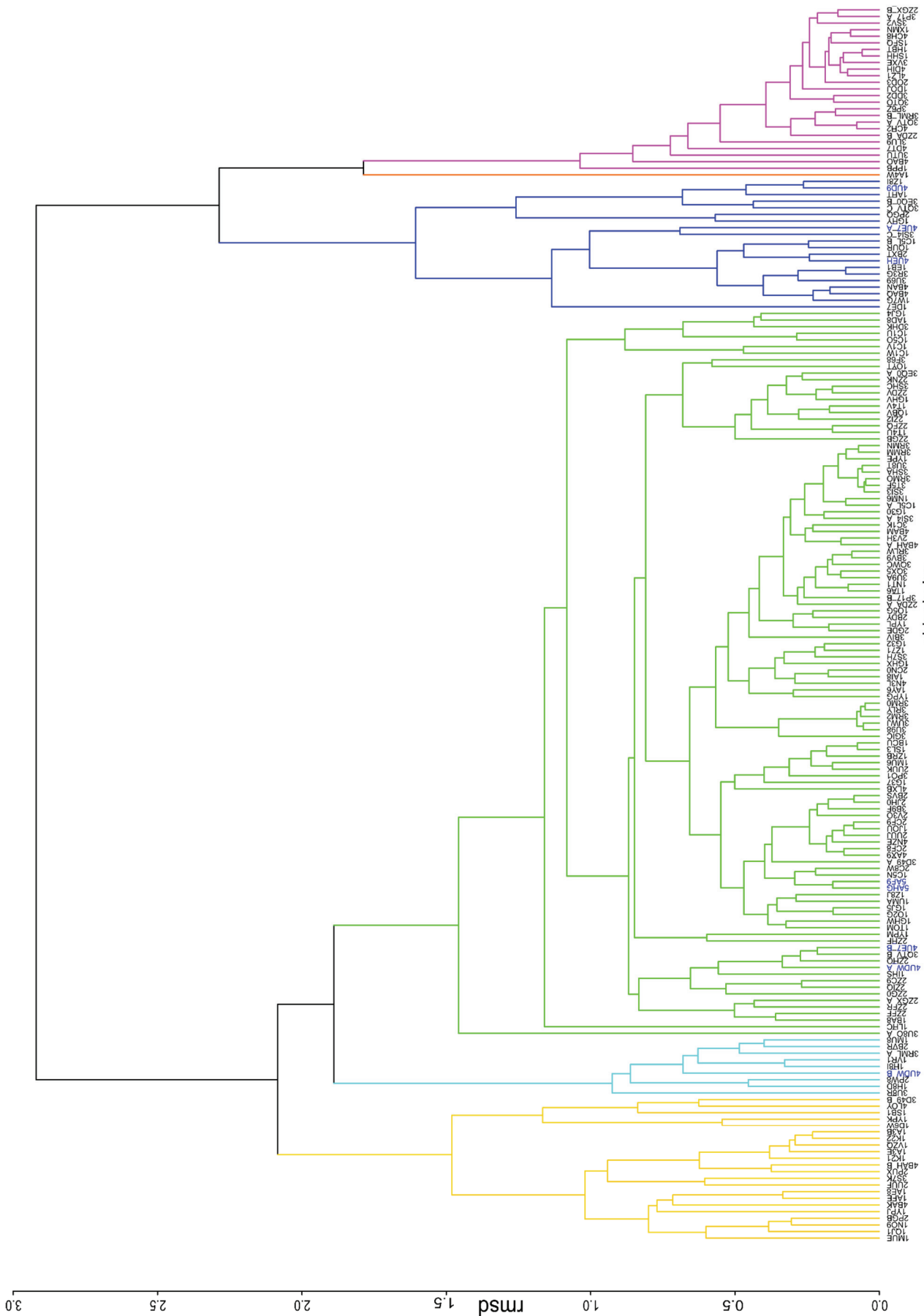


Figure 3.5 Cluster dendrogram
Rectangles around the branches highlight the clusters colored corresponding to the side chains in Figure 3.6. The structures discussed in this paper are labeled blue in the dendrogram sidechains

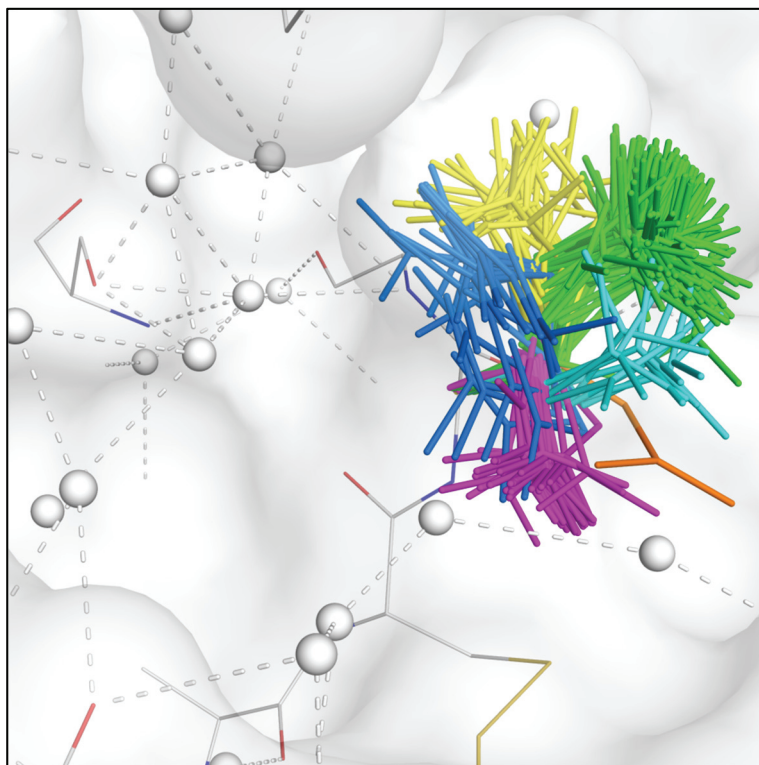


Figure 3.6 Distribution of Glu192 side chain orientations as sticks colored by cluster number. The surface of the thrombin *apo*-structure (2UUF), carbon atoms of the amino acids, and water molecules are depicted in white.

3.5 Discussion

The thermodynamic characterization of the reference ligands **A-C** and the four fragments **1-4** has been described in our previous contribution (chapter two). [13] Fragments **5** and **6** are characterized by too low affinities to record their thermodynamic signature and accurate affinity data. According to a fluorescence-based assay [13] we estimate that their binding must be about 10 mM. Thus, both are by 5-100 fold less potent than the other fragments studied. A summary of the binding data of **A-C** and **1-4** is given in **Figure 3.7**.

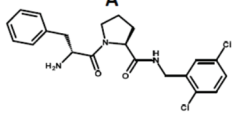
Compound	ITC Protocol	ΔG°	ΔH°	$-T\Delta S^{\circ}$	K_i Assay	K_d ITC
	direct	-37.5 ± 0.5	-45.1 ± 0.8	7.6 ± 0.5	0.06 ± 0.02	0.27 ± 0.05
	direct	-39.7 ± 0.4	-28.8 ± 0.8	-10.9 ± 0.6	0.18 ± 0.03	0.11 ± 0.02
	direct	-38.8 ± 0.6	-20.7 ± 0.9	-18.1 ± 1.1	0.14 ± 0.07	0.16 ± 0.03

Figure 3.7 a) Chemical structures of reference ligands, thermodynamic parameters (ΔG° in blue, ΔH° in green, $-T\Delta S^{\circ}$ in red in kJ/mol; the size of the columns indicates the absolute value of the parameter) determined by direct ITC titrations for the strong binding reference ligands **A**, **B** and **C** and their estimated binding constants in μM are shown. The estimated standard deviations were calculated from at least triplicate measurements.

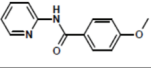
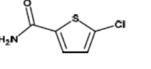
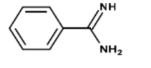
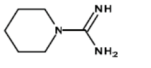
Fragment	ITC Protocol	ΔG°	ΔH°	$-T\Delta S^{\circ}$	K_i Assay	K_d ITC displacement	K_d ITC direct
	direct	-17.2 ± 0.6	-15.7 ± 4.1	-1.5 ± 4.7	2430 ± 309	1582 ± 441	1003 ± 224
	disp. A	-16.7 ± 1.6	-18.4 ± 6.4	1.7 ± 6.6			
	disp. B	-15.8 ± 1.1	-31.8 ± 7.0	16.0 ± 7.1			
	disp. C	-16.2	-20.7	4.5			
	direct	-19.0 ± 0.6	-29.7 ± 2.0	10.7 ± 2.3	431 ± 27	507 ± 139	475 ± 110
	disp. A	-18.2 ± 0.6	-36.3 ± 2.5	18.1 ± 2.6			
	disp. B	-18.9	-28.8	9.9			
	disp. C	-19.6 ± 0.7	-36.7 ± 1.9	17.1 ± 2.1			
	direct	-19.8 ± 0.2	-13.9 ± 1.3	-5.9 ± 1.5	258 ± 1	455 ± 109	355 ± 29
	disp. A	-18.7 ± 0.6	-4.3 ± 1.0	-14.4 ± 1.2			
	disp. B	-19.0 ± 1.2	-17.4 ± 2.9	-1.6 ± 3.1			
	disp. C	-19.3 ± 0.7	-16.6 ± 2.1	-2.7 ± 2.1			
	direct	-22.8 ± 0.4	-15.4 ± 2.1	-7.4 ± 2.1	111 ± 2	197 ± 74	107 ± 29
	disp. A	-22.0 ± 0.6	-15.9 ± 1.7	6.1 ± 1.9			
	disp. B	-20.9 ± 0.7	-25.7 ± 1.6	4.8 ± 3.1			
	disp. C	-21.5	-20.7	-0.8			

Figure 3.7 b)⁴ Chemical formulas of fragments **1-4** with their thermodynamic parameters derived from different titration protocols (direct titrations and displacement titrations with ligands **A**, **B** and **C**) in kJ/mol and their estimated binding constants in μM . In case the enthalpies were taken as similar to the binding enthalpy of the reference ligand (**C-1**, **B-2** and **C-4**) no error estimations are indicated.

Independent of the selected titration strategy (direct/displacement) or the chosen reference ligand, the free energy of binding is determined with rather small deviations in all cases, rendering ITC as a reliable method to determine ΔG and dissociation constants. The characterization of the enthalpy/entropy partitioning is more difficult and seems to be dependent on the applied titration protocol. It has to be kept in mind that in ITC enthalpy

⁴ **Figure 3.7 b)** contains a corrected summary of thermodynamic data according to footnote 1 on page 29.

and entropy do not result from independent measurements, they are calculated as the numerical difference of ΔG and ΔH . We therefore have to consider with respect to the relative comparisons only one of the two properties ΔH or $-T\Delta S$. To reduce the influence of possibly superimposed protonation effects all titration experiments were performed in pyrophosphate buffer which exhibits a small heat of ionization. Multiple titrations of ligands to thrombin have indicated that His57 generally releases approximately 0.3 - 0.5 mol of protons upon ligand binding. [18, 19] For ligands with a scaffold like **A** and **B** entrapment of a similar molar quantity of protons has been recorded, compensating the proton release from His57. Thus, for the studied fragments some small uncertainties of the thermodynamic profile can be given owing to superimposed and not corrected protonation effects.

Interestingly, quite consistent thermodynamic profiles are found for fragment **2** by all titration protocols. In the case of **1** this conclusion is difficult to make as the accuracy of the measurements is limited due to the low potency of this fragment. Thus difference in the profiles resulting from the deviating protocols have to be discussed with some care. Nonetheless, the data suggest that this ligand is a predominantly enthalpic binder. Considering the direct titration, both amidine-type fragments exhibit a similar profile, although **4** is more potent owing to both an enthalpic and entropic advantage over **3**. The displacement titrations provide a more complex picture. Here the reference ligand **A** featuring a chloroaromatic P1 portion reveals for both fragments **3** and **4** an entropically more favored profile, whereas the benzamidine derivatives **B** and **C** result in stronger enthalpically favored binding. The absolute values differ for both fragments, however the relative entropic advantage using **A** compared to the enthalpic components using **B** and **C** are similar (≈ 8 - 10 kJ/mol) (**Figure 3.7b**).

3.5.1 Methoxy fragment 1

The methoxy group of **1** falls very close to the location of the chlorine substituent found in **A**, **2**, **5** and **6** (**Figure 3.3e**). In all ligands the attached substituents are able to displace the water molecule located at this position on top of the aromatic moiety of Tyr228 in the *apo*-

protein (PDB Codes: 2UUF, 3D49) and the two amidino complexes. The methoxy group attached to fragment **1** is also found in the marketed factor Xa inhibitor Apixaban rendering this substituent as competent to collect sufficient affinity contribution in the S1 pocket of serine proteases. [22] Fragment **1** has been refined with full occupancy and results in a B-factor ratio between fragment and protein ($B_{\text{fragment}}/B_{\text{protein}} = 1.24$) which is the highest within the studied series. This indicates that either the occupancy is below 100% and/or this fragment exhibits a significant residual mobility in the pocket due to less tight fixation. This observation would suggest a reduced enthalpic signature in favor of an enhanced entropic contribution. At least compared to **2** that also displaces the named water molecule (s. below), it is indeed a less enthalpic binder.

3.5.2 Amidine fragments 3 and 4

The conformationally more flexible amidinopiperidine **4** loses a larger amount of degrees of freedom upon binding as it restricts to only one distinct chair conformation in the S1 pocket. This should be entropically unfavorable. In comparison, the nearly planar scaffold of benzamidine allows at most some wiggling motions of the amidino group with respect to the adjacent aromatic portion and part of this mobility will still be experienced at the binding site which should be entropically beneficial. Additionally, the binding of **4** allows establishment of two conformations of Glu192 whereas for **3** only one conformation could be observed. Possibly the latter effect overwhelms and provides an explanation for the higher entropic binding contribution of **4** (cf. below comparison **2** and **5**). Both molecules show a comparable B factor ratio of the fragment compared to the surrounding protein ($B_{\text{fragment}}/B_{\text{protein}}$ for **3** = 0.87 and **4** = 0.85) and comparably low maximum-likelihood based coordinate errors (0.09 Å for **3** and 0.07 Å for **4** see **Table 3.1**). A closer analysis of the anisotropically refined B-factors unveils that the part of the benzamidine ring pointing toward Gly219 is moving more strongly out of plane, leading most likely to a higher residual mobility of the fragment. Planes assigned to the ligand's amidino and the carboxylate group of Asp189 exhibit a stronger mutual twist from coplanarity for **3** compared to **4**. It is tempting to speculate whether this observation results in a weaker salt bridge formed to Asp189 in case of **3**, particularly as **3** is also the enthalpically less favorable binder

compared to **4**. Overall **4** shows more potent binding to thrombin which stimulated the researchers at Roche to use the amidinopiperidine moiety as a P1 head group in the development compound napsagatran. [23] Taking the data from direct titrations (as they are not further affected by superimposed implications of the displacement ligands) as the basis to compare the thermodynamic signatures of **4** over **3**, the affinity advantage of **4** seems to result from both an enthalpic and entropic benefit.

Interestingly, a direct comparison of the water networks formed above the S1 pocket shows no substantial differences (**Figure 3.3a**), and will hardly contribute to the differences in the thermodynamic profiles of **3** and **4**.

3.5.3 Chlorine-substituted fragments 2, 5 and 6

The less potent fragment **2** shows interestingly an even lower $B_{\text{fragment}}/B_{\text{protein}}$ ratio of 0.82 compared to **3** and **4**, whereas the very weakly binding fragments **1**, **5** and **6** show significantly higher ratios (see **Table 3.1**). The properties of **2** might be determined by a fairly tight fixation of the terminal carboxamide group in a network of well-established H-bonds between Glu192, Gly219 and, mediated by two water molecules, to Ser195OH and its backbone NH. This fairly rigid fixation in combination with the enthalpically favored formation of the chlorine- π interaction owing to the replacement of a water molecule which is located in the *apo*-protein on top of Tyr228 has to be compared with the binding pose of **5**, a fragment binding by 1-2 orders of magnitude less strongly. The chloroaromatic portion accommodates virtually the same position in both structures (**Figure 3.3b**). Due to the deviating geometry of the five and six-membered rings, the exit vectors orienting the carboxamide groups differ strongly. In consequence, the carboxylate of Glu192 is not recruited to interact with **5** in contrast to **2**, it remains disordered and likely oriented toward the solvent. However, as modeling attempts show, it could potentially form an H-bond to the ligand with about 3.1 Å length. On first glance, the surrounding water network, apart from the involvement of Glu192, seems to be similar in both complexes, but in detail it appears geometrically less ideal in case of **5**. The water molecule bridging a putative contact to Glu192 is lacking. In the complex with **2** this water molecule succeeds to firmly

cap the S1 pocket from above. Obviously these differences in the solvation pattern and the contact to Glu192 enhance the potency of **2** significantly, suggesting this fragment as superior to **5**.

In **5** the exit vector for the carboxamide is located *meta* with respect to the chlorine substituent. In **6**, the second weak binding fragment, *para*-topology is realized. Accordingly, its polar sulfamide group, even better competent to establish more polar contacts, only establishes a hydrogen bond with Gly219 and via an interstitial water molecule a contact to Ser195OH and its backbone NH. Due to its larger steric demand, the sulfamide group perturbs the water network more strongly compared to that in **2** and **5**. The inability of **6** to recruit Glu192 for binding, likely owing to repulsive interactions, the stronger perturbation of the capping water network and the higher costs for at least partial desolvation of the sulfamide group make this fragment overall a weaker binder.

All ligands containing a chloroaromatic P1 portion replace, with respect to the *apo*-protein, at least four well-ordered water molecules from the S1 pocket. Interestingly enough, the amidine derivative **3** and **4** as well as the reference ligands **B** and **C** all show one residual water molecule in S1, occupying the position of the chlorine atom in **2**, **5**, **6**, and **A** (**Figures 3.3e** and **3.3d**). This water molecule, found on top of the aromatic ring of Tyr228, is slightly shifted in the complexes compared to its position adopted in the *apo*-structure. The desolvation cost required to shed the hydration shell around the charged amidino group in **3**, **4**, **B**, and **C** will be significantly larger than for those with the chloroaromatic P1 group. Desolvation is predominantly enthalpically unfavorable and will therefore reduce the enthalpic binding signal. [12] Remarkably, the chloro-thiophene anchor used in **2** has also been applied as P1 head group in the marketed factor Xa inhibitor rivaroxaban. [24]

3.5.4 *Deviating thermodynamic profiles obtained using different reference ligands*

We suppose that we record with these displacement titration experiments another case of non-additivity in protein-ligand binding. In case every individual step in a displacement titration procedure would be at equilibrium, the thermodynamic signature obtained for

the displacement of the bound fragments should be independent of the properties of the applied stronger binding reference ligand.

A displacement titration starts by incubating the protein with the probe fragment. Subsequently, the selected reference ligand is added and repels the fragment from the pocket. If intermediately the binding pocket would return to the unperturbed solvation pattern observed for the *apo*-protein, the displacement titration should reveal - independent of the used displacement ligand - the same results. This is due to the fact that in the evaluation the results of the titration using the reference ligand alone are subtracted from the combined titration protocol. However, within experimental accuracy this is not the case, we obtain deviations using different reference ligands. This suggests that the protein does not return intermediately to the unperturbed solvation structure seen in the *apo*-protein.

The chloro-thiophene fragment **2** displaces all four well-ordered water molecules found in the S1 pocket of the *apo*-protein. Displacement by the chloroaromatic reference ligand **A** occurs without changing the solvation pattern in the S1 pocket. Displacement by the two benzamidine-type reference ligands **B** and **C** involves, however, the displacement of **2** from the fully desolvated pocket and the benzamidine reference ligand enters the site together with a water molecule. Thus, the latter ligand does not fully release its solvation shell. This situation is different if the amidine-type fragments **3** and **4** are displaced from the S1 pocket. They incorporate already the water molecule on top of Tyr228. This has to be repelled once the pocket accommodates the chloroaromatic reference ligand **A**. This release is likely combined with an entropic advantage. In contrast the displacement of **3** and **4** by the benzamidine-type reference ligand **B** and **C** does not change the solvation pattern in S1, the water molecule on top of Tyr228 remains. Obviously, this shifts the thermodynamic signature relative to the displacement with **A** toward a more enthalpically favored process of about 8-10 kJ/mol. It remains interesting to see that the displacement of **3** and **4** obtains the same relative changes in the enthalpy/entropy inventory. But compared on an absolute scale a difference is experienced. Obviously, additional effects are overlaid to the process, resulting from further changes of the protein and/or the

solvation structure, possibly involving the conformational transitions of Glu192. This underlines again that thermodynamic signatures should not be compared on an absolute scale but only relative to each other. [12] The example shows that the mutual displacement of ligands in binding pockets can follow a rather complex and not necessarily an additive protocol. Instead cooperativity resulting from differences in the solvation pattern can be given.

3.6 Conclusion

High-resolution crystal structures of six fragments binding to the S1 pocket of thrombin were determined and analyzed with respect to the thermodynamic binding profile observed for the bound fragments. Two rather weak binding fragments (**5**, **6**) address the S1 pocket via their chloroaromatic portion, the remaining fragment moieties do not succeed to establish strong interactions with the enzyme. In consequence, they show significant residual mobility and their weak potency makes a detailed thermodynamic characterization impossible. Another chloro-thiophene fragment (**2**) is similarly recognized in the S1 pocket through a chloro- π interaction with Tyr228. However, additional strong interactions beyond the S1 pocket and an extensive water network capping the fragment's binding position turn this fragment into the most enthalpic binder of the series and renders **2** a much better binder compared to the highly isostructural **5**. Part of the enthalpic advantage of **2** is compensated by an unfavorable positive entropic contribution making this fragment the third most potent of the series. A fourth fragment (**1**) exhibiting a methoxy aromatic P1 head group binds with less affinity to the protein, however its binding is also dominated by enthalpy. In addition to its S1 binding, this fragment experiences several polar contacts with the protein beyond the S1 pocket. Important enough, all four fragments displace four well defined water molecules found in the S1 pocket of the *apo*-structure of thrombin, particularly including the water molecule located on top of the aromatic ring of Tyr228. The two most potent fragments studied are of the amidino-type molecules **3** (benzamidine) and **4** (amidinopiperidine). The latter fragment turns out to be the best binder of the series owing to an enthalpic and entropic advantage over **3**. For both charged fragments a huge enthalpic desolvation price will be required

which should be much larger than for the other four fragments. [12] This cost will, however, be overcome by the strong salt bridge formed to Asp189. Compared to the *apo*-structure, the water molecule on top of Tyr228 remains in the pocket, while it is slightly shifted out of the *apo*-position. In addition, both fragments allow for a perfect water network capping the S1 pocket even though they are not involved in any further polar contacts.

The side chain of Glu192 which exhibits large conformational flexibility across the large variety of thrombin crystal structures adopts different configurations in the six fragment complexes, partly involving its carboxylate group into interactions with the fragments. Supposedly, also this residue, in addition to the residual solvation structure, takes an impact on the thermodynamic signature of the fragments. Remarkably, the fragments exhibiting a rather perfect solvation pattern in the adopted binding pose also experience the highest potency. Likely in the case of fragment binding the influence of the quality and completeness of the water network wrapping around the bound fragment is even more determining for the thermodynamic signature than in case of drug-size ligand binding. [12]

This observation is further emphasized by the fact that displacement titrations using different reference ligands produce deviating absolute signatures for the fragments being replaced. The displacement of amidine-based fragments by a benzamidine-type ligand does not alter the S1 solvation pattern, whereas a chloroaromatic ligand used for displacement will induce a change of the solvation pattern in the S1 pocket as the water on top of Tyr228 is repelled. These changes are superimposed to the overall titration and have to be regarded. Overall we can conclude that at least a semi-quantitative estimate on the thermodynamic signature of the fragments is possible, however, a detailed control of the involved structural changes is mandatory, particular with respect to the residual solvation structure. Only then the assignment of a fragment to be a more enthalpic or entropic starting point for subsequent lead optimization makes sense. However, for such weak binders, a modulation of the solvation structure can easily shift the thermodynamic signature from “more enthalpic” to “more entropic”.

3.7 Experimental Section / Materials & Methods

3.7.1 ITC measurements

ITC experiments were performed using an ITC200™ system from GE Healthcare, Northampton, MA, USA. Details about the applied measuring protocols have been described elsewhere. [13]

Thrombin was obtained from CSL Behring (Marburg, Germany) and purified from Beriplast®. Thrombin was extracted by dialysis using an experimental buffer of 50 mM tetra sodium pyrophosphate, 100 mM NaCl, 0.1 % polyethylene glycol 8000 at pH 7.8. Subsequently, the protein could be used for all titration experiments. The protein concentration was measured by absorbance at 280 nm using a NanoDrop 2000c spectrophotometer from Thermo Scientific.

For studying the influence of bound hirudin, 250 µM hirudin peptide (54-65) (sulfated) from Bachem (Switzerland) in TSP buffer was preincubated with 27.5 µM thrombin and afterwards titrated. Displacement ligands from 50 mM DMSO stock solutions were diluted to 0.25 mM in ITC buffer and subsequently the DMSO concentration in all solutions was adjusted to 3 %. All titrations were performed in duplicate in order to save hirudin material.

All ITC experiments were performed at 25 °C with a reference power of 5 kcal/s after a stable baseline had been reached. The pre-titration delay was set to 300 s. For titrations, ligand injections of 0.3 µL (to prevent artefacts arising from small syringe leakages or air in the syringe) were followed by 19 to 27 injections of 1.5 – 2.0 µL with at least a 180 s interval between each injection. Data integration and evaluation was performed with Nitpic [25] and Sedphat [26].

3.7.2 X-ray Data

3.7.2.1 Crystallization and soaking

Human α -thrombin (from Enzyme Research Laboratories, South Bend, USA) was dissolved in the crystallization buffer (20 mM NaH₂PO₄, 350 mM NaCl, 2 mM benzamidine, pH 7.5)

at 10 mg/ml. The hirudin fragment Hirudin (54-65) (sulfated) obtained from Bachem (Bubendorf, Switzerland) was dissolved in the crystallization buffer at 2.5 mg/mL. In the next step, 40 μ L of the solution of the hirudin fragment was mixed with 160 μ L of the thrombin solution. After incubation for 1 h at 4°C, crystallization was carried out at 4°C by the hanging-drop method. The hirudin/thrombin solution was mixed 1:1 with the reservoir solution (20 mM NaH₂PO₄, 27 % polyethylene glycol 8000, pH 7.5) and 2 μ L of this solution were placed to the center of a cover slide. Immediately after mixing of protein and reservoir buffer microseeding was performed with a horse hair. The wells of the crystallization trays were filled with 500 μ L of the reservoir buffer. Subsequently the cover slips were placed over the wells and sealed. Crystals of good diffracting quality could be obtained after 10 to 14 days.

For soaking, a DMSO stock solution of the inhibitors (**A**, **B** and **C** (50 mM) and fragments **1-6** (500 mM)) was diluted 1:10 with a solution containing crystallization and reservoir buffer at a 1:1 ratio resulting in the final soaking concentration containing 5 mM or 50 mM of the respective compound and 10 % DMSO. A crystal without visible imperfections was selected and transferred into the soaking solution for 24 h.

3.7.2.2 Data collection and processing

For data collection at 110 K, crystals were prepared using a cryoprotectant solution of 20 % glycerol in reservoir buffer for two minutes.

The data sets of **1**, **2**, **5** and **A** were collected at the synchrotron DESY (Hamburg) at beamline P14, a wavelength of 0.976 Å and a Pilatus 6M detector. Data sets of **3**, **4**, **5** and **C** have been collected at the synchrotron BESSY (Berlin) at beamline 14.2 (wavelength of 0.9184 Å) and beamline 14.3 (wavelength of 0.895 Å), both using Rayonix MX-225 detectors.

For compounds **1**, **2**, **3**, **5**, **6**, **A** and **C** data processing and scaling was performed with the XDS program package. [27] For fragment **4** data processing and scaling was performed using the HKL2000 package. [28]

3.7.2.3 Structure determination and refinement

The coordinates of human thrombin (PDB codes 1H8D [29] and 4UE7) were used for molecular replacement with Phaser from the CCP4 program package. [30]

For initial rigid body refinement of the protein molecule, followed by repeated cycles of maximum likelihood energy minimization simulated annealing and B-factor refinement the program PHENIX [31] was used. For determining the temperature factors for the structures anisotropic B-factor refinement was applied. A randomly chosen 5 % set of all data was used for the calculation of R_{free} and not used during the refinement. Amino acid side chains were fit into σ -weighted $2F_o - F_c$ and $F_o - F_c$ electron density maps using Coot. [32] After the first refinement cycle, water molecules and subsequently ions and ligands were located in the electron density and added to the model. Restraints were applied to bond lengths and angles, planarity of aromatic rings and van der Waals contacts. Multiple side chain conformations were built in case an appropriate electron density was observed and maintained during the refinement, and if the minor populated side chain showed at least 20% occupancy. The final models were validated using PHENIX own validation options and Coot. The Ramachandran plot was calculated with MolProbity [33] integrated in the PHENIX package. Data collection, unit cell parameters and refinement statistics are given in **Table 3.1**. Analysis of temperature factors was performed with MolProbity. The naming of the protein amino acids was done according to Bode et al. [29]

3.7.2.4 Protein Data Bank and accession codes

Coordinates and structure factors have been deposited in the Protein Data Bank with the following accession codes: **1** (5AF9), **2** (4UD9), **3** (4UEH), **4** (4UE7), **5** (5AFY), **6** (5AHG), **A** (4UDW) and **C** (5AFZ).

Fragment Binding Can Be Either More Enthalpy-Driven or Entropy-Driven: Crystal Structures and Residual Hydration Patterns Suggest Why.

<i>Compound (PDB-Code)</i>	<i>1 (5AF9)</i>	<i>2 (4UD9)</i>	<i>3 (4UEH)</i>	<i>4 (4UE7)</i>
Wavelength (Å)	0.976	0.976	0.9184	0.9184
Resolution range (Å)^a	70.96 - 1.18 (1.22 - 1.18)	35.68 - 1.12 (1.17 - 1.12)	35.76 - 1.16 (1.20 - 1.16)	19.56 - 1.13 (1.15 - 1.13)
Space group	C 1 2 1	C 1 2 1	C 1 2 1	C 1 2 1
Unit cell a, b, c (Å) α, β, γ (°)	70.0 71.3 72.1 90 100.4 90	70.1 71.4 72.1 90 100.4 90	70.4 71.5 72.5 90 100.5 90	70.3 71.5 72.4 90 100.3 90
Total reflections	497434 (40339)	804825 (42252)	444822 (35244)	392964 (19572)
Unique reflections	108698 (9826)	127839 (10054)	116712 (10231)	130988 (6524)
Multiplicity	4.6 (4.1)	6.3 (4.2)	3.8 (3.4)	3.0 (2.7)
Completeness (%)	95.1 (86.1)	96.7 (76.3)	95.7 (84.3)	99.5 (99.5)
Mean I/sigma(I)	19.6 (2.2)	21.8 (3.1)	17.7 (2.5)	24.5 (2.5)
Wilson B-factor	13.4	12.7	13.6	10.7
R-merge	0.033 (0.565)	0.036 (0.414)	0.032 (0.440)	0.038 (0.390)
R-work	0.121 (0.223)	0.125 (0.186)	0.121 (0.204)	0.124 (0.189)
R-free	0.137 (0.247)	0.141 (0.197)	0.138 (0.214)	0.139 (0.200)
Number of non-hydrogen atoms	2778	2815	2810	2816
macromolecules	2434	2465	2447	2434
ligands	58	58	22	55
water	280	291	331	321
Protein residues	289	294	289	291
RMS(bonds) (Å)	0.011	0.009	0.011	0.01
RMS(angles) (°)	1.45	1.38	1.46	1.45
Ramachandran favored (%)^b	97	97	98	97
Ramachandran outliers (%)^b	0	0	0	0
Ramachandran allowed (%)^b	3	2	2	3
Average B-factor (Å³)	20.6	20.1	22.0	17.9
macromolecules	19.2	18.9	20.3	16.0
ligands	26.8	27.8	22.9	20.0
solvent	31.8	28.4	33.8	30.9
B_{fragment} / B_{protein}	1.24	0.81	0.87	0.85
Coordinate errors (Å)^c	0.09	0.07	0.09	0.07

^a Statistics for the highest-resolution shell are shown in parentheses.

^b Calculated by MOLPROBITY

^c Maximum-likelihood based

Table 3.1 Data collection and refinement statistics.

<i>Compound (PDB-Code)</i>	<i>5 (5AFY)</i>	<i>6 (5AHG)</i>	<i>A (4UDW)</i>	<i>C (5AFZ)</i>
Wavelength (Å)	0.9184	0.976	0.976	0.895
Resolution range (Å)^a	18.45 - 1.12 (1.16 - 1.12)	43.41 - 1.24 (1.28 - 1.24)	71.13 - 1.16 (1.20 - 1.16)	21.47 - 1.53 (1.58 - 1.53)
Space group	C 1 2 1	C 1 2 1	C 1 2 1	C 1 2 1
Unit cell a, b, c (Å)	70.4 71.5 72.6	70.2 71.3 72.2	69.6 71.1 72.3	70.3 71.6 72.4
α, β, γ (°)	90 100.5 90	90 100.5 90	90 100.4 90	90 100.5 90
Total reflections	417074 (37162)	326876 (29013)	506370 (36860)	201127 (19424)
Unique reflections	134540 (13272)	95099 (8909)	116096 (10454)	53119 (5257)
Multiplicity	3.1 (2.8)	3.4 (3.3)	4.4 (3.5)	3.8 (3.7)
Completeness (%)	99.3(98.5)	96.1 (90.2)	97.1 (87.9)	99.8 (99.4)
Mean I/sigma(I)	25.7 (2.5)	14.1 (2.2)	25.9 (2.7)	17.7 (2.8)
Wilson B-factor	11.4	13.3	13.2	17.0
R-merge	0.040 (0.396)	0.043 (0.528)	0.025 (0.472)	0.048 (0.527)
R-work	0.122 (0.186)	0.123 (0.229)	0.124 (0.214)	0.145 (0.176)
R-free	0.138 (0.190)	0.144 (0.250)	0.141 (0.230)	0.174 (0.191)
Number of non-hydrogen atoms	2780	2659	2808	2645
macromolecules	2416	2294	2456	2383
ligands	37	61	41	43
water	325	300	309	218
Protein residues	290	280	290	287
RMS(bonds) (Å)	0.011	0.011	0.012	0.006
RMS(angles) (°)	1.45	1.43	1.51	1.04
Ramachandran favored (%)^b	97	97	97	98
Ramachandran outliers (%)^b	0	0	0	0
Ramachandran allowed (%)^b	3	3	3	2
Average B-factor (Å²)^c	18.9	19.9	21.0	25.0
macromolecules	17.1	17.9	19.5	24.3
ligands	23.1	28.8	16.8	26.3
solvent	31.8	32.8	33.3	32.0
B_{fragment} / B_{protein}	0.98	1.18	-	-
Coordinate errors (Å)^c	0.08	0.11	0.08	0.12

^a Statistics for the highest-resolution shell are shown in parentheses.

^b Calculated by MOLPROBITY

^c Maximum-likelihood based

Continuation Table 3.1 Data collection and refinement statistics.

3.8 References

- [1] J.E. Ladbury, G. Klebe, E. Freire, Adding calorimetric data to decision making in lead discovery: a hot tip, *Nat Rev Drug Discov* 9 (2010) 23–27.
- [2] D.C. Rees, M. Congreve, C.W. Murray, R. Carr, Fragment-based lead discovery, *Nat Rev Drug Discov* 3 (2004) 660–672.
- [3] J.D. Taylor, P.J. Gilbert, M.A. Williams, W.R. Pitt, J.E. Ladbury, Identification of novel fragment compounds targeted against the pY pocket of v-Src SH2 by computational and NMR screening and thermodynamic evaluation, *Proteins* 67 (2007) 981–990.
- [4] T.S.G. Olsson, M.A. Williams, W.R. Pitt, J.E. Ladbury, The thermodynamics of protein-ligand interaction and solvation: insights for ligand design, *J. Mol. Biol.* 384 (2008) 1002–1017.
- [5] G.G. Ferenczy, G.M. Keserű, How are fragments optimized? A retrospective analysis of 145 fragment optimizations, *Journal of medicinal chemistry* 56 (2013) 2478–2486.
- [6] J.B. Chaires, Calorimetry and thermodynamics in drug design, *Annual review of biophysics* 37 (2008) 135–151.
- [7] V. Borsi, V. Calderone, M. Fragai, C. Luchinat, N. Sarti, Entropic contribution to the linking coefficient in fragment based drug design: a case study, *Journal of medicinal chemistry* 53 (2010) 4285–4289.
- [8] S.F. Martin, J.H. Clements, Correlating structure and energetics in protein-ligand interactions: paradigms and paradoxes, *Annual review of biochemistry* 82 (2013) 267–293.
- [9] G.E. de Kloe, D. Bailey, R. Leurs, de Esch, Iwan J P, Transforming fragments into candidates: small becomes big in medicinal chemistry, *Drug discovery today* 14 (2009) 630–646.
- [10] M. Congreve, G. Chessari, D. Tisi, A.J. Woodhead, Recent developments in fragment-based drug discovery, *Journal of medicinal chemistry* 51 (2008) 3661–3680.
- [11] E. Freire, A thermodynamic approach to the affinity optimization of drug candidates, *Chemical biology & drug design* 74 (2009) 468–472.
- [12] G. Klebe, Applying thermodynamic profiling in lead finding and optimization, *Nature reviews. Drug discovery* 14 (2015) 95–110.
- [13] E. Rühmann, M. Betz, M. Fricke, A. Heine, M. Schäfer, G. Klebe, Thermodynamic signatures of fragment binding: Validation of direct versus displacement ITC titrations, *Biochimica et biophysica acta* 1850 (2015) 647–656.

- [14] J. Broecker, C. Vargas, S. Keller, Revisiting the optimal c value for isothermal titration calorimetry, *Analytical biochemistry* 418 (2011) 307–309.
- [15] J. Tellinghuisen, Isothermal titration calorimetry at very low c , *Analytical biochemistry* 373 (2008) 395–397.
- [16] W.B. Turnbull, A.H. Daranas, On the value of c : can low affinity systems be studied by isothermal titration calorimetry? *Journal of the American Chemical Society* 125 (2003) 14859–14866.
- [17] Y.L. Zhang, Z.Y. Zhang, Low-affinity binding determined by titration calorimetry using a high-affinity coupling ligand: a thermodynamic study of ligand binding to protein tyrosine phosphatase 1B, *Analytical biochemistry* 261 (1998) 139–148.
- [18] B. Baum, L. Muley, A. Heine, M. Smolinski, D. Hangauer, G. Klebe, Think twice: understanding the high potency of bis(phenyl)methane inhibitors of thrombin, *Journal of molecular biology* 391 (2009) 552–564.
- [19] A. Biela, F. Sielaff, F. Terwesten, A. Heine, T. Steinmetzer, G. Klebe, Ligand binding stepwise disrupts water network in thrombin: enthalpic and entropic changes reveal classical hydrophobic effect, *Journal of medicinal chemistry* 55 (2012) 6094–6110.
- [20] Y.N. Imai, Y. Inoue, I. Nakanishi, K. Kitaura, Cl- π interactions in protein-ligand complexes, *Protein science : a publication of the Protein Society* 17 (2008) 1129–1137.
- [21] H.U. Ahmed, M.P. Blakeley, M. Cianci, D.W.J. Cruickshank, J.A. Hubbard, J.R. Helliwell, The determination of protonation states in proteins, *Acta crystallographica. Section D, Biological crystallography* 63 (2007) 906–922.
- [22] D.J.P. Pinto, M.J. Orwat, S. Koch, K.A. Rossi, R.S. Alexander, A. Smallwood, P.C. Wong, A.R. Rendina, J.M. Luettgen, R.M. Knabb, K. He, B. Xin, R.R. Wexler, P.Y.S. Lam, Discovery of 1-(4-methoxyphenyl)-7-oxo-6-(4-(2-oxopiperidin-1-yl)phenyl)-4,5,6,7-tetrahydro-1H-pyrazolo[3,4-c]pyridine-3-carboxamide (apixaban, BMS-562247), a highly potent, selective, efficacious, and orally bioavailable inhibitor of blood coagulation factor Xa, *Journal of medicinal chemistry* 50 (2007) 5339–5356.
- [23] K. Hilpert, J. Ackermann, D.W. Banner, A. Gast, K. Gubernator, P. Hadvary, L. Labler, K. Muller, G. Schmid, T.B. Tschoop, Design and synthesis of potent and highly selective thrombin inhibitors, *Journal of medicinal chemistry* 37 (1994) 3889–3901.
- [24] S. Roehrig, A. Straub, J. Pohlmann, T. Lampe, J. Pernerstorfer, K.-H. Schlemmer, P. Reinemer, E. Perzborn, Discovery of the novel antithrombotic agent 5-chloro-N-((5S)-2-oxo-3-[4-(3-oxomorpholin-4-yl)phenyl]-1,3-oxazolidin-5-yl)methyl)thiophene-2-carboxamide (BAY 59-7939): an oral, direct factor Xa inhibitor, *Journal of medicinal chemistry* 48 (2005) 5900–5908.

- [25] S. Keller, C. Vargas, H. Zhao, G. Piszczek, C.A. Brautigam, P. Schuck, High-precision isothermal titration calorimetry with automated peak-shape analysis, *Analytical chemistry* 84 (2012) 5066–5073.
- [26] J.C.D. Houtman, P.H. Brown, B. Bowden, H. Yamaguchi, E. Appella, L.E. Samelson, P. Schuck, Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: application to adaptor protein complexes in cell signaling, *Protein science : a publication of the Protein Society* 16 (2007) 30–42.
- [27] W. Kabsch, XDS, *Acta crystallographica. Section D, Biological crystallography* 66 (2010) 125–132.
- [28] Z. Otwinowski, W. Minor, [20] Processing of X-ray diffraction data collected in oscillation mode, in: *Methods in Enzymology*, vol. 276, *Macromolecular Crystallography Part A*, Elsevier, 1997, pp. 307–326.
- [29] W. Bode, I. Mayr, U. Baumann, R. Huber, S.R. Stone, J. Hofsteenge, The refined 1.9 Å crystal structure of human alpha-thrombin: interaction with D-Phe-Pro-Arg chloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion segment, *The EMBO journal* 8 (1989) 3467–3475.
- [30] The CCP4 suite: programs for protein crystallography, *Acta crystallographica. Section D, Biological crystallography* 50 (1994) 760–763.
- [31] P.D. Adams, P.V. Afonine, G. Bunkóczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.-W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, P.H. Zwart, PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta crystallographica. Section D, Biological crystallography* 66 (2010) 213–221.
- [32] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, *Acta crystallographica. Section D, Biological crystallography* 60 (2004) 2126–2132.
- [33] V.B. Chen, W.B. Arendall, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, D.C. Richardson, MolProbity: all-atom structure validation for macromolecular crystallography, *Acta crystallographica. Section D, Biological crystallography* 66 (2010) 12–21.

4 Boosting affinity by correct ligand preorganization for the S2 pocket of thrombin: A study by ITC, MD and high resolution crystal structures

4.1 Einleitende Bemerkungen/Introductory Remarks

Das nachfolgende Kapitel wurde zur Publikation vorbereitet. Die Berechnungen zur Volumenanalyse der S2 Spezifitätstasche wurden von Michael Betz durchgeführt. MD Simulationen der untersuchten Liganden wurden von Melinda Rupp durchgeführt.

Die Messung der thermodynamischen Daten, Bestimmung der Affinitätsdaten und Erstellung der Röntgenstrukturen und des Manuskripts erfolgte durch den Autor dieser Dissertation.

4.2 Abstract

Structural preorganization fixing bioactive conformations at protein-binding sites is a popular strategy to enhance binding affinity during late-stage optimization. The rationale for this enhancement relates to entropic advantages assigned to rigidified versus flexible ligands. We analyzed a narrow series of peptidomimetics binding to thrombin. The individual ligands exhibit at P2 a conformationally flexible glycine, more restricted alanine, N-methyl glycine, N-methyl *homo*-alanine and largely rigidified proline moiety. Overall, affinity boosts by a factor of 1000, partly explained by an entropic advantage. All ligands adopt the same binding mode with tiny deviations. Residual mobility of the bound ligands is reduced across the series and a protein sidechain differs in its order/disorder behavior along with changes in the surface-water network pattern established across the newly generated protein-ligand surfaces. The enthalpy/entropy inventory displays a rather complex picture and emphasizes that thermodynamics can only be compared relative to each other in close ligand series.

4.3 Introduction

A large number of natural products are based on rigid chemical scaffolds that have limited conformational variability, e.g., alkaloids and steroids.[1] With respect to their binding to a protein target, this conformational restriction has resulted in an optimal structural preorganization to fix the bioactive conformation. By this, the conformational entropy penalty to be paid for the transition from the ligand's unbound to bound state is lowered.[2] As this aspect is an important component of the observed binding affinity, the understanding of the conformational transformations on molecular level that a ligand has to undergo upon receptor binding will be an important prerequisite to rationalize structure-activity relationships and to design ligands with higher affinity.[3]

It is generally assumed that a preorganized flexible ligand when bound to a macromolecular receptor, will experience an increased binding affinity, because a rigid molecule is expected to benefit from a lesser entropic penalty as such a ligand will sacrifice a smaller amount of conformation degrees of freedom during complex formation.[4] But there is also experimental evidence which demonstrates that this belief is not universally applicable.[5] Ligand preorganization may also be accompanied by an unfavorable entropy of binding, even when the constrained ligand exhibits a higher binding affinity compared to its more flexible analogue.[6,7] Therefore, the impact that ligand preorganization can have on the energetics and structure of protein–ligand interactions must be studied carefully.[5]

The aim of this study is to examine the influence on the thermodynamic properties of inhibitors rigidified to deviating extent on their binding to the serine protease thrombin. This pharmacological important protease is involved in the regulation of blood clot formation and hemostasis, accordingly it has been addressed by synthetic inhibitors as a target for an antithrombotic therapy.[8] Its active site consists of three well-shaped pockets: The S1, S2 and S3/4 pocket. In the current study, we selected a series of potent peptidomimetic thrombin inhibitors that gradually increases hydrophobicity and sterical restrictions at the P2 position. We used isothermal titration calorimetry (ITC) and X-ray crystallography accompanied with MD simulations. The inhibitors vary systematically at the

P2 position by use of a glycine (**C**), alanine (**D**), N-methyl glycine (**E**), N-methyl *homo*-alanine (**F**) and proline (**G**) residue (**Figure 4.1**) which address the rather hydrophobic S2 pocket of the *apo*-enzyme. With respect to the remaining pockets of the protease all inhibitors exhibit the same substituents and show virtually the same binding poses. In particular, the water structure, the residual mobility and the filling of the S2 pocket of the bound ligands in their corresponding complexes along their conformational properties are essential to interpret the thermodynamic data we obtained by isothermal titration calorimetry (ITC). With this analysis, we want to ask the question: What are the contributions to binding of a rigid inhibitor molecule compared to a structurally closely related one exhibiting a less rigid scaffold.

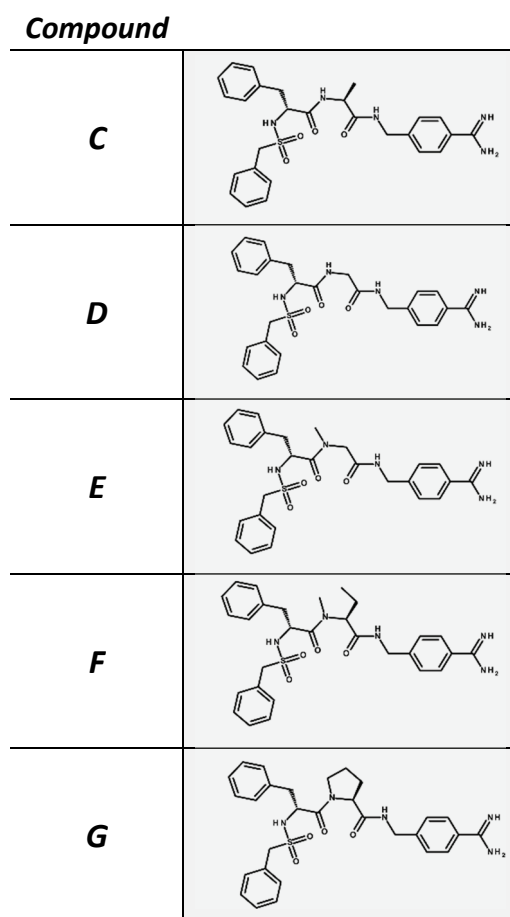


Figure 4.1: Chemical structures of ligands C-G.

4.4 Results

4.4.1 Crystal structures

As previously mentioned, the studied ligand series differs only in their chemical structure with respect to the P2 portion and therefore ligand **G** will be used to describe the key interactions shared in common by all ligands. For the remaining four ligands only the deviations from this parent reference will be reported.

Ligand **G** (resolution 1.43 Å) accommodates the S1 pocket of thrombin using its benzamidine moiety (**Figure 4.2e**). It forms a salt bridge with distances of 2.9 and 2.8 Å to the deprotonated Asp189 at the bottom of this pocket. Additional weak H-bonds are formed between the amidine NH groups and the carbonyl functions of Ala190 (3.3 and 3.1 Å) and Gly219 with a distance of 2.9 Å. A water molecule is entrapped between one of the amidino NH groups and the carbonyl oxygen of Phe227, which can be observed across a large number of benzamidine derivatives binding to thrombin. Furthermore, a hydrogen bond is formed from the NH group of this P1 group to the carbonyl oxygen of Ser214 (3.0 Å). The conformationally most restricted ligand **G** of the series shows a proline moiety at P2 to address the S2 pocket of thrombin. This P2 portion is quite common in thrombin inhibitors and it fits nicely under the so-called 60's loop (**Figure 4.2d**). The 60 loop is a specific insertion in the structure of thrombin which morphs the S2 pocket into a narrow hydrophobic crevice and features thrombin with a high substrate specificity. The P2 portion fits under this part of the 60 loop and shields the ligand's central amide NH (3.1 Å) and the adjacent carbonyl function (2.8 Å) of the D-Phe portion in P3. Both hydrogen bonds form the typical β -ladder-type binding motif to the backbone amide group of Gly216 and Trp217.[9] The S3 pocket accommodates the phenyl ring of the D-Phe residue of the ligand and the adjacent S4 pocket hosts a terminally attached benzylsulfonamide group. This part of the ligand shows some kind of back folding in a U-shape orienting towards the S1 pocket. Additionally, the sulfoxy moiety is able to form a hydrogen bond (3.0 Å) to the backbone carbonyl of Gly219. Glu192, localized adjacent to the S1 pocket, shows one well-ordered conformation in this structure pointing towards the solvent. The conformational properties

and the spatial distribution of this residue have been extensively discussed in a previous study.[10] To quantify the residual mobility of a bound ligand, we compared the B-factors of the ligand atoms with those of the surrounding protein as previously suggested by Baum et al. [9] Principally, B-factors can be strongly affected by a reduced population of the bound ligands which can deteriorate such a comparison, however, to our experience in the case of highly potent and well-soluble ligands full occupancy of the protein binding site can be assumed. Following this analysis, ligand **G** exhibits the lowest B-factor ratio $B_{\text{ligand}}/B_{\text{protein}}$ of 0.86 (**Table 4.1**) suggesting a reduced residual mobility of **5** compared to the neighboring protein residues.

The crystal structure of ligand **C**, has been presented in detail in a previous study. [10] Thus, only the differences regarding the P2 position will be shortly highlighted in this section. In contrast to the proline derivative **G**, ligand **C** (**Figure 4.2a**) possesses an NH group in the central glycine P2 residue, which addresses the hydroxyl moiety of Tyr60A above the S2 pocket via a water-mediated H-bond. Both H-bond distances measure equal distances of 3.0 Å. Apart from this important difference, **C** adopts a very similar binding pose as **G**. For ligand **C**, the B-factor ratio $B_{\text{ligand}}/B_{\text{protein}}$ amounts to 1.08 (**Table 4.1**). It displays the second largest ratio of the ligand series indicating a fairly high residual mobility of the ligand. Quite remarkable is the fact, that due to an ill-defined difference electron density, only for the structure of **C** it was not possible to assign a well-defined side-chain conformation of Glu192 localized above the S1 pocket. We even collected a second data set using a different crystal, however refinement of this data set suggested the same disorder for Glu192 in the complex with **C** (for further details, s. below).

As expected also the binding mode of **D** (resolution 1.59 Å) is very similar (**Figure 4.2b**). In P2 position **D** exhibits an alanine residue allowing the inhibitor to form a similar water-mediated H-bond to Tyr60A of the 60's loop as similarly observed for **C** with distances of 2.9 and 3.1 Å. The methyl group of the P2 alanine fits under Trp60D of the 60's loop. The remaining interactions formed to the protein backbone match well with those established by **G**. In contrast to ligand **C**, but in agreement with the complexes of **E**, **F**, and **G**, Glu192 could be refined with one better defined conformation pointing towards the solvent. The

B-factor ratio $B_{\text{ligand}}/B_{\text{protein}}$ of 1.16 (**Table 4.1**) is the highest within the series suggesting a significantly enhanced residual mobility of this ligand.

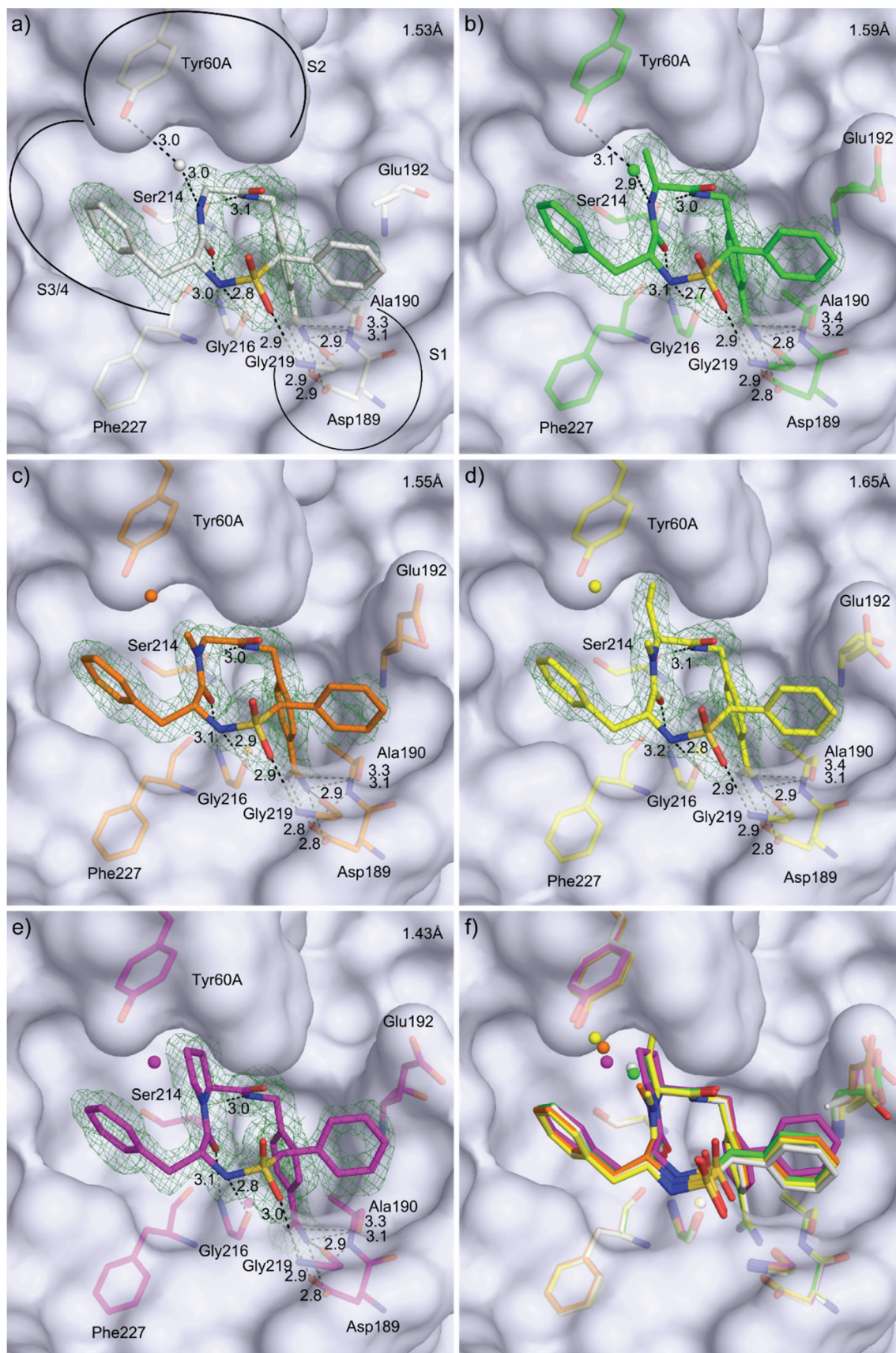


Figure 4.2 a) Ligands C (white), b) D (green), c) E (orange), d) F (yellow), e) G (magenta) bound to the active site of thrombin. f) Shows the superposition of ligand C-G. The protein is indicated by its solvent-accessible surface (white). Oxygen atoms red, nitrogen atoms blue, sulfur atoms yellow are indicated. In a) also the specificity pockets are indicated. Water molecules are shown as spheres in the color of the respective ligand. Electron density difference maps ($F_o - F_c$) are shown at 2 (σ) (ligands C, D, E and F) or 3 σ (Ligand G) as green meshes, respectively.

Ligand **E** (resolution 1.55 Å), which is an N-methyl-glycine derivative at P2, is not able to form the water-mediated H-bond to Tyr60A which was discovered with **C** and **D** (**Figure 4.2 a, b**). Nonetheless, the water molecule previously involved in the contact to Tyr60A is slightly pushed out of the S2 pocket into a new position, where it is still able to maintain an H-bond to the hydroxyl group of Tyr60A (**Figure 4.2c**).

Ligand **F** (resolution 1.65 Å) presents an N-methylated *homo*-alanine residue in P2 position (Figure 2d). It combines the structural P2 features of ligand **D** and **E**. Furthermore, it provides the same number of carbon atoms as ligand **G**. Also here the water molecule, which mediates the H-bond contact between ligand and protein in case of **C** and **D**, is positionally shifted in the S2 pocket similar to **E**, so that it can still form the H-bond to Tyr60A. The ethyl side chain at C α of the N-methylated *homo*-alanine is deeply penetrating into the S2 pocket beneath Trp60D (4.8 Å distance of C γ to ring center) and Tyr60A (4.2Å). Ligand **G**, already described above, is also alkylated at the P2 amide nitrogen, thus also the water-mediated H-bond to Tyr60AOH cannot be formed. Instead, also this ligand shows the water molecule at the shifted position, still being involved in a contact to Tyr60A. Ligand **E** and **F** show similar B-factor ratios $B_{\text{ligand}}/B_{\text{protein}}$ of 1.00 for **E** and 0.97 for **F** (**Table 4.1**).

4.4.2 Surface area and volume analysis

To assess the quality and to quantify the hydrophobic contacts formed by the P2 substituents below the 60-loop we calculated the van der Waals surface area (VSA) and the solvent accessible surface area (ASA) of the respective ligands explicitly including the hydrogen atoms. In addition, we determined the unoccupied volume below the 60-loop (**Figure 4.3 a**) and **4.3 b**) and Experimental section).

Ligand **C** has the smallest surface area (VSA= 165 Å² and ASA= 323 Å²). In this complex the unoccupied volume below the 60-loop is calculated as the largest (32 Å³) across the series. Ligands **D** (VSA= 184 Å² and ASA= 349 Å²) and **E** (VSA= 181 Å² and ASA= 342 Å²) show comparable surface areas. With respect to the unoccupied volume below the 60 loop, a difference can be recognized (**D**= 9 Å³; **E**= 22 Å³). The methyl group of the alanine derivative **D** fills the S2 pocket significantly better than the N-methylated glycine derivative **E**.

Interestingly, for ligand **F** the largest surface areas have been calculated (VSA= 215 Å² and ASA= 384 Å²), which are even larger than those of the proline derivative **G** (VSA= 203 Å² and ASA= 373 Å²). Accordingly, ligand **F** leaves hardly any volume unoccupied (2 Å³), however, only little less than that found for **G** (5 Å³).

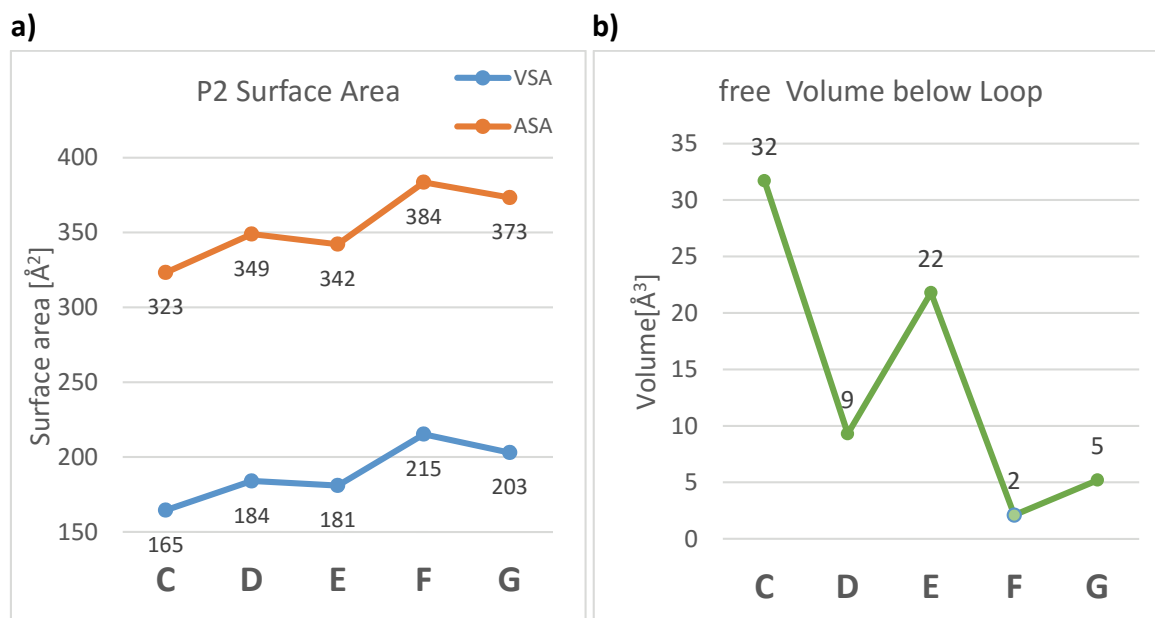


Figure 4.3: **a)** The calculated van der Waals surface area (VSA) (blue) and solvent accessible surface area (ASA) (orange) of the P2 portions of the respective ligands is plotted in Å² for ligands **C-G**. **b)** The calculated free volume below the 60's loop for the P2 portions of the respective ligands is plotted in Å³ (green).

4.4.3 ITC results

In the following, the ΔG° and ΔH° values will only be discussed, as they are the actually observed properties in ITC experiments. The entropic component, usually discussed as $-T\Delta S^\circ$, is calculated as the numerical difference of the Gibbs free energy and enthalpic binding contribution according to $\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ$. All five ligands have been titrated in a way to unravel a possibly superimposed protonation change. This can be discovered by performing the titration experiments from different buffer conditions. Therefore, the titrations of the five ligands have been carried out in three buffer systems exhibiting ionization enthalpies of deviating amount (TRIS, TRICINE and HEPES). [11]

Similar ligands from this inhibitor class have already been investigated in a previous study by us.[12] These investigations suggested a release of ca. 0.5 mole protons upon ligand binding which are liberated from His57, a residue that is partially protonated in the uncomplexed state. In the current inhibitor series, no basic terminal nitrogen is present which could putatively mask the proton release of His57. Similar to the derivative with a terminal amide group,[13] the present sulfonamide ligands **C-G** do not pick-up a proton upon binding and a net buffer dependence is observed. By correcting the measured binding enthalpies, we reveal a mean proton release from His57 to the buffers of approximately 0.5 moles (**Figure 4.4**).

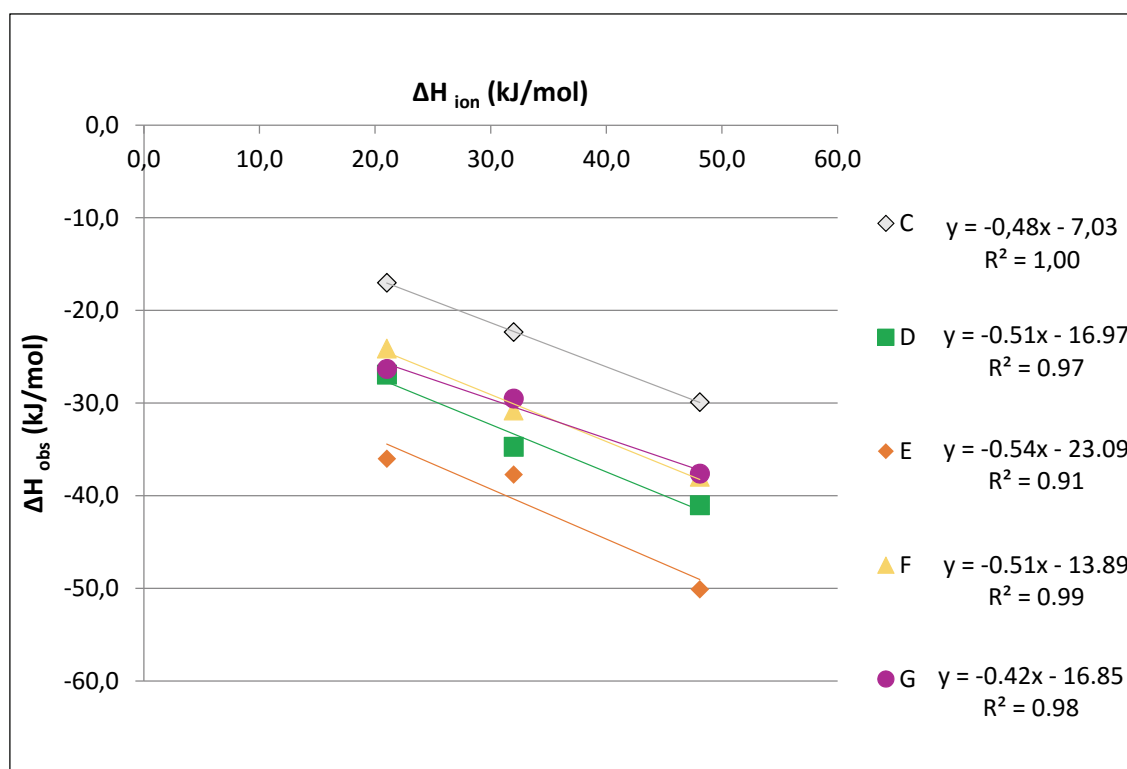


Figure 4.4: The measured heat signal is plotted against the ionization enthalpy of the buffer system (TRIS, HEPES and TRICINE) in order to extract the enthalpy of binding. Ligand **1** (white), **2** (green), **3** (orange), **4** (yellow) and **5** (magenta).

Ligand **C** ($K_i=140\pm 70$ nM; mean $\Delta G^\circ=-38.3$ kJ/mol) shows weak enthalpically ($\Delta H^\circ_{\text{ion. corr.}}=-7.03$ kJ/mol) and strong entropically ($-T\Delta S^\circ=-31.3$ kJ/mol) favored binding contribution after correction for the superimposed buffer ionization effects (**Figure 4.5**). Ligand **D** ($K_i=7 \pm 4$ nM; mean $\Delta G^\circ=-43.8$ kJ/mol) offers a stronger enthalpic ($\Delta H^\circ_{\text{ion. corr.}}=-17.0$) and a slightly

less entropic ($-T\Delta S^\circ = -26.8 \text{ kJ/mol}$) binding contribution compared to **C**. It also binds with a 20-fold higher affinity to the target. For ligand **E** ($K_i = 11 \pm 4 \text{ nM}$; mean $\Delta G^\circ = -44.8 \text{ kJ/mol}$) balanced enthalpic ($\Delta H^\circ_{\text{ion. corr.}} = -23.1 \text{ kJ/mol}$) and entropic ($-T\Delta S^\circ = -21.3 \text{ kJ/mol}$) binding contributions are observed. With a corrected binding enthalpy of -23.1 kJ/mol ligand **E** is the most enthalpic binder of the series showing a comparable binding affinity as **D** and **F**. Interestingly, the N-methylated *homo*-alanine derivative **F** ($K_i = 7 \pm 4 \text{ nM}$; mean $\Delta G^\circ = -43.0 \text{ kJ/mol}$) shows a thermodynamic signature more closely similar to that of ligand **D** than **E**, with a significantly smaller enthalpic ($\Delta H^\circ_{\text{ion. corr.}} = -13.9$) and a stronger entropic binding contribution ($-T\Delta S^\circ = -29.2 \text{ kJ/mol}$).

The rigid proline derivative **G** is by far the most potent binder of the series ($K_i = 0.21 \pm 0.04 \text{ nM}$; mean $\Delta G^\circ = -53.9 \text{ kJ/mol}$). The thermodynamic profile shows, that the binding of **G** is entropically most strongly favored ($-T\Delta S^\circ = -37.0 \text{ kJ/mol}$) compared to the other derivatives. Due to the high potency of ligand **G**, it was necessary to carry out a displacement titration [14] [15] in order to measure reliable thermodynamic data with respect to ΔG° . For this displacement procedure, [16] which has been described in detail previously for inhibitors of closely related scaffolds, [12] ligand **C** has been used as the weak ligand to be displaced.

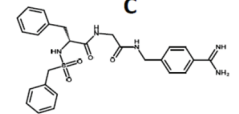
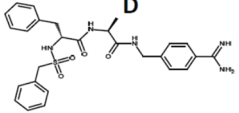
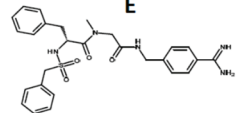
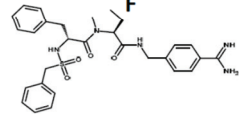
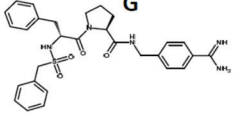
Compound	ITC Protocol/Buffer	ΔG^0	ΔH^0	$-T\Delta S^0$
	direct TRIS	-38.5 ± 0.3	-29.9 ± 0.3	-8.6 ± 0.4
	direct TRICINE	-39.2 ± 1.0	-22.3 ± 0.3	-16.9 ± 1.2
	direct HEPES	-38.3 ± 0.6	-17.1 ± 0.1	-21.2 ± 0.5
	ion. corrected	-38.3 ± 0.6	-7.0 ± 0.2	-31.3 ± 0.7
	direct TRIS	-43.3 ± 0.2	-41.0 ± 1.0	-3.3 ± 1.1
	direct TRICINE	-43.7 ± 0.1	-34.7 ± 0.5	-9.0 ± 0.4
	direct HEPES	-44.4 ± 0.7	-26.9 ± 0.8	-17.5 ± 1.1
	ion. corrected	-43.8 ± 0.3	-17.0 ± 0.8	-26.8 ± 0.9
	direct TRIS	-44.1 ± 1.0	-50.1 ± 1.7	6.0 ± 2.7
	direct TRICINE	-43.2 ± 0.3	-37.7 ± 1.8	-5.5 ± 2.1
	direct HEPES	-44.7 ± 1.3	-36.0 ± 1.8	-8.7 ± 3.1
	ion. corrected	-44.4 ± 0.8	-23.1 ± 1.8	-21.3 ± 2.6
	direct TRIS	-43.1 ± 0.6	-37.9 ± 1.1	-5.2 ± 1.5
	direct TRICINE	-43.2 ± 0.2	-30.8 ± 1.0	-12.4 ± 0.8
	direct HEPES	-42.8 ± 0.4	-24.1 ± 0.8	-18.7 ± 0.4
	ion. corrected	-43.1 ± 0.4	-13.9 ± 1.0	-29.2 ± 0.9
	direct TRIS	-54.8 ± 0.3	-37.6 ± 3.5	-17.2 ± 3.2
	direct TRICINE	-54.2 ± 3.3	-29.5 ± 2.3	-24.7 ± 1.0
	direct HEPES	-52.7 ± 2.9	-26.3 ± 0.2	-26.4 ± 2.7
	ion. corrected	-53.9 ± 2.2	-16.9 ± 2.0	-37.0 ± 2.3

Figure 4.5 Chemical structures of the ligands, thermodynamic parameters (ΔG^0 in blue, ΔH^0 in green, $-T\Delta S^0$ in red in kJ/mol) determined by direct ITC titrations in TRIS, TRICINE and HEPES buffer for ligands **C**, **D**, **E**, and **F**. Values for **G** have been determined by displacement titrations in the respective buffers using **C** as weak ligand. The estimated standard deviations were calculated from at least triplicate measurements.

4.4.4 MD simulation

In order to estimate the amount of degrees of freedom lost during the binding process by the different ligands exhibiting P2 portions of deviating flexibility, we carried out MD simulations to determine the rotational degrees of freedom of the crucial torsion angles in water. Differences in the thermodynamic properties can be used to assess some part of the entropic contribution to binding. For the five differently substituted ligands it can be expected that the flexibility of the central P2 portion takes significant influence on the accessible conformational space of the ligands prior to protein binding. We therefore analyzed the phi, psi, and omega dihedral angles of the P2 part for a simulation in water. Each dihedral angle has been defined in AMBER 12[17] using the four adjacent atoms and measured in degrees. They have been evaluated explicitly and the results are summarized by considering every frame along the trajectory in terms of a histogram (**Figure 4.6**).

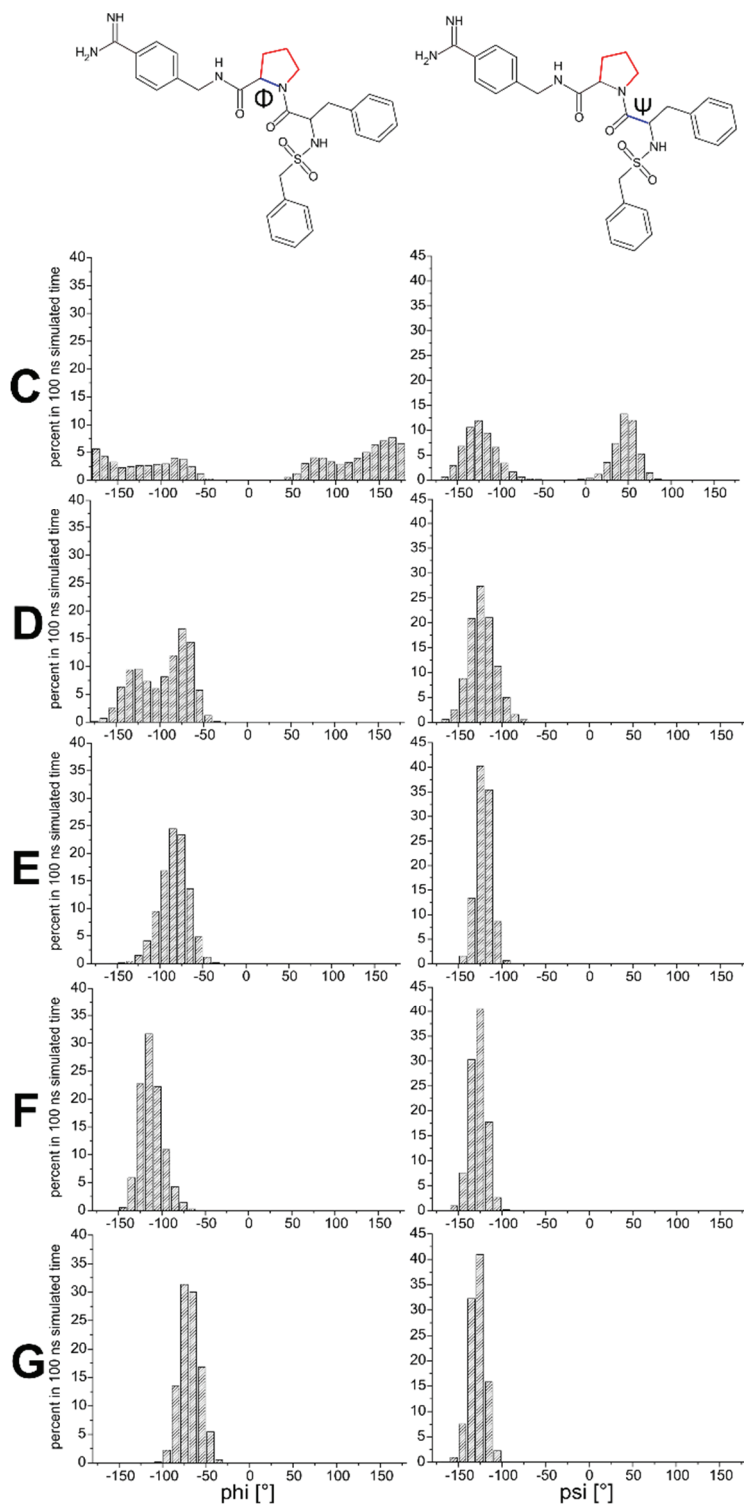


Figure 4.6 The calculated conformational distributions for the dihedral angles ϕ and ψ of the ligands C-G are shown in [°]. The respective bonds for the angle calculations are indicated in the chemical structure in blue.

The phi angle records the geometry between the two differently substituted P2 positions in the five derivatives. We therefore expect for this angle the largest differences. The least restricted glycine derivative **C** shows a broad distribution of accessible conformations and populates nearly the entire angular range, apart from values around 0°. This conformational gap most likely results from sterically unfavorable orientations of the adjacent carbonyl groups with respect to each other. In the other derivatives only angles with negative values are populated and any value in the mirrored positive torsion angle space remains unoccupied. As expected, the proline derivative **G** is conformationally most restricted and only populates at the preferred angular range values around -75°. Ligand **F** adopts a preferred angular phi range next to -120°, where the three attached substituents (ethyl, N-methyl and carbonyl) experience maximal distances between each other. Ligands **D** and **E** also prefer this orientation, for **D** the phi angle is scattered across a larger range between -50° to -150°.

The psi angle describes the positioning of the carbonyl group relative to the two P3/P4 substituents. Also with respect to this angle, the glycine derivative **C** shows the largest flexibility as a bimodal distribution is indicated. Two families of conformations are populated, suggesting that the two substituents can pass each other upon single bond rotation, whereas for all other ligands only one conformational range can be adopted.

We also recorded the distribution of the omega angle, however, along the simulated trajectory we did not observe for any of the five ligands a flip of the peptide bond between the *cis* and *trans* orientation. Actually, the angular variations (maximally $\pm 5^\circ$) are very similar across the compound series.

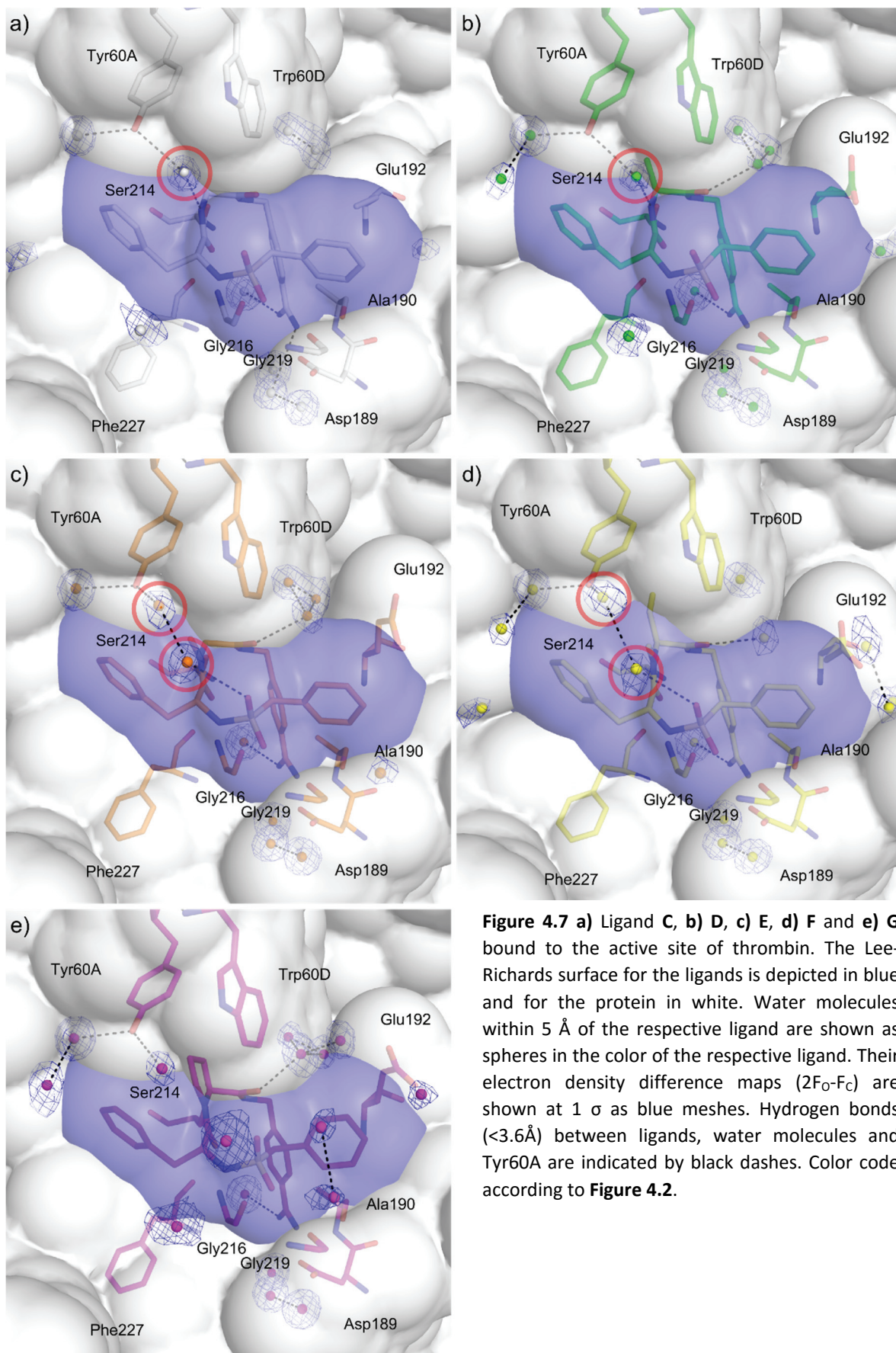
4.5 Discussion

At first glance, the chemical modifications of the P2 moiety appear rather small across this congeneric series of ligands, however, the affinity differences and the relative partitioning in enthalpy and entropy are huge. First of all, the preorganization and thus rigidification of the ligands with respect to the bound conformation at the binding site alters strongly across the series. The MD simulations performed in aqueous solution indicate large differences of the inhibitors in terms of their accessible conformational space as expressed by the observed scatter of the torsion angles in the P2 portion (**Figure 4.6**). The glycine derivative **C** is clearly more flexible than the other substituted derivatives, the proline derivative **G** is the conformationally most strongly constrained one. This difference takes impact on the entropic contribution to binding as **C** will sacrifice a larger proportion of its degrees of freedom compared to **G**. Interestingly, the residual mobility at the binding site shows an opposite trend as indicated by the B-factor analysis. The glycine and alanine derivatives exhibit the largest residual mobility, whereas the proline ligand is the most firmly fixed one of the series. This behavior should be reflected in controversial trends in the entropic signatures across the series as the loss of degrees of freedom is entropically unfavorable whereas the preservation of these degrees will provide an entropic advantage. Finally, the complex of the glycine derivative shows remarkable singularity in the series as only in this complex the side chain of Glu192 is completely disordered whereas in all other complexes this residue adopts a more confined orientation. This surprising observation was confirmed by collecting a second data set using another crystal, nevertheless, leading to the same results. The disorder phenomenon will clearly provide the complex of **C** with an entropic advantage compared to the other members of the series.

The different ligands bury an increasing amount of hydrophobic surface beneath the 60-loop of thrombin which forms a small hydrophobic cleft flanked by Tyr60A and Trp60D. Important enough, this cleft accommodates in the *apo*-enzyme a well-defined water molecule flanked by the two aromatic residues. This ordered water molecule must be released upon ligand binding and will produce, most likely, an entropic signal. Whereas the glycine derivative **C** leaves this niche largely unoccupied the N-Me and C α substituted

ligands fill this pocket increasingly efficient. The N-Me *homo*-alanine derivative **F** achieves the best filling closely followed by **G** and **D**, and with some difference to **E**. Interestingly, at the far end of this pocket the catalytic residue His57 is located which is involved in a proton release upon ligand binding. Possibly, the water release and the filling of the adjacent pocket contributes to the induced pKa shift of this residue resulting in the release of protons (**Figure 4.4**).

Finally, the hydrogen bond inventory is changed across the series. The Gly and Ala derivatives form a water-mediated hydrogen bond with Tyr60A OH. (**Figure 4.7 a, b**)



The N-alkylated ligands **E**, **F** and **G** abandon this hydrogen bond but the interstitial water molecule is not repelled from the complexes. Instead, it is spatially shifted to still interact with Tyr60AOH. At the new position, the water molecule finds a favorable van der Waals contact with the attached N-Me group. Furthermore, it is able to recruit an additional water molecule to bridge the gap to the sulfonyl group of the ligand via a chain of H-bond connected water molecules (**Figure 4.7c**). This kind of belt spanning across the ligand is well established with **E** and **F**. (**Figure 4.7 c, d**). In case of **G** the recruited water network covering the bound ligand takes an even more complex pattern, however, most likely this network helps to fix **G** with the lowest residual mobility in the series. (**Figure 4.7 e**).

Remarkably, the affinity of **C** increases by -15.6 kJ/mol upon introduction of the central P2 proline moiety, even though the water-mediated H-bond of **C** is lost compared to **G**. The affinity gain of **G** compared to the other three open chain derivatives **D**, **E**, and **F** still amounts to about -10 kJ/mol. One aspect of the huge Gibbs free energy gain of **G** is the correct preorganization resulting from the rigid proline moiety. An entropic advantage seen for this ligand compared to the other members of the series underlines the benefit of the preorganization and that this ligand loses less conformational degrees of freedom. Therefore, its binding will hardly counteract any other entropically favored steps which parallel ligand binding. However, the entropy inventory across the series is more complex due to several superimposed and partly controversial effects (see above, e.g., the residual mobility, entrapped water surface layer and conformational properties of Glu192).

A direct comparison of **C** and **D** results in a $\Delta\Delta G_{C \rightarrow D}$ of -5.5 kJ/mol. Both ligands form the water-mediated H-bond to Tyr60A. However, **D** fills the hydrophobic cleft below the 60-loop with a methyl group which is lacking in **C**. The $\Delta\Delta G$ increment attributed to this methyl group is quite large, however the group becomes well buried in the hydrophobic niche which might explain the pronounced affinity contribution of a single methyl group.[18] In addition, the C $_{\alpha}$ substituted alanine derivative **D** can only access a reduced conformational space compared to the glycine analog owing to this additional methyl group. This difference matters prior to protein binding in aqueous solution.

The methylation of **C** to feature **E** results in a similar affinity enhancement ($\Delta\Delta G_{C \rightarrow E} = -6.1$ kJ/mol), however, the free energy gain factorizes differently. Most likely, the methyl contribution is smaller as the burial beneath the 60-loop is less efficient (**Figure 4.3b**). Obviously, the lost H-bond of **C** is not penalized, possibly it is enthalpically even overcompensated in the complex with **E** by forming a new water-mediated contact to the SO₂ group of the ligand (**Figure 4.7c**). This requires the entrapment of an additional water molecule that might explain part of the entropic disadvantage of **E** compared to **C**. In addition, **E** should be energetically cheaper to desolvate which may further explain its more enthalpic binding profile. Moreover, the N-methylation should restrict to larger extent the accessible conformational space of **E** compared to **C** prior to binding, disfavoring **C** entropically relative to **E**.

Overall the alanine **D** and the two N-Me derivatives **E** and **F** bind with rather similar affinities. However, that this is owing to mutually compensating effects is underscored by the factorization of enthalpy/entropy, which varies quite considerably across these three ligands. As mentioned above, multiple features are in operation to determine the actual thermodynamic signature of these three ligands, thus a detailed explanation of the inventory appears difficult to establish. Finally, the preorganized proline derivative **G** improves affinity by nearly two orders of magnitude compared to **D**, **E**, and **F** and almost three orders with respect to **C**. The adopted binding mode and interaction pattern with Tyr60A resembles closely those found in the complexes of **F** and **G**. Furthermore, the desolvation of both ligands, formally showing the same number of carbon atoms in the P2 moiety should be very similar. Nonetheless, the binding profile of **G** is still predominantly improved by an entropic advantage that we ascribe to the conformational preorganization of the proline derivative in consequence of the introduced cyclic ring structure in the center of the inhibitor.

The strong entropically favored signature of **C** appears surprising on first glance, considering its conformational flexibility experienced in the unbound state compared to the other series members, as this ligand has to sacrifice a larger amount of its conformational degrees of freedom. Obviously, other entropically beneficial contributions outweigh this handicap. First of all, the unique behavior of the Glu192 side chain, scattering

over multiple orientations, may produce an entropic advantage. Second, the residual mobility of this ligand is quite large and also enhances the entropic inventory, particularly in comparison with **G**.

4.6 Conclusions

Several messages can be learnt from the current study. Binding affinity is not solely determined by the properties of the formed protein-ligand complex but some differences can already discriminate ligands in aqueous solution prior to any protein binding. Conformationally restricted ligands such as **G** can experience a significant binding advantage over more flexible competitors if the correct protein-bound geometry is preorganized in a rigidified skeleton. The two ligands **F** and **G** possess apart from two aliphatic hydrogens the same chemical formulae, thus a very similar desolvation enthalpy will be required to shed their solvation shell. Nonetheless, they differ in binding affinity by more than 10 kJ/mol; nearly two orders of magnitude in the binding constant. The advantage of the conformational restriction of **G** is expressed by an entropic benefit of nearly 8 kJ/mol.

It is also remarkable to see that methylation of a polar amide nitrogen does not necessarily result in a drop of affinity, even though the free NH group of the corresponding ligand has been involved in an H-bond to the protein. An interstitial water molecule used to mediate this hydrogen bond to Tyr60AOH is slightly shifted in the N-Me derivative to a new position where it can establish a kind of interaction belt wrapping around and in consequence fixing the ligand in its binding pose.

Obviously, the efficiency how the newly formed solvation layer arranges across the protein complex with its bound ligand also takes an impact on binding. In the present case the coverage of the ligand by a well-established water layer obviously reduces the residual mobility of the bound ligand as indicated by the B-factor analysis. Several studies have already underscored the importance of the latter feature to the entropic term of the binding signature.[19] [20] Overall, the current narrow ligand series suggests that even

small changes involving the addition or reshuffling of methyl groups at a given parent scaffold can take remarkable influence on the affinity and even more on the enthalpy/entropy partitioning in ligand binding.

4.7 Experimental Section / Materials & Methods

4.7.1 *Synthesis*

The synthesis of the current compound series has been described elsewhere [21] and follows a very similar protocol as described previously. [12]

4.7.2 *ITC*

ITC experiments were conducted using an ITC200™ system from GE Healthcare, Northampton, MA, USA. Thrombin was obtained from CSL Behring (Marburg, Germany) and purified from Beriplast®. Thrombin was extracted by dialysis using an experimental buffer of 50 mM TRIS, TRICINE or HEPES, 100 mM NaCl, 0.1 % polyethylene glycol 8000 at pH 7.8. Subsequently, the protein could be used for all titration experiments. The protein concentration was measured by absorbance at 280 nm using a NanoDrop 2000c spectrophotometer from Thermo Scientific.

Ligands from 50 mM DMSO stock solutions were diluted to 0.25-0.50 mM in ITC buffer and subsequently the DMSO concentration in all solutions was adjusted to 3 %. All titrations were performed at least in triplicate. For the strong binding ligand **G**, a displacement approach has been applied, where the protein has been saturated with the weaker binding ligand **C** in advance.

All ITC experiments were started at 25 °C with a reference power of 5 kcal/s after a stable baseline had been achieved. The pre-titration delay was set to 300s. For titrations ligand injections of 0.3 µL (to prevent artefacts arising from small syringe leakages or air in the

syringe) were followed by 19 to 27 injections of 1.2 – 1.8 μ L with at least 150s interval between each injection. Data integration and evaluation was done with Origin 7.

In a previous study, we reported on the thermodynamic data of **C** in TSPP buffer.[10] We first included these data in our buffer correction using the literature-reported value of 5 kJ/mol for the heat-of-ionization.[11] Unexpected deviations from linearity, described by the three other buffers, were observed. In consequence, we measured all remaining four ligands in TSPP buffer but now we experienced in all cases similar deteriorations from linearity caused by the TSPP data. A possible explanation might be an incorrectly reported heat-of-ionization for TSPP. Finally, we decided to remove the TSPP buffer data from our analysis and considered only the TRIS, TRICINE and HEPES data.

4.7.3 X-ray Data

4.7.3.1 Crystallization and soaking

Human α -thrombin (from Enzyme Research Laboratories, South Bend, USA) was dissolved in the crystallization buffer (20 mM NaH₂PO₄, 350 mM NaCl, 2 mM benzamidine, pH 7.5) at 10 mg/ml. The hirudin fragment Hirudin (54-65) (sulfated), purchased from Bachem (Bubendorf, Switzerland), was dissolved in crystallization buffer at 2.5 mg/mL. In the next step, 40 μ L of the solution of the hirudin fragment was mixed with 160 μ L of the thrombin solution. After incubation for 1 h at 4 °C, crystallization was carried out at 4 °C by the hanging-drop method. The hirudin/thrombin solution was mixed 1:1 with the reservoir solution (20 mM NaH₂PO₄, 27 % polyethylene glycol 8000, pH 7.5) and 2 μ L of this solution were placed in the centre of a cover slip. Immediately after mixing of protein and reservoir buffer microseeding was performed by streaking with a horse hair. The wells of the crystallization trays were filled with 500 μ L of the reservoir buffer. Subsequently, the cover slips were placed over the wells and sealed. Crystals of good diffracting quality could be obtained after 10 to 14 days.

For soaking, a 50 mM DMSO stock solution of the inhibitors was diluted 1:10 with a solution containing crystallization and reservoir buffer 1:1 resulting in the final soaking concentration containing 5 mM of the respective compound and 10 % DMSO. A crystal

without visible imperfections was selected and transferred into the soaking solution for 24h.

4.7.3.2 Data collection and processing

Crystals were prepared for data collection at 110 K using a cryoprotectant solution of 20 % glycerol in reservoir buffer for two minutes.

As mentioned previously the collection of the data set of **C** has been described in a previous contribution.[10] We collected a second data set of this complex at BESSY (Berlin), however, the same binding geometry with the enhanced disorder of the Glu192 sidechain has been recorded. The data of this structure have not been deposited.

The data sets of **D**, **E** and **F** were collected at the synchrotron BESSY (Berlin) at beamline 14.2 (wavelength of 0.9184Å) and beamline 14.3 (wavelength of 0.895Å) using Rayonix MX-225 detectors. The data set of **G** has been collected at the synchrotron SLS (Villingen) at beamline X06DA (wavelength of 1.00 Å) using a MAR 225 detector.

For compounds **E** and **F** data processing and scaling was performed with the XDS program package.[22] For compounds **D** and **G** data processing and scaling was performed using the HKL2000 package.[23]

4.7.3.3 Structure determination and refinement

The coordinates of human thrombin (PDB codes 1H8D [24] and 4UE7 [10]) were used for molecular replacement with Phaser from the CCP4 program package.[25]

For initial rigid body refinement of the protein molecule, followed by repeated cycles of maximum likelihood energy minimization simulated annealing and B-factor refinement the program PHENIX was used.[26] For determining the temperature factors for the structures of **D**, **E** and **F** TLS refinement was applied. The definition of the TLS groups was done with PHENIX. For compound **G**, anisotropic B factor refinement was applied. A randomly chosen 5 % of all data were used for the calculation of R_{free} and were not considered in the refinement. Amino acid side chains were fit into σ -weighted $2F_o - F_c$ and $F_o - F_c$ electron

density maps using Coot.[27] After the first refinement cycle, water molecules and subsequently ions and ligands were located in the electron density and added to the model. Restraints were applied to bond lengths and angles, planarity of aromatic rings and van der Waals contacts. Multiple side chain conformations were built in case an appropriate electron density was observed and maintained during the refinement, and if the minor populated side chain showed at least 20% occupancy. The final model was validated using PHENIX own validation options and Coot. The Ramachandran plot was calculated with MolProbity[28] integrated in the PHENIX package. Data collection, unit cell parameters and refinement statistics are given in **Table 4.1**. Analysis of temperature factors was performed with MolProbity. The naming of the protein amino acids was done according to Bode et al.[29]

<i>Compound (PDB-Code)</i>	<i>C* (5AFZ)</i>	<i>D (4UFE)</i>	<i>E (4UFF)</i>	<i>F (4UFG)</i>	<i>G (4UFD)</i>
<i>Wavelength (Å)</i>	0.895	0.9184	0.895	0.9184	1.00
<i>Resolution range^a (Å)</i>	21.47 - 1.53 (1.58 - 1.53)	29.00 - 1.59 (1.65 - 1.59)	35.62 - 1.55 (1.61 - 1.55)	35.68 - 1.65 (1.709 - 1.65)	49.66 - 1.43 (1.48 - 1.43)
<i>Space group</i>	C 1 2 1	C 1 2 1	C 1 2 1	C 1 2 1	C 1 2 1
<i>Unit cell α, b, c (Å)</i>	70.3 71.6 72.4	70.0 71.4 72.5	70.4 71.6 72.5	70.1 71.3 72.6	70.0 71.6 71.9
<i>α, β, γ (°)</i>	90.0 100.5 90.0	90.0 100.5 90.0	90.0 100.5 90.0	90.0 100.5 90.0	90.0 100.0 90.0
<i>Total reflections</i>	201127 (19424)	138973 (10986)	134837 (13259)	117138 (11449)	230688 (19646)
<i>Unique reflections</i>	53119 (5257)	42965 (3788)	50817 (5040)	41835 (4089)	62348 (5613)
<i>Multiplicity</i>	3.8 (3.7)	3.2 (2.9)	2.7 (2.6)	2.8 (2.8)	3.7 (3.5)
<i>Completeness (%)</i>	99.8 (99.4)	91.5 (80.9)	99.00 (99.0)	98.9 (97.2)	96.5 (88.1)
<i>Mean I/sigma(I)</i>	17.7 (2.8)	14.9 (2.6)	12.7 (2.3)	18.8 (3.7)	30.1 (4.1)
<i>Wilson B-factor</i>	17.0	14.5	15.3	14.0	12.5
<i>R-merge</i>	0.048 (0.527)	0.038(0.509)	0.055 (0.549)	0.051 (0.497)	0.033 (0.320)
<i>R-work</i>	0.145 (0.176)	0.158 (0.214)	0.160 (0.228)	0.162 (0.189)	0.134 (0.147)
<i>R-free</i>	0.174 (0.191)	0.182 (0.258)	0.187 (0.247)	0.187 (0.222)	0.159 (0.176)
<i>Number of non-hydrogen atoms</i>	2645	2730	2658	2672	2774
<i>macromolecules</i>	2383	2396	2390	2392	2409
<i>ligands</i>	43	44	44	46	46
<i>water</i>	218	289	222	234	305
<i>Protein residues</i>	287	290	288	290	289
<i>RMS(bonds) (Å)</i>	0.006	0.007	0.006	0.006	0.009
<i>RMS(angles) (°)</i>	1.04	1.08	1.08	1.09	1.13
<i>Ramachandran favored (%)^b</i>	98	98	98	98	97
<i>Ramachandran outliers (%)</i>	0	0	0	0	0
<i>Ramachandran allowed (%)</i>	2	2	2	2	3
<i>Clashscore</i>	1.06	1.88	1.26	1.69	2.26
<i>Average B-factor (Å³)</i>	25.0	21.8	22.9	19.9	19.9
<i>macromolecules</i>	24.3	20.7	22.2	19.2	18.3
<i>ligands</i>	26.3	23.5	23.1	18.5	20.8
<i>solvent</i>	32.0	30.5	30.1	26.6	31.5
<i>B ligand / B protein</i>	1.08	1.16	1.00	0.97	0.86
<i>Coordinate errors^c (Å)</i>	0.12	0.14	0.14	0.15	0.10

* Published in a previous contribution [10], data are shown only for comparison purposes.

^a Statistics for the highest-resolution shell are shown in parentheses.

^b Calculated by MOLPROBITY

^c Maximum-likelihood based

Table 4.1 Data collection and refinement statistics.

4.7.4 MD Simulation

The molecular dynamics (MD) simulations of the ligands **C**, **D**, **E**, **F** and **G** have been performed with the use of Amber12 [17] and the DPDP precision model.[30] The trajectory was analyzed with AmberTools12 [31] using *ptraj* and *cpptraj*. [32] The ligands have been prepared using MOE (v. 2012.10; Chemical Computing Group, Montreal, Canada). 100 ns MD simulations were carried out for each ligand. The histograms were prepared with Origin 7 (OriginLab).

The ligands were protonated with MOE using the Protonate3D-tool with the default values. After the preparation in MOE, the ligands were parameterized with the programs *antechamber* and LEaP of the AmberTools12-package. The charge and partial charges of the ligands were defined with *antechamber*. AM1-BCC was used as charge method; the net charge was specified as +1. The generalized Amber force field (GAFF) was used.[17] A chloride was added as a counterion with *antechamber*. Then, a truncated octahedral periodic box of TIP3P water molecules as a solvent was allocated around the ligand.[30] The closest distance between the solvent atoms to the solute atoms (i.e. to the ligands) was set to 10 Å. The calculations of the MD simulations were performed with *pmemd.cuda* on Nvidia K20 graphic cards with the particle mesh Ewald method [33] under periodic boundary conditions with a cut-off of 10 Å. The SHAKE algorithm [34] was used to constrain bonds involving hydrogen. The temperature and the isotropic position scaling was set using Langevin dynamics [35] with a predefined collision frequency of 5 ps⁻¹ between external heat reservoir and simulated system. The time steps for the calculations of the atomic coordinates were set to 2 fs.

The first minimization procedure consisted of 400 cycles of steepest descent algorithm with subsequent 100 cycles of conjugate gradient algorithm. The second minimization procedure comprised 1000 cycles of steepest descent- and 200 cycles of conjugate gradient strategy. The system was heated in three steps from 0 to 300 K in 150 ps with use of Cartesian restraints of 500 kcal mol⁻¹ Å⁻² and subsequent adjustment of the pressure to 1

bar during 100 ps. In the last preliminary simulations, the Cartesian restraints were decreased to zero during 300 ps. The coordinates of the trajectory were saved every 10 ps.

The dihedral angles between the two planes of the atoms a, b, c and b, c, d have been calculated for every frame with *ptraj*.

4.7.5 Volume Calculation

In order to determine the differences of the respective unoccupied volume below the 60 loop across the five complexes, we used the program *POVME*. [36] All structures were aligned by fitting the C α atoms to each other. Protonation states were consistently assigned using MOE. A narrow-spaced grid of 0.15Å, spanning the region between the ligand, His57, and the 60 loop was computed. A probe sphere of 0.54 Å radius (corresponding to half of the van der Waals radius of a hydrogen atom) was used to assess whether a grid point was still accessible or coincided with an atom. The volume of the unoccupied site was calculated based on the remaining freely accessible and contiguously connected grid points.

4.8 Biochemical Assays

Details to the biochemical photometric and fluorogenic assays can be found in previous publications.[12,37]

4.9 Accession codes

The atomic coordinates and structure factors of thrombin in complex with ligands **C** to **G** were deposited in the RCSB Protein Data Bank (PDB) with accession codes 5AFZ (**C**), 4UFE (**D**), 4UFF (**E**), 4UFG (**F**), 5UFD (**G**).

4.10 Acknowledgement

We kindly acknowledge CSL Behring, Marburg, Germany, for supplying us with generous amounts of human thrombin from the production of Beriplast. We are grateful to the

beamline staff at BESSY II (Helmholtz-Zentrum Berlin) in Berlin and SLS in Villingen, Switzerland, for providing us outstanding support during the data collection. We also thank the Helmholtz-Zentrum Berlin for the travel support. We are grateful to the Chemical Computing Group for making the MOE program available to us. The usage of the AMBER program suite is acknowledged. The present study was kindly supported by the ERC Advanced Grant no. 268145-DrugProfilBind awarded to GK.

4.11 References

- [1] J. Clardy, C. Walsh, Lessons from natural molecules, *Nature* 432 (2004) 829–837.
- [2] J. Boström, P.-O. Norrby, T. Liljefors, Conformational energy penalties of protein-bound ligands, *Journal of Computer-Aided Molecular Design* 12 (1998) 383–383.
- [3] K.A. Brameld, B. Kuhn, D.C. Reuter, M. Stahl, Small Molecule Conformational Preferences Derived from Crystal Structure Data. A Medicinal Chemistry Focused Analysis, *J. Chem. Inf. Model.* 48 (2008) 1–24.
- [4] A. Reichelt, S.F. Martin, Synthesis and properties of cyclopropane-derived peptidomimetics, *Acc. Chem. Res.* 39 (2006) 433–442.
- [5] S.F. Martin, Preorganization in biological systems: Are conformational constraints worth the energy? *Pure and Applied Chemistry* 79 (2007) 193–200.
- [6] J.E. DeLorbe, J.H. Clements, M.G. Teresk, A.P. Benfield, H.R. Plake, L.E. Millspaugh, S.F. Martin, Thermodynamic and Structural Effects of Conformational Constraints in Protein–Ligand Interactions. Entropic Paradoxy Associated with Ligand Preorganization, *J. Am. Chem. Soc.* 131 (2009) 16758–16770.
- [7] S.F. Martin, J.H. Clements, Correlating structure and energetics in protein-ligand interactions: paradigms and paradoxes, *Annu. Rev. Biochem.* 82 (2013) 267–293.
- [8] H.S. Gurm, D.L. Bhatt, Thrombin, an ideal target for pharmacological inhibition: a review of direct thrombin inhibitors, *Am. Heart J.* 149 (2005) S43–53.
- [9] B. Baum, L. Muley, A. Heine, M. Smolinski, D. Hangauer, G. Klebe, Think twice: understanding the high potency of bis(phenyl)methane inhibitors of thrombin, *J. Mol. Biol.* 391 (2009) 552–564.
- [10] E. Rühmann, M. Betz, A. Heine, G. Klebe, Fragment Binding Can Be Either More Enthalpy-Driven or Entropy-Driven: Crystal Structures and Residual Hydration Patterns Suggest Why, *Journal of medicinal chemistry* 58 (2015) 6960–6971.
- [11] H. Fukada, K. Takahashi, Enthalpy and heat capacity changes for the proton dissociation of various buffer components in 0.1 M potassium chloride, *Proteins* 33 (1998) 159–166.
- [12] A. Biela, F. Sielaff, F. Terwesten, A. Heine, T. Steinmetzer, G. Klebe, Ligand binding stepwise disrupts water network in thrombin: enthalpic and entropic changes reveal classical hydrophobic effect, *J. Med. Chem.* 55 (2012) 6094–6110.
- [13] B. Baum, L. Muley, M. Smolinski, A. Heine, D. Hangauer, G. Klebe, Non-additivity of Functional Group Contributions in Protein–Ligand Binding: A Comprehensive Study by

Crystallography and Isothermal Titration Calorimetry, *Journal of Molecular Biology* 397 (2010) 1042–1054.

[14] Y.L. Zhang, Z.Y. Zhang, Low-affinity binding determined by titration calorimetry using a high-affinity coupling ligand: a thermodynamic study of ligand binding to protein tyrosine phosphatase 1B, *Anal. Biochem.* 261 (1998) 139–148.

[15] B.W. Sigurskjold, Exact analysis of competition ligand binding by displacement isothermal titration calorimetry, *Anal. Biochem.* 277 (2000) 260–266.

[16] A. Velazquez-Campoy, E. Freire, Isothermal titration calorimetry to determine association constants for high-affinity ligands, *Nat Protoc* 1 (2006) 186–191.

[17] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general amber force field, *J. Comput. Chem.* 25 (2004) 1157–1174.

[18] C.S. Leung, S.S.F. Leung, J. Tirado-Rives, W.L. Jorgensen, Methyl Effects on Protein–Ligand Binding, *J. Med. Chem.* 55 (2012) 4489–4500.

[19] A. Biela, M. Betz, A. Heine, G. Klebe, Water makes the difference: rearrangement of water solvation layer triggers non-additivity of functional group contributions in protein–ligand binding, *ChemMedChem* 7 (2012) 1423–1434.

[20] S.G. Krimmer, M. Betz, A. Heine, G. Klebe, Methyl, ethyl, propyl, butyl: futile but not for water, as the correlation of structure and thermodynamic signature shows in a congeneric series of thermolysin inhibitors, *ChemMedChem* 9 (2014) 833–846.

[21] M.T. Khayat, Understanding Molecular Recognition: Thermodynamics and Binding Kinetics of Potent Thrombin Inhibitors, (Dissertation) 01.01.2013.

[22] W. Kabsch, XDS, *Acta Crystallogr. D Biol. Crystallogr.* 66 (2010) 125–132.

[23] Z. Otwinowski, W. Minor, [20] Processing of X-ray diffraction data collected in oscillation mode, in: *Methods in Enzymology*, vol. 276, *Macromolecular Crystallography Part A*, Elsevier, 1997, pp. 307–326.

[24] H.U. Ahmed, M.P. Blakeley, M. Cianci, D.W.J. Cruickshank, J.A. Hubbard, J.R. Helliwell, The determination of protonation states in proteins, *Acta Crystallogr. D Biol. Crystallogr.* 63 (2007) 906–922.

[25] The CCP4 suite: programs for protein crystallography, *Acta Crystallogr. D Biol. Crystallogr.* 50 (1994) 760–763.

[26] P.D. Adams, P.V. Afonine, G. Bunkóczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.-W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, P.H. Zwart, PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr. D Biol. Crystallogr.* 66 (2010) 213–221.

- [27] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, *Acta Crystallogr. D Biol. Crystallogr.* 60 (2004) 2126–2132.
- [28] V.B. Chen, W.B. Arendall, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, D.C. Richardson, MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallogr. D Biol. Crystallogr.* 66 (2010) 12–21.
- [29] W. Bode, I. Mayr, U. Baumann, R. Huber, S.R. Stone, J. Hofsteenge, The refined 1.9 Å crystal structure of human alpha-thrombin: interaction with D-Phe-Pro-Arg chloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion segment, *EMBO J.* 8 (1989) 3467–3475.
- [30] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.* 79 (1983) 926–926.
- [31] R. Salomon-Ferrer, D.A. Case, R.C. Walker, An overview of the Amber biomolecular simulation package, *WIREs Comput Mol Sci* 3 (2013) 198–210.
- [32] D.R. Roe, T.E. Cheatham, PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data, *J. Chem. Theory Comput.* 9 (2013) 3084–3095.
- [33] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems, *J. Chem. Phys.* 98 (1993) 10089–10089.
- [34] J.-P. Ryckaert, G. Ciccotti, H.J. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *Journal of Computational Physics* 23 (1977) 327–341.
- [35] R.J. Loncharich, B.R. Brooks, R.W. Pastor, Langevin dynamics of peptides: the frictional dependence of isomerization rates of N-acetylalanyl-N'-methylamide, *Biopolymers* 32 (1992) 523–535.
- [36] J.D. Durrant, C.A.F. de Oliveira, J.A. McCammon, POVME: an algorithm for measuring binding-pocket volumes, *J. Mol. Graph. Model.* 29 (2011) 773–776.
- [37] E. Rühmann, M. Betz, M. Fricke, A. Heine, M. Schäfer, G. Klebe, Thermodynamic signatures of fragment binding: Validation of direct versus displacement ITC titrations, *Biochim. Biophys. Acta* 1850 (2015) 647–656.

5 Kristallographische und thermodynamische Untersuchungen von MI067 im Vergleich zu MI001 und MI002

5.1 Einleitung

In diesem Kapitel wird ein Thrombin Inhibitor (**MI067**) in Kooperation mit der Arbeitsgruppe von Prof. Dr. Steinmetzer thermodynamisch und kristallographisch charakterisiert und mit zwei verwandten Thrombin-Inhibitoren verglichen, die bereits durch Dr. Adam Biela [1] (**MI001**) und Dr. Bernhard Baum [2] (**MI002**) untersucht worden sind.

Die Synthese und Messung der Enzymassays der Inhibitoren erfolgte in der Arbeitsgruppe von Prof. Dr. Steinmetzer.

In der P4 Position zeigen alle Inhibitoren ein Benzylsulfonamid und in der P2 Position einen für Thrombin Inhibitoren sehr oft verwendeten Prolinrest. Für alle Inhibitoren wurde in P3 jedoch keine typische hydrophobe Aminosäure, sondern das polare *D*-Arginin verwendet. Um die S1 Tasche zu adressieren wurden für die bereits untersuchten Liganden (**MI002** und **MI001**) ein 4-AMBA- (4-Amidinobenzylamid) und ein ACB-Anker (2-Aminomethyl-5-chlorobenzylamid) verwendet und für die neu untersuchte Struktur (**MI067**) eine HCB Gruppe (2-Hydroxymethyl-5-chlorobenzylamid) (**Abbildung 5.1**).

Von diesen verschiedenen P1 Gruppen wird erwartet, dass sie unterschiedliche Effekte auf den Bindungsmodus und die thermodynamischen Eigenschaften der Liganden zeigen, da sie unterschiedliche intramolekulare Interaktionen ausbilden können.

Besonders interessant sind hierbei die thermodynamischen Unterschiede der ACB- und HCB-Gruppen, welche unterschiedliche Wasserstoffbrückendonator- und Wasserstoffbrückenakzeptoreigenschaften besitzen.

Für eine weitere Serie wurden von der Arbeitsgruppe Steinmetzer nur Affinitätsdaten bestimmt. Bei diesen Liganden handelt es sich um **MI073**, **MI074** und **MI066**, welche in P1 und P2 identisch zu ersten Serie aufgebaut sind, jedoch in P3 ein *D*-Cyclohexylalanin aufweisen und in P4 unbesetzt verbleiben. Die unbesetzte P4 Position schließt daher für alle Inhibitoren eine intramolekulare Interaktion mit dieser Position aus.

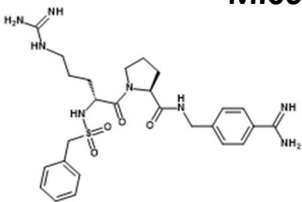
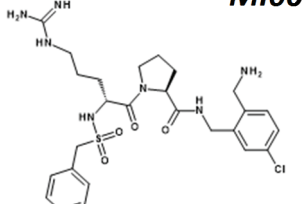
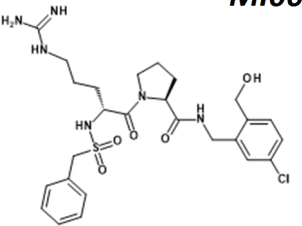
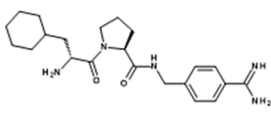
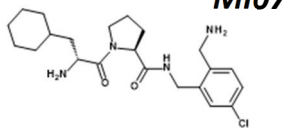
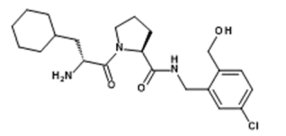
Compound ID	Sequence	K_i nM
<p>MI001</p> 	Bzls-D-Arg-Pro-4-Amba	2.22 ± 2.40
<p>MI002</p> 	Bzls-D-Arg-Pro-ACB	0.88 ± 0.55
<p>MI067</p> 	Bzls-D-Arg-Pro-HCB	2541.50 ± 846.47
<p>MI073</p> 	H-D-Cha-Pro-4-Amba	0.28 ± 0.05
<p>MI074</p> 	H-D-Cha-Pro-ACB	0.19 ± 0.11
<p>MI066</p> 	H-D-Cha-Pro-HCB	24.14 ± 2.19

Abbildung 5.1 Gezeigt sind die Strukturformeln, Sequenzen und Affinitätskonstanten der Inhibitoren **MI001**, **MI002** und **MI067** sowie von **MI073**, **MI074** und **MI066**.

5.2 Kristallstrukturen

Die Komplexstruktur von **MI067** konnte mit einer Auflösung von 1.40 Å bestimmt werden (**Abbildung 5.2**).

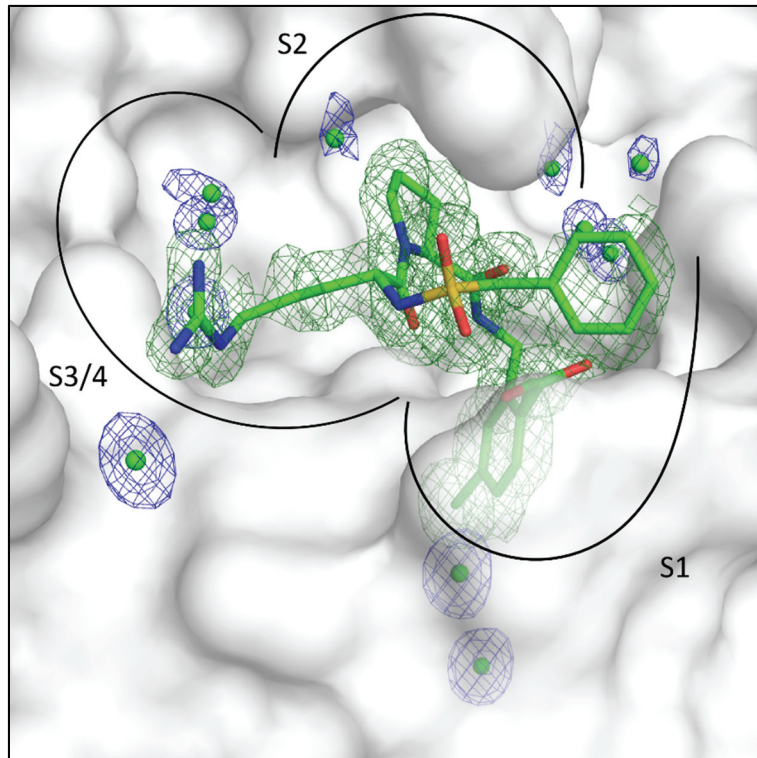


Abbildung 5.2: Inhibitor **MI067** (grün) im Komplex mit Thrombin. Die F_o-F_c Differenzelektronendichte des Liganden ist bei 2σ in grün gezeigt. Wassermoleküle im Radius von 5\AA um den Liganden sind als grüne Kugeln und die dazugehörige $2F_o-F_c$ Elektronendichte bei 1σ in blau dargestellt. Stickstoffatome bei 1σ in blau, Sauerstoffatome in rot, Schwefelatome in gelb und Kohlenstoffatome in der Farbe des Inhibitors gezeigt. Die spezifischen Bindetaschen S1, S2 und S3/4 sind ebenfalls hervorgehoben.

MI067 zeigt die für m-Chloraromaten in P1 typische Chlor- π Interaktion zu Tyr228. (**Abbildung 5.3**) Der 2-Hydroxymethylsubstituent zeigt sich in zwei Konformationen und ist in der Lage, Wasserstoffbrücken zum Amidstickstoff von Glu192 (3.0\AA , Besetzung 44%) und zum Carbonylsauerstoff von Gly219 (2.8\AA , Besetzung 56%) auszubilden. Darüber hinaus ist eine schwache intramolekulare Wasserstoffbrücke der zweiten Konformation (3.4\AA , Besetzung 44%) zum Carbonylsauerstoff des P3 D-Arg Restes möglich. Der zentrale Amidstickstoff zeigt die typische Wasserstoffbrücke zum Carbonylsauerstoff von Ser214 (2.9\AA). Das P2 Prolin liegt wie erwartet unter dem 60er Loop in der S2 Tasche. Das Peptid

des P3 Arginin zeigt die typischen β -Faltblatt-Wasserstoffbrücken zum Gly216 des Proteinrückgrats (3.0 Å und 2.7 Å). Der Hydroxymethylsubstituent in P1 verschiebt zu dem die Peptidbindung dieses Restes um 1.0 Å (**Abbildung 5.3**).

Wie auch für **MI001** beobachtet, ist die Arginin-Seitenkette zu lang um die zudem unpolare S3 Tasche direkt zu adressieren. Das Arginin bindet daher am unteren Rand der Tasche und bildet mit seiner Guanidiniumfunktion ladungsgestützte Wasserstoffbrücken zu Glu217 aus (3.0 Å und 3.2 Å) oder bei dem verwendeten pH von 7.5 vielleicht sogar eine Salzbrücke. Letztlich bindet der Benzylsulfonylrest in P4 wie bei Inhibitoren dieses Typs oft beobachtet U-förmig oberhalb der S1 Tasche, wobei eine Wasserstoffbrücke zwischen einem der Sulfonylsauerstoffe und dem Amidstickstoff von Gly219 ausgebildet wird (3.0 Å). Hierdurch ist es für Glu192 nur möglich mit einer Konformation in die umgebende Lösung zu zeigen, da der Phenylring der P4 Gruppe eine Orientierung oberhalb der S1 Tasche verhindert.

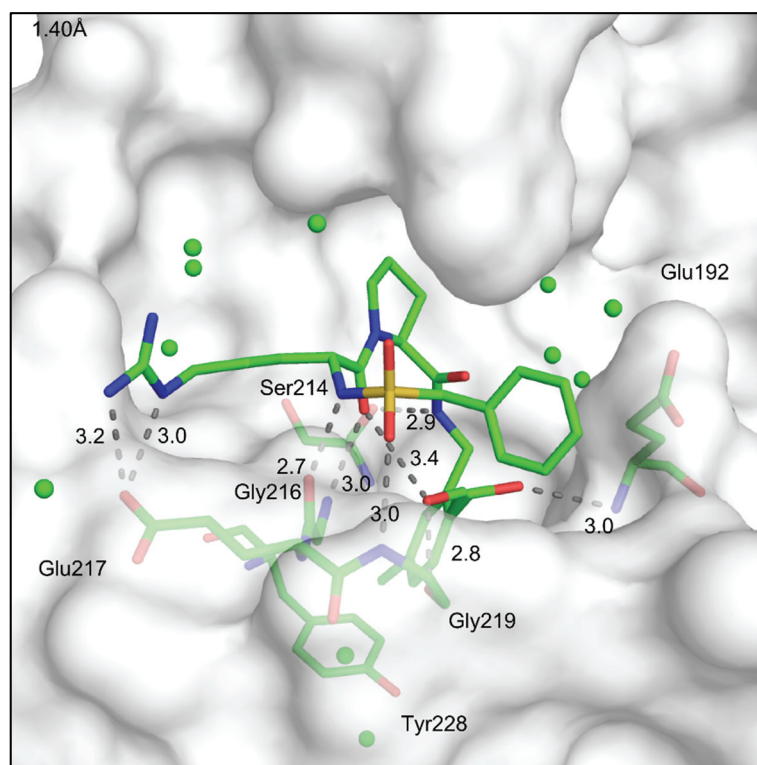


Abbildung 5.3 Bindungsmodus des Inhibitors **MI067** im Komplex mit Thrombin. Wassermoleküle im Radius von 5Å um den Liganden sind als Kugeln, der Ligand und die interagierenden Proteinreste sind in weiß dargestellt. Interaktionen mit dem Protein sind als graue gestrichelte Linien gezeigt. Stickstoffatome sind in blau, Sauerstoffatome in rot, Schwefelatome in gelb und Kohlenstoffatome in der Farbe des Inhibitors gezeigt.

Die Komplexstruktur von **MI001** (PDB-Code 3U98; $K_i = 2.22 \pm 2.40$ nM) wurde von Dr. Biela mit einer Auflösung von 1.40 Å bestimmt (**Abbildung 5.4**). Der Bindungsmodus zeigt den P1 Benzamidinanker in der S1 Tasche und das Benzylsulfonylamid in der Nähe einer Disulfidbrücke. Dieser U-förmige Bindungsmodus wurde bisher bei allen Derivaten beobachtet, die in der P3 Position eine D-Aminosäure und das Benzylsulfonylamid in P4 enthielten (siehe auch Kapitel 3 und 4). Die polare Seitenkette des D-Arg wurde hier auch nicht in die unpolare S3/4 Tasche platziert. Die Guanidinium-Gruppe des D-Arg interagiert ebenfalls nur mit dem Glu217 am unteren Rand der Bindetasche. Aus der S3/4 Tasche werden auch hier keine Wassermoleküle durch **MI001** verdrängt. Interessanterweise stimmt der Bindungsmodus des Restmoleküls, abgesehen von der P1 Gruppe mit dem Bindungsmodus von **MI067** komplett überein (**Abbildung 5.4**).

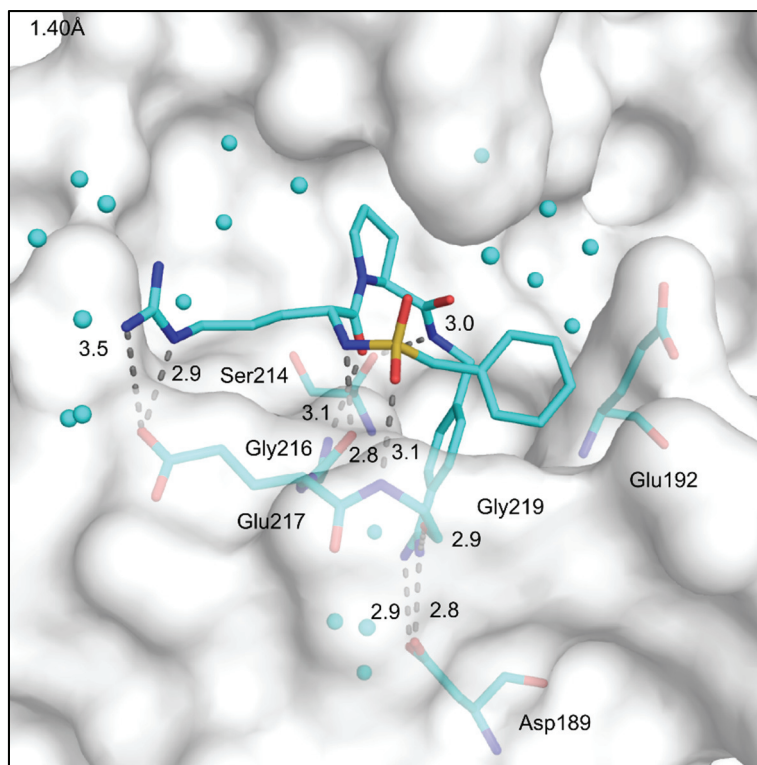


Abbildung 5.4 Bindungsmodus des Inhibitors **MI001** im Komplex mit Thrombin. Wassermoleküle im Radius von 5 Å um den Liganden sind als Kugeln, der Ligand und die interagierenden Proteinreste sind in türkis dargestellt. Interaktionen mit dem Protein sind als graue gestrichelte Linien gezeigt. Stickstoffatome sind in blau, Sauerstoffatome in rot, Schwefelatome in gelb und Kohlenstoffatome in der Farbe des Inhibitors gezeigt

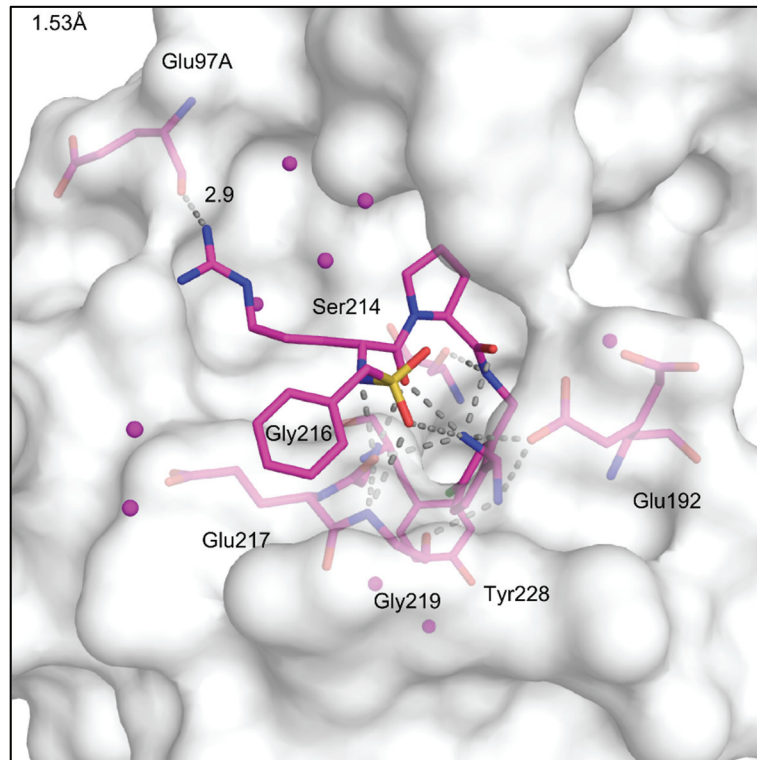


Abbildung 5.5 Bindungsmodus des Inhibitors **MI002** im Komplex mit Thrombin. Wassermoleküle im Radius von 5 Å um den Liganden sind als Kugeln, der Ligand und die interagierenden Proteinreste sind in magenta dargestellt. Interaktionen mit dem Protein sind als graue gestrichelte Linien gezeigt. Stickstoffatome sind in blau, Sauerstoffatome in rot, Schwefelatome in gelb und Kohlenstoffatome in der Farbe des Inhibitors gezeigt.

Der Inhibitor **MI002** (PDB-Code 3EQ0; $K_i = 0.88 \pm 0.55$ nM) wurde bereits in der Doktorarbeit von Dr. Baum kristallographisch untersucht und mit einer Auflösung von 1.53 Å bestimmt (**Abbildung 5.5**). Die Aminomethylgruppe des P1 Restes ist in der Lage, eine Konformation des Glu192 zu rekrutieren, was den Benzylsulfonylamidrest in P4 in eine unterschiedliche Position zwingt. Der Inhibitor mit dem ACB Anker hat so im Komplex nicht die Möglichkeit einen U-förmigen Bindungsmodus einzunehmen, da die Glu192-Rekrutierung diese Geometrie aus sterischen Gründen verhindert. Der D-Arg-Rest zeigt hierdurch bedingt ebenfalls einen unterschiedlichen Bindungsmodus, kann aber hier eine Wasserstoffbrücke zum Carbonylsauerstoff von Glu97A ausbilden.

Während für **MI001** und **MI067** eindeutig nur eine Konformation des Argininrestes zu sehen ist, wurden in der Struktur **MI002** von Dr. B. Baum zwei Konformationen in die Dichte modelliert. Eine Überlagerung aller Strukturen ist in **Abbildung 5.6** zu sehen. Die zweite

Konformation in **MI002** ist allerdings nur zu 30 % besetzt und zeigt eine ähnliche Position wie das D-Arg in **MI001** und **MI067**. Neben dieser für ein Ligand Molekül problematisch niedrigen Besetzung bestehen zusätzlich zwei weitere Argumente, die diese zweite Konformation in fraglich erscheinen lassen. Das D-Arg ist in einer ungewöhnlich kurzen Distanz nahe dem Benzyl-Rest in der P4 Position Inhibitors (2.4 Å) und die $2F_o - F_c$ Elektronendichte zeigt keine überzeugenden Daten für eine zweite Konformation. Wie bereits von Dr. A. Biela diskutiert wurde, sprechen diese Beobachtungen eher für ein zweites Wassermolekül in der Nähe der Guanidinium-Gruppe von **MI002** als einer zweiten Konformation der Arginin Seitenkette. [1]

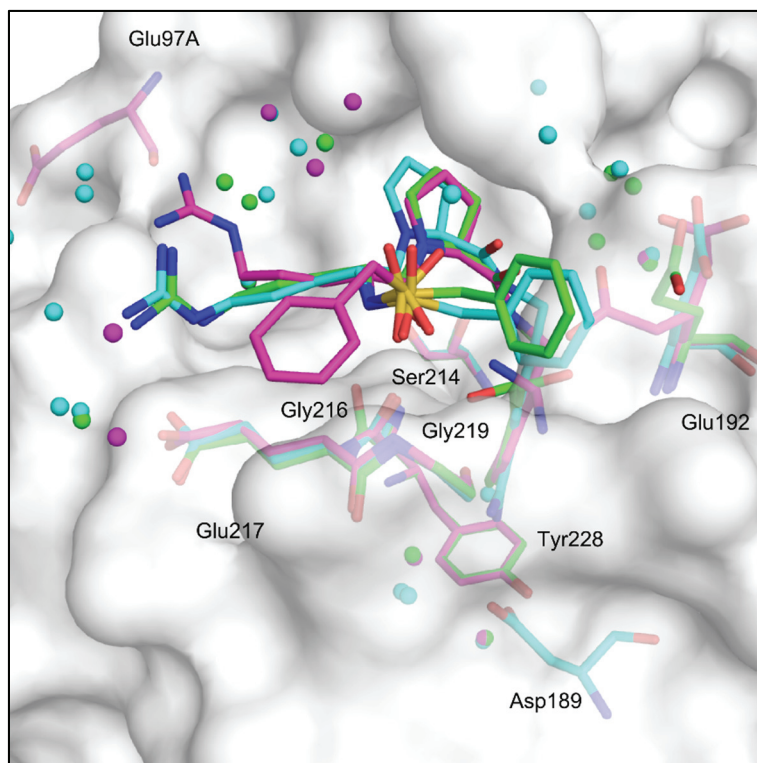


Abbildung 5.6 Überlagerung der Strukturen von **MI002** (magenta) mit sinnvoller D-Arg Konformation, **MI001** (türkis) und **MI067** (grün). Wassermoleküle im Radius von 5Å um den Liganden sind als Kugeln, der Ligand und die interagierenden Proteinreste sind in den Farben des jeweiligen Liganden dargestellt. Die Farbcodierung entspricht den vorherigen Abbildungen. Die gezeigte Proteinoberfläche stammt von **MI002**.

Aus diesem Grund soll diese zweifelhafte zweite Konformation hier nicht weiter diskutiert werden und wurde für den besseren direkten Vergleich der Strukturen entfernt und fehlt in den **Abbildungen 5.5** und **5.6**.

5.3 Thermodynamik und Affinitätsdaten

Trotz der strukturellen Ähnlichkeit zeigen die Inhibitoren **MI001** ($K_i = 2.22 \pm 2.40$ nM), **MI002** ($K_i = 0.88 \pm 0.55$ nM) und MI067 ($K_i = 2541,50 \pm 846,47$ nM) große Unterschiede in ihrer Bindungsaffinität. Die Bindungsaffinität von **MI002** ist mit Abstand am höchsten und um etwa den Faktor drei höher als für **MI001**, beziehungsweise sogar um den Faktor 3000 höher als für **MI067**. Dies ist umso bemerkenswerter, da der strukturelle Unterschied zwischen **MI002** und **MI067** lediglich in einem Schweratom und einem Wasserstoffatom besteht.

Für die zweite Serie **MI073** ($K_i = 0.28 \pm 0.05$ nM), **MI074** ($K_i = 0.19 \pm 0.11$ nM) und **MI066** ($K_i = 24.14 \pm 2.19$ nM) wurden in der Arbeitsgruppe Steinmetzer nur Affinitätsdaten bestimmt. Bei den Bindungsaffinitäten zeigen sich ebenfalls für die P1 Substituenten Unterschiede, jedoch sind sie weitaus kleiner als für die erste Serie. Das Aminomethylderivat **MI074** weist gegenüber dem Benzamidinderivat **MI073** nur eine etwa um den Faktor 1.5 höhere und gegenüber dem Hydroxymethylderivat **MI066** eine um etwa den Faktor 100 höhere Affinität auf.

Die thermodynamische Charakterisierung (**Abbildung 5.7**) zeigt ebenfalls enorme Unterschiede zwischen den drei S1 Anker auf. Die Enthalpie der Interaktion ist bei allen Liganden pufferabhängig aber vergleichbar (**Abbildung 5.8**), was auf die Deprotonierung des katalytischen His57 während der Ligandbindung zurück zu führen ist.

Die thermodynamischen Daten zeigen, dass die Interaktion mit dem ACB Anker von **MI002** wesentlich enthalpischer ($\Delta H_{\text{Korr.}} = -46.5$ kJ/mol) abläuft als die Interaktion mit dem korrespondierenden AMBA-Derivat **MI001** ($\Delta H_{\text{Korr.}} = -16.4$ kJ/mol). Das ACB-Derivat **MI002** ist zudem stärker affin ($\Delta G = -53.3 \pm 0.3$ kJ/mol) als das analoge AMBA-Derivat **MI001** ($\Delta G = -46.5 \pm 0.1$ kJ/mol), was die Ergebnisse der Affinitätsmessungen aus dem Enzymassay bestätigt. Diese Unterschiede in der Bindungsaffinität zwischen ACB- und AMBA-Anker zeigten sich auch bei verwandten Inhibitoren, die bereits von Dr. Biela vermessen wurden.

[1]

Das HCB Derivat **MI067** ($\Delta G = -30.0 \pm 0.6$ kJ/mol) zeigt auch hier eine bedeutend schwächere Affinität im Vergleich zu den anderen Inhibitoren. Die Korrelation der Daten der freien Bindungsenergie aus dem Enzymassay und den ITC Messungen ist durchwegs sehr gut ($R^2=0.98$) (**Abbildung 5.9**).

Die gemessene Bindungsenthalpie ($\Delta H_{\text{Korr.}} = -10.0$ kJ/mol) ist nochmals niedriger als die des Benzamidinderivats. Interessanterweise zeigen die Enthalpie- und Entropieterme der beiden Liganden das gleiche Aufspaltungsmuster, wobei die gemessene Bindungsenthalpie nur in etwa halb so groß wie der entropische Bindungsbeitrag ausfällt (**Abbildung 5.7**).

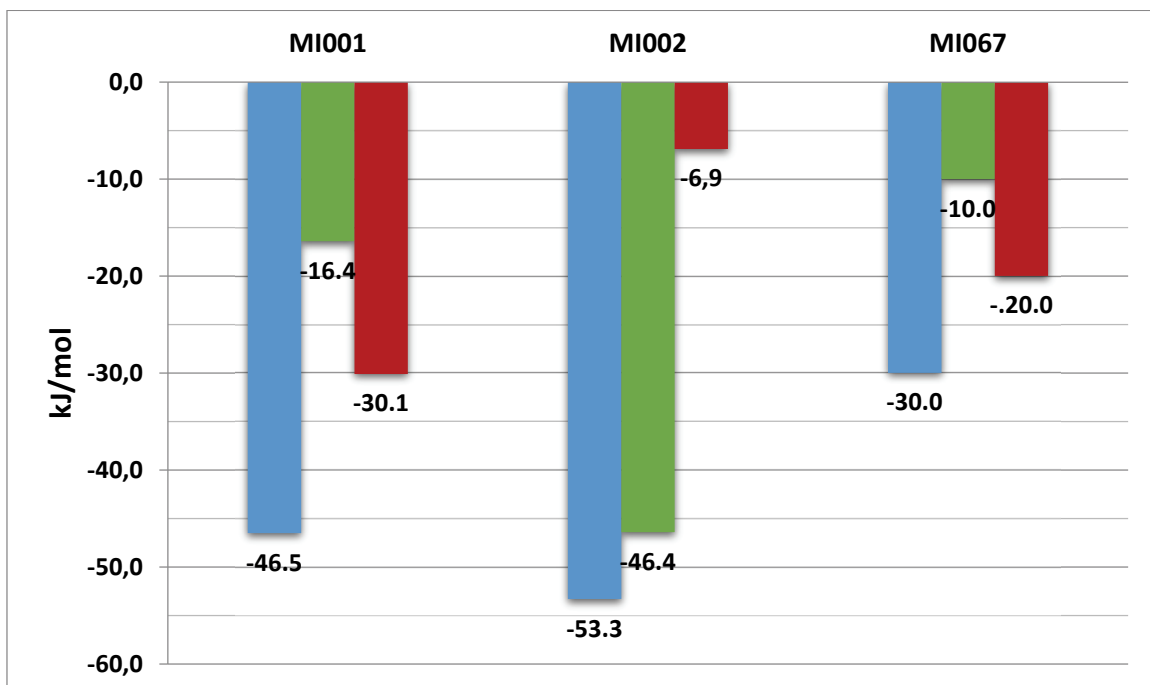


Abbildung 5.7 Im Diagramm sind die mit ITC gemessenen Bindungsterme ΔG^0 (blau); $\Delta H^0_{\text{Korr.}}$ (grün) und $-T\Delta S^0$ (rot) (kJ/mol) der untersuchten Liganden aufgetragen. (MI001 und MI002 wurden von Dr. A Biela vermessen [1])

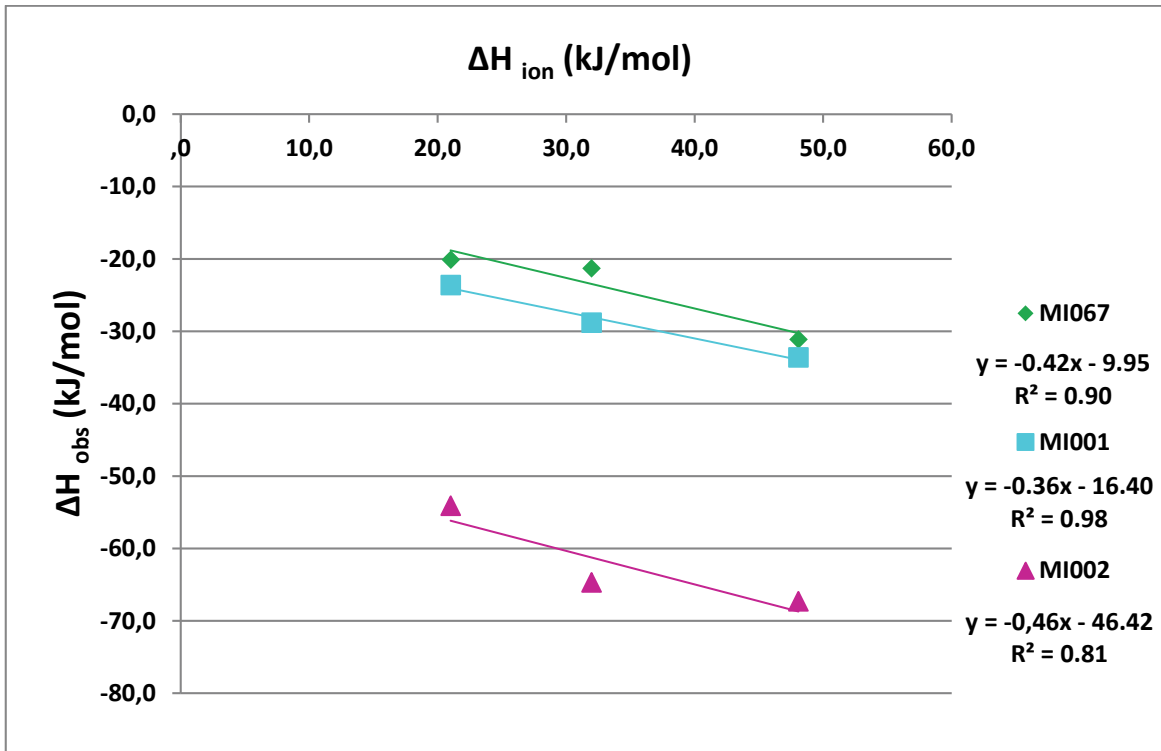


Abbildung 5.8 Im Diagramm sind die gemessenen Enthalpien (ΔH_{obs}) der Liganden gegen die Ionisierungsenthalpien (ΔH_{ion}) der jeweiligen Puffer (TRIS, TRICIN und HEPES) aufgetragen.

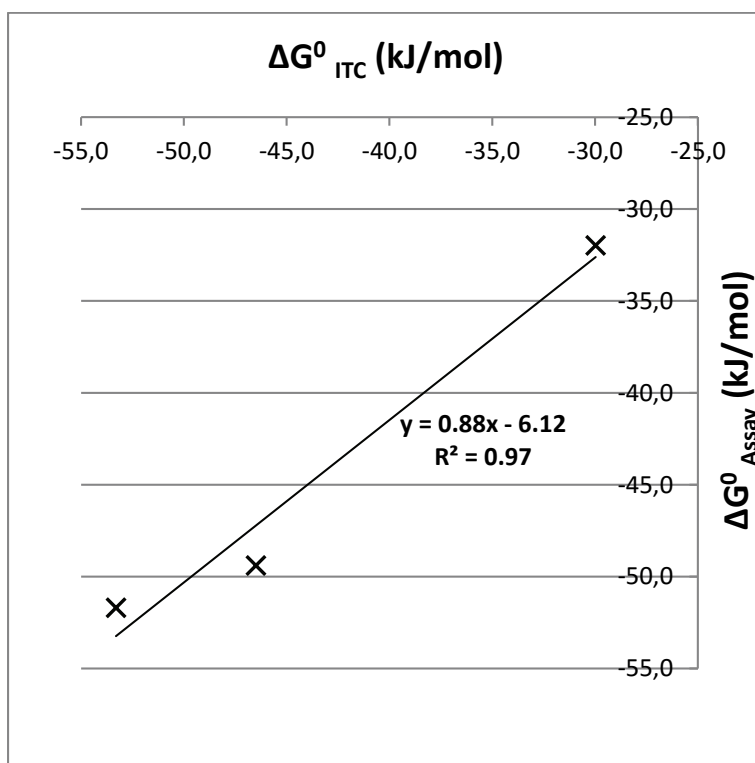


Abbildung 5.9 Die in den Enzymassays gemessene Bindungsaffinität der Liganden (**MI001**, **MI002** und **MI067**) ist umgerechnet in ΔG^0 (kJ/mol) gegen die in den ITC Experimenten ermittelte freie Bindungsenergie ΔG^0 (kJ/mol) aufgetragen.

5.4 Diskussion und Zusammenfassung der Daten

Die normalerweise für reine m-Chloraromaten in P1 beobachtete hohe Bindungsenthalpie, die auf vorteilhafte Desolvatation und Verdrängung eines weiteren Wassermoleküls aus der S1 Tasche zurückzuführen ist, lässt sich für **MI067** nicht erkennen. Eine intramolekulare Wechselwirkung wie für **MI002** konnte nicht beobachtet werden und ist wahrscheinlich neben den anderen Bindungsmodi der P3 und P4 Substituenten für die schwache Bindungsaffinität von **MI067** verantwortlich. Die Bindungsgeometrie des Hydroxymethylrestes von **MI067** scheint darüber hinaus für ungünstige Interaktionen mit dem Proteinerückgrat verantwortlich zu sein, welche durch eine Verschiebung von Gly216 zum Ausdruck kommt.

Die Arginyl-Seitenkette kann trotz ausreichendem Platz bei allen Liganden nicht in die S3/S4-Tasche binden, stattdessen verbleibt die Tasche vollständig solvatisiert und die

Argininseitenketten binden an den Rändern der Tasche an Glutamatreste und müssen so nur teilweise desolvatisiert werden. Scheinbar ist der Prozess der Desolvatisierung der Seitenkette und der Bindetasche insgesamt ein energetisch ungünstiger Prozess.

Die hohe Affinität des Inhibitors **MI002** resultiert daher wahrscheinlich aus optimalen intra- und intermolekularen Wasserstoffbrücken im zentralen Teil des Komplexes. Sollte diese Konformation des Inhibitors derjenigen des Inhibitors in Lösung entsprechen, könnte dieser nur bei diesem Inhibitor mögliche Präorganisationseffekt eine weitere mögliche Erklärung für die hohe Bindungsaffinität liefern.

Diese Vermutung wird durch die Affinitätsmessungen der zweiten Serie gestützt, wo die Unterschiede in der Bindungsaffinität ohne den P4 Substituenten, der für einen Teil der intramolekularen Fixierung essentiell ist, deutlich geringer ausfallen. Um diese Aspekte weiter zu überprüfen sollten für die Inhibitoren der zweiten Serie noch ITC-Messungen durchgeführt werden und hochaufgelöste Röntgenstrukturen gesammelt werden.

5.5 Methoden

5.5.1 Assays

Die Bestimmung der Affinitätswerte erfolgte mittels photometrischem und fluoreszenzbasierten Assays analog wie in den vorangegangenen Kapiteln 2 und 4 beschrieben.

5.5.2 ITC

Alle untersuchten Liganden wurden mittels Verdrängungstitrationen vermessen. Für den schwach bindenden Liganden **MI067** wurde der stark bindende Ligand **MI331** (BzIs-Gly-Pro-ACB) als Verdrängungsligand verwendet. Der Ligand **MI331** sowie **MI001** und **MI002** wurden von Dr. Biela [1] thermodynamisch charakterisiert.

Die ITC-Experimente wurden unter Verwendung eines ITC200™-System von GE Healthcare (Northampton, MA, USA) durchgeführt. Thrombin wurde aus Beriplast® von CSL Behring (Marburg, Deutschland) aufgereinigt. Es wurde hierfür durch Dialyse unter Verwendung

eines experimentellen Puffers aus 50 mM TRIS, TRICINE oder HEPES, 100 mM NaCl, 0.1% Polyethylenglykol 8000 bei pH 7.8 extrahiert. Anschließend wurde das Protein für alle Titrations verwendet.

Liganden aus 50mM DMSO-Stammlösungen wurden zu 0.50 mM in ITC-Puffer verdünnt und anschließend die DMSO-Konzentration in allen Lösungen auf 3% eingestellt. Die Titrations wurden mindestens dreimal durchgeführt.

Alle ITC-Experimente wurden bei 25 °C mit einer Referenzleistung von 5 kcal/s gestartet, nachdem eine stabile Basislinie erreicht wurde. Die initiale Verzögerung wurde auf 300 s eingestellt. Injektionen von 1,4µL mit mindestens 150s Intervall zwischen jeder Injektion wurden für Titrations der Liganden verwendet. Die Datenintegration und Evaluierung wurde mit Origin 7 durchgeführt.

5.5.3 Röntgenstruktur

5.5.3.1 Kristallisation

Humanes α -Thrombin (von Enzyme Research Laboratories, South Bend, USA) wurde in Kristallisationspuffer (20 mM NaH_2PO_4 , 350 mM NaCl, 2 mM Benzamidin, pH 7.5) mit einer Konzentration 10 mg/mL des Lyophilisats gelöst.

Das sulfatierte Hirudinfragment (54-65; Bestellnummer: H-7425) von Bachem (Bubendorf, Schweiz) wurde ebenfalls in Kristallisationspuffer mit 2,5 mg/mL gelöst.

Im nächsten Schritt wurden 40 µL der Lösung des Hirudinfragments mit 160 µL der Thrombinlösung gemischt. Nach Inkubation von 1 h bei 4 °C wurde die Kristallisation bei 4 °C mittels der Hanging-Drop-Methode durchgeführt.

Die Hirudin/Thrombin-Lösung wurde 1:1 mit der Reservoirlösung (20 mM NaH_2PO_4 , 27% Polyethylenglykol 8000, pH 7.5) vermischt und anschließend wurden 2 µL dieser Lösung in der Mitte eines Deckglases platziert. Unmittelbar nach dem Mischen des Proteins und der Reservoirpufferlösung wurde ein Microseeding (Übertragen von Kristallisationskeimen) mit einem Pferdehaar durchgeführt. Die Vertiefungen der Kristallisationsplatten wurden mit

500 μL der Reservoirlösung gefüllt. Anschließend wurde mit den Deckgläsern die Vertiefungen abgedeckt und mit Silikon versiegelt. Kristalle von guter Streuqualität konnten nach 10 bis 14 Tagen beobachtet werden.

Zum Erreichen der endgültigen Inhibitorkonzentration (5 mM des Inhibitors und 10% DMSO) für den Kristall wurde eine DMSO-Lösung des Inhibitors (50 mM) im Verhältnis 1:10 mit einer 1:1 Lösung aus Kristallisations- und Reservoirpuffer verdünnt.

Ein mittelgroßer Kristall ohne sichtbare Mängel wurde ausgewählt und in die Inhibitor Lösung für 1h überführt.

5.5.3.2 Datensammlung und Verfeinerung

Der Datensatz für **MI067** wurde am Synchrotron BESSY (Berlin) an der Beamline 14.2 (Wellenlänge 0.9184Å) mit einem Rayonix MX-225 Detektor gesammelt. Die Prozessierung und Skalierung der Daten erfolgte mit dem Programm XDS [3].

Für die initiale Verfeinerung wurde ein von Dr. Johannes Schiebel entwickeltes automatisches Refinement Skript verwendet, das auf den Programmen Phaser (CCP4 Suite) [4], PHENIX [5] und Coot [6] basiert und die folgenden Schritte beinhaltet:

1. Molecular replacement mit dem Programm Phaser (Modell 4UE7)
2. Simulated annealing mit dem Programm Phenix
3. Standard Verfeinern (isotrope B-Faktoren, Besetzung und Koordinaten) mit dem Programm Phenix
4. TLS Verfeinern mit dem Programm Phenix
5. Automatisches hinzufügen von Wassermolekülen mit dem Programm Coot und anschließendes Verfeinern mit Phenix
6. Entfernen von falsch gesetzten Wassermolekülen mit dem Programm Coot und anschließendes Verfeinern mit Phenix
7. Anisotropes Verfeinern (außer Wassermoleküle) mit dem Programm Phenix
8. Anisotropes Verfeinern aller Atome mit dem Programm Phenix

9. Automatisches hinzufügen von Wasserstoffatomen und anschließendes Verfeinern mit Phenix
10. Erneutes automatisches hinzufügen von Wassermolekülen mit dem Programm Coot und anschließendes Verfeinern mit Phenix
11. Entfernen von falsch gesetzten Wassermolekülen mit dem Programm Coot und anschließendes Verfeinern mit Phenix

Nach diesen Verfeinerungszyklen wurden weitere Wassermoleküle und anschließend Ionen und Liganden in die Elektronendichte eingebettet und dem Modell hinzugefügt.

Multiple Seitenkettenkonformationen wurden modelliert, wenn während der Verfeinerung eine geeignete Elektronendichte beobachtet wurde und wenn die Besetzung der Seitenkette mindestens 20% zeigte.

Das endgültige Modell wurde mit in PHENIX implementierten Validierungsoptionen oder MolProbity [7] validiert. Die Ramachandran-Plots wurden mit MolProbity berechnet. Die Parameter der Datensammlung, der Elementarzelle und die Verfeinerungsstatistiken sind in **Tabelle 6.1** ersichtlich. Die Analyse der Temperaturfaktoren wurde mit MolProbity durchgeführt. Die Benennung der Protein-Aminosäuren erfolgte nach Bode et al. [8]. Koordinaten und Strukturfaktoren wurden in der Proteindatenbank mit dem folgenden Code hinterlegt: 5AM2.

<i>Compound (PDB-Code)</i>	<i>MI067 (5A2M)</i>
Wavelength (Å)	0.9184
Resolution range (Å) ^a	35.63 - 1.40 (1.45 - 1.40)
Space group	C 1 2 1
Unit cell a, b, c (Å)	70.1 71.2 72.4
α, β, γ (°)	90.0 100.1 90.0
Total reflections	232047 (22863)
Unique reflections	68384 (6775)
Multiplicity	3.4 (3.4)
Completeness (%)	99.1 (98.3)
Mean I/sigma(I)	19.1 (2.6)
Wilson B-factor	15.4
R-merge	0.037 (0.520)
R-work	0.137 (0.243)
R-free	0.159 (0.263)
Number of non-hydrogen atoms	2589
Macromolecules	2306
Ligands	49
Water	234
Protein residues	279
RMS(bonds) (Å)	0.008
RMS(angles) (°)	1.28
Ramachandran favored (%) ^b	96
Ramachandran outliers (%) ^b	0
Ramachandran allowed (%) ^b	4
Average B-factor (Å ³)	25.0
Macromolecules	24.3
Ligands	26.2
Solvent	34.9
Coordinate errors (Å) ^c	0.13

^a Statistics for the highest-resolution shell are shown in parentheses.

^b Calculated by MOLPROBITY

^c Maximum-likelihood based

Tabelle 5.1 Datensammlungs- und Verfeinerungsstatistiken für **MI067**.

5.6 Referenzen

- [1] A. Biela, Dissertation, Philipps Universität Marburg, Germany, (2012)
- [2] B. Baum, Dissertation, Philipps Universität Marburg, Germany, (2009)
- [3] W. Kabsch, XDS, *Acta Crystallogr. D Biol. Crystallogr.* 66 (2010) 125–132.
- [4] The CCP4 suite: programs for protein crystallography, *Acta Crystallogr. D Biol. Crystallogr.* 50 (1994) 760–763.
- [5] P.D. Adams, P.V. Afonine, G. Bunkóczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.-W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, P.H. Zwart, PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr. D Biol. Crystallogr.* 66 (2010) 213–221.
- [6] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, *Acta Crystallogr. D Biol. Crystallogr.* 60 (2004) 2126–2132.
- [7] V.B. Chen, W.B. Arendall, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, D.C. Richardson, MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallogr. D Biol. Crystallogr.* 66 (2010) 12–21.
- [8] W. Bode, I. Mayr, U. Baumann, R. Huber, S.R. Stone, J. Hofsteenge, The refined 1.9 Å crystal structure of human alpha-thrombin: interaction with D-Phe-Pro-Arg chloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion segment, *EMBO J.* 8 (1989) 3467–3475.

6 Röntgenstruktur des Inhibitors MI432 in Thrombin zur Modellierung eines Matriptase-Komplexes

6.1 Einleitende Bemerkungen/Introductory Remarks

Diese Arbeit wurde in Kooperation mit der Arbeitsgruppe von Prof. Dr. Steinmetzer durchgeführt. Das Manuskript für die Veröffentlichung wurde, wie mit unserem Kooperationspartner abgestimmt von Dr. Maya Hammami als Erstautor in englischer Sprache angefertigt und in ChemMedComm publiziert. [1] Die Synthese und Messung der Enzymassays der Inhibitoren erfolgte in der Arbeitsgruppe von Prof. Dr. Steinmetzer.

Dieses Kapitel dient der detaillierten Beschreibung der vom Autor für diese Kooperation erstellten Thrombinkristallstruktur, welche in dieser Publikation nicht in aller Ausführlichkeit beschrieben werden konnte. Der vorliegende Text wird daher in dieser Form in keiner Zeitschrift, sondern nur im Rahmen dieser Dissertation dokumentiert.

6.2 Einleitung

Neue vom 3-Amidinophenylalanin abgeleitete Matriptase Inhibitoren wurden in der Arbeitsgruppe von Prof. Dr. Steinmetzer entwickelt und an den Trypsin-ähnlichen Serin-Proteasen Matriptase, Matriptase 2, Thrombin und Faktor Xa auf deren Affinität und Selektivität getestet.

Die Röntgenkristallstruktur eines repräsentativen Analogons im Komplex mit Thrombin wurde mit der Matriptasestruktur eines verwandten Inhibitortyps (PDB-Code 2gv6 [2]) überlagert. Dieses Vorgehen war nötig, da Matriptase nicht in ausreichenden Mengen für Kristallisationsversuche zur Verfügung stand. Aus der Überlagerung dieser Strukturen konnten Informationen hinsichtlich der in dieser Studie beobachteten Selektivitätsprofile und des möglichen Bindungsmodus in Matriptase gezogen werden.

Die Strukturlösung dieses Thrombinkomplexes gestaltete sich zunächst schwierig, da durch das *Soaking* (durch osmotischen Druck und Konzentrationsunterschiede bedingtes Eindiffundieren in den Kristall) des Inhibitors in den Thrombinkristall die Kristallpackung derart verändert wurde, dass die normalerweise für Thrombin unter diesen Kristallisationsbedingungen gefundene Raumgruppe C2 nicht mehr gegeben war.

Die Struktur konnte schließlich allerdings in der Raumgruppe P1 bestimmt werden, was zur Folge hatte, dass zwei Proteinmoleküle in der asymmetrischen Einheit gefunden wurden.

(Abbildung 6.1)

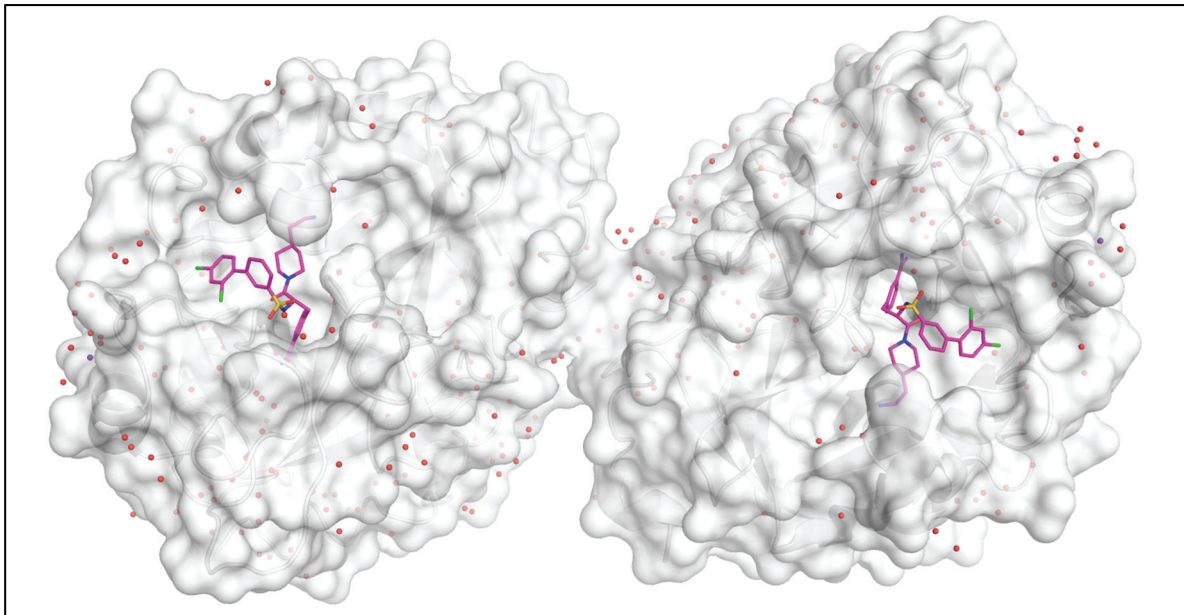


Abbildung 6.1 Gezeigt sind die zwei Thrombinmoleküle der asymmetrischen Einheit. Die Proteinoberfläche ist in weiß gezeig, Wassermoleküle als rote Kugeln und der Inhibitor in Magenta.

Bei anderen Matriptase Inhibitoren, die im Bereich des P3 Restes sterisch noch anspruchsvoller waren, führte das *Soaking* sogar zu einer Zerstörung der Kristalle. Eine Cokristallisation kam in Ermangelung entsprechender Bedingungen für Thrombin nicht in Frage.

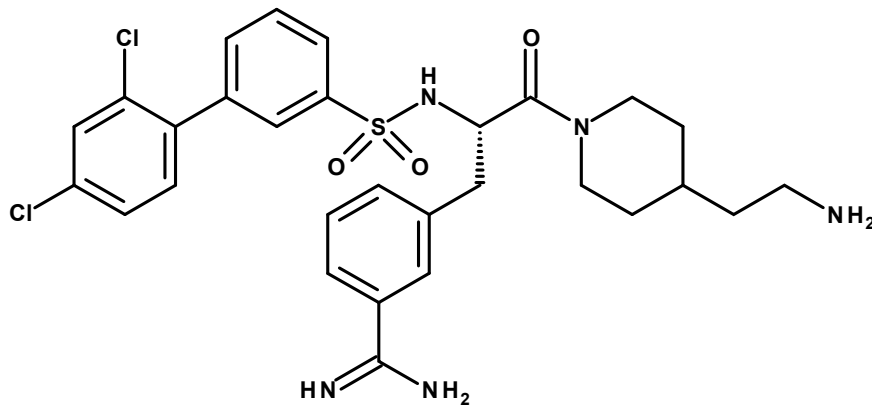
Der für die Untersuchung verwendete Inhibitor **MI432** zeigte gegen Thrombin eine hohe Affinität von 20 nM (Matriptase: 2 nM; Matriptase 2: 110 nM; Faktor Xa: 50 nM).

6.3 Röntgenstruktur

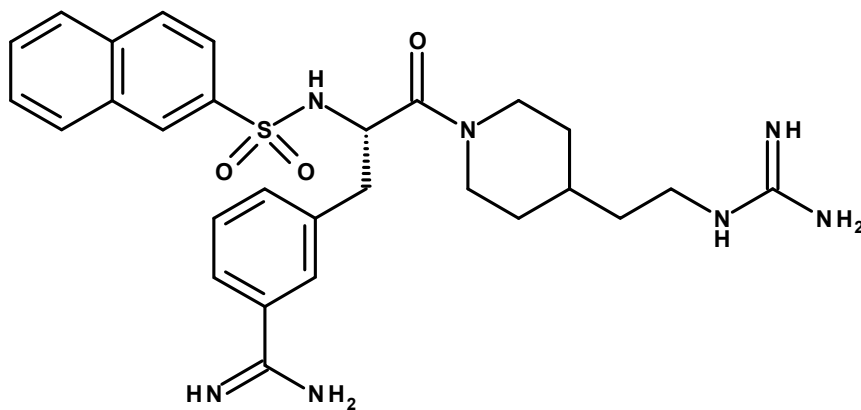
Durch Zeitdruck für die Veröffentlichung der Röntgenstruktur wurden die Daten nicht am Synchrotron gemessen. Daher konnte bedingt durch die intensitätsschwache, hauseigene Röntgenstrahlungsquelle und die vergleichsweise schlechten R-Werte durch die Veränderung der Kristallpackung (**Tabelle 6.1**) die Struktur nur mit einer verhältnismäßig niedrigen Auflösung von 2.25 Å gelöst werden.

6.3.1 Interaktion mit Thrombin

Der Inhibitor **MI432** zeigt nicht das für Thrombin typische Peptidgrundgerüst, eines Tripeptidanalogs. (**Abbildung 6.2**)



MI432



CJ-730 (PDB-Code 2gv6)

Abbildung 6.2 Gezeigt sind die Strukturformeln des untersuchten Inhibitors MI432 und des verwandten Matriptaseinhibitors CJ-730 (PDB-Code 2gv6).

In P1 Position trägt der Inhibitor ein in 3-Position mit dem P2 Rest substituiertes Benzamidin, welches die S1 Tasche von Thrombin adressiert. (**Abbildung 6.3**)

Dort bildet das Amidin eine Salzbrücke zum deprotonierten Asp189 (2.9 Å und 2.8 Å) am Boden der S1 Tasche aus. Die unterschiedlichen Abstände resultieren aus einer

Abweichung von der Koplanarität von 21.9° der Amidinstickstoffatome zur Carboxylatgruppe von Asp189.

Zusätzlich werden Wasserstoffbrücken von einem der Amidinstickstoffe zu den Carbonylsauerstoffen von Ala190 (3.0 Å) und Gly219 (2.9 Å) ausgebildet.

Durch die Substitution in 3-Position, welche auch für Piperidinoamidinhibitoren (Napsagatran) Anwendung findet, deckt Glu192 die S1 Tasche teilweise ab.

Der Piperidinring in P2 liegt teilweise unter dem 60iger *Loop* der S2 Tasche, wobei sich der Aminoethylensubstituent in Richtung S1' Tasche orientiert ohne eine direkte Interaktion mit der Proteinbindetasche einzugehen. Der Piperidid Sauerstoff bildet zusammen dem Sulfonylamidstickstoff des P3 Restes β -faltblattartige Wasserstoffbrücken zu Gly216 (2.9 Å und 2.8 Å) des Proteinrückgrats, welches eine sehr wichtige und häufig beobachtete Interaktion für viele Thrombininhibitoren darstellt. (**Abbildung 6.3**)

Ein Sauerstoffatom der Sulfonylgruppe des P3/4 Substituenten kann darüber hinaus eine Wasserstoffbrücke zum Stickstoff von Gly219 (2,7 Å) ausbilden.

Das 2,4-Dichloro-biphenyl-Ringsystem dieses Substituenten adressiert schließlich mit seinem 2,4-Dichlorophenylring den unteren Rand der S3/4 Tasche oberhalb des aromatischen Ringsystems von Trp215, wobei sich der zweite Phenylring um etwa 60° aus den Ringebeine des ersten Rings heraus verdreht und in Richtung des Lösungsmittels orientiert.

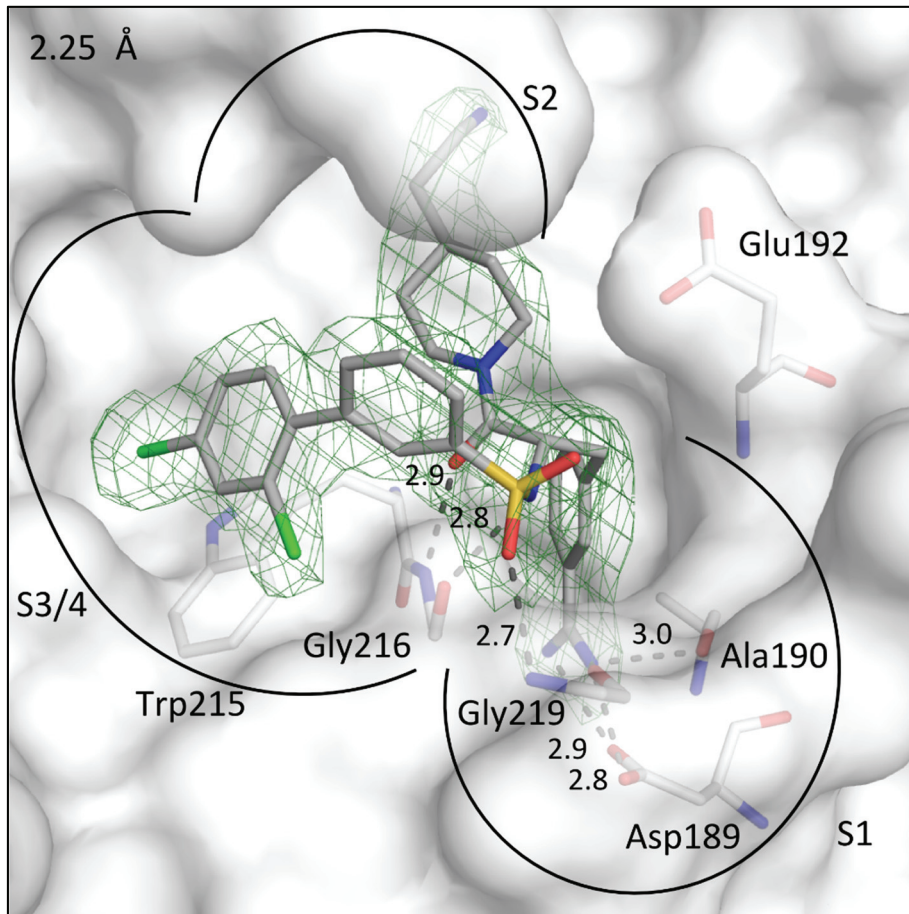


Abbildung 6.3 Ansicht des Thrombin/MI432 Komplexes. Thrombin wird mit seiner lösungsmittelzugänglichen Oberfläche (weiß) gezeigt, die Atomtypen der interagierenden Aminosäuren und des Inhibitors sind farbig hervorgehoben (Kohlenstoff in grau/weiß, Sauerstoff in rot, Stickstoff in blau, Schwefel in gelb, Chlor in grün). Die $F_o - F_c$ Elektronendichte des Inhibitors ist bei 2σ als grünes Netz gezeigt. Die spezifischen Bindetaschen S1, S2 und S3/S4 sind ebenfalls angedeutet.

6.3.2 Interaktion mit Matriptase

Der Bindungsmodus des Inhibitors mit der überlagerten Matriptasestruktur zeigt einen fast identischen Bindungsmodus, obwohl die Bindetasche der Matriptase deutlich offener ist und eine weniger zerklüftete Bindetasche aufweist. (**Abbildung 6.4**)

Aus diesem Grund werden nachfolgend nur die Unterschiede zum Thrombinbindungsmodus hervorgehoben. Die Aminosäuren in der S1 Tasche sind bis auf Position 190, wo das Alanin gegen ein Serin ausgetauscht ist, identisch und somit werden auch vergleichbare Interaktionen eingegangen. Die Hydroxylfunktion der Serin-Seitenkette

in Matriptase bietet neben dem Carbonylsauerstoff der Hauptkette die Möglichkeit, eine weitere Wasserstoffbrücke zu dem zweiten Amidinstickstoff des Inhibitors auszubilden.

Oberhalb der S1 Tasche ist in Matriptase der 192er Rest ein Glutamin anstatt eines Glutamats in Thrombin. Der 60iger *Loop*, der in Thrombin die S2 Tasche nach oben abschließt, fehlt in Matriptase komplett.

Es besteht die Möglichkeit, dass im Falle des Matriptase-Komplexes eine zusätzliche Halogen- π -Wechselwirkung zwischen dem ortho-Chlorsubstituenten und dem benachbarten Phenylring von Phe99 hergestellt wird, welche dann neben dem zusätzlichen Kontakt zu Ser190 zu einer verbesserten Affinität beiträgt. Im Vergleich zur Thrombinstruktur müsste allerdings eine Drehung des Phenylrings im Matriptase-Komplex auftreten.

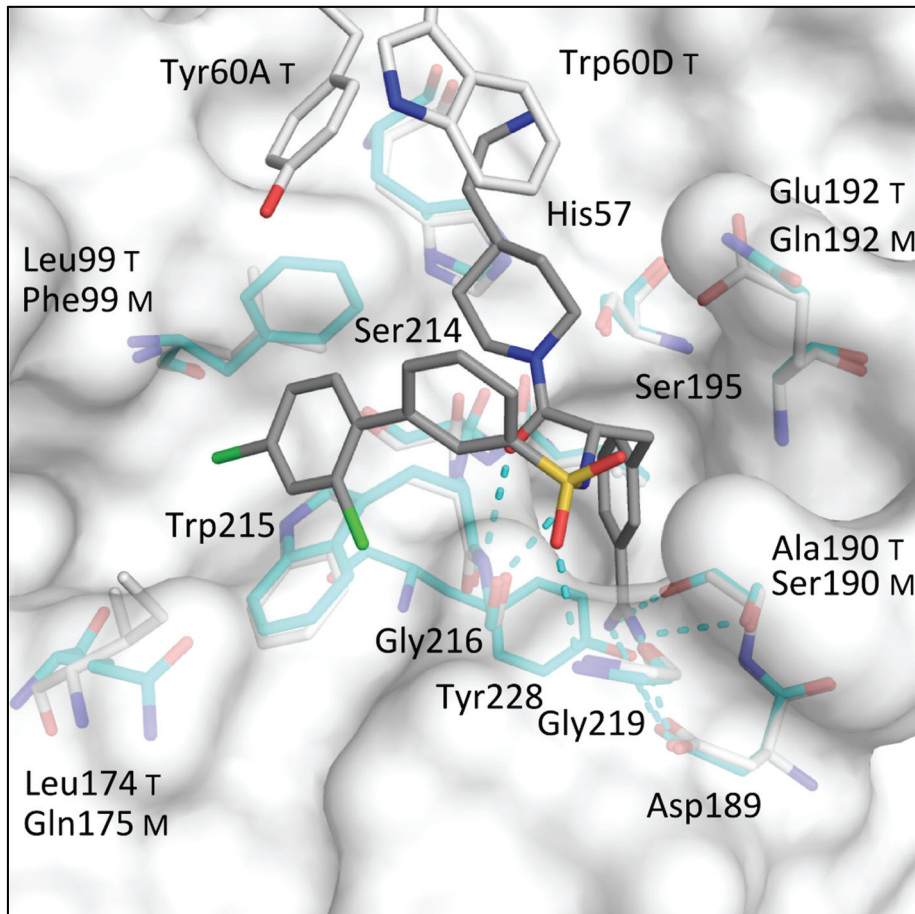


Abbildung 6.4 Gezeigt ist ein Modell des Inhibitors **MI432** in Matriptase mit seiner lösungsmittelzugänglichen Oberfläche, welches durch Überlagerung des Thrombin Komplexes mit den Matriptase Koordinaten (PDB-Code 2gv6) [2] erzeugt wurde. Identische Reste in beiden Enzymen werden durch den gemeinsamen Namen gekennzeichnet, spezifische Matriptase (hellblau) und Thrombin (weiß) Reste jeweils durch eine zusätzliche M oder T Markierung. Die möglichen polaren Wechselwirkungen sind als hellblaue Striche angezeigt.

6.4 Zusammenfassung

Durch die große Homologie der mit dem Inhibitor interagierenden Aminosäuren in der Bindetasche von Thrombin und Matriptase können zuverlässige Vorhersagen über den möglichen Bindungsmodus in Matriptase getroffen werden.

Die höhere Affinität des Inhibitors zu Thrombin scheint auf die zusätzliche Wasserstoffbrücke des Amidins zu Ser190 zurückzuführen zu sein, wohingegen die fehlende Fixierung des P2 Substituenten durch den nicht vorhandenem 60iger Loop überraschender Weise nur eine untergeordnete Rolle zu spielen scheint. Dass gerade

dieser Bereich für die Affinität von Thrombininhibitoren eine wichtige Rolle einnimmt, wurde in Kapitel 4 diskutiert.

Über eine Chlor- π -Interaktion des Dichlorphenylrings mit Phe99 der Matriptase kann hingegen nur spekuliert werden, da der Ring für diese Wechselwirkung eine um etwa 180° gedrehte Orientierung aufweisen müsste. Dies ließe sich nur mit einer Kristallisation des Inhibitors in Matriptase bestätigen.

Die Verwendung von sterisch anspruchsvollen Inhibitoren kann zu einer unerwarteten Veränderung der Raumgruppe eines Proteinkristalls durch das Eindiffundieren in den Zwischenräumen der Kristallpackung zur Bindetasche führen.

Beim *Soaking* von anderen sterisch anspruchsvollen Matriptasesinhibitoren kam es sogar zu einer Zerstörung der Proteinkristalle, was allerdings nicht weiter systematisch untersucht worden ist.

6.5 Material & Methoden

6.5.1 Assay

Details zu den verwendeten Affinitätsassays sind ausführlich in der zugehörigen Publikation dargestellt. [1]

6.5.2 Kristallisation

Humanes α -Thrombin (von Enzyme Research Laboratories, South Bend, USA) wurde in Kristallisationspuffer (20 mM NaH₂PO₄, 350 mM NaCl, 2 mM Benzamidin, pH 7.5) mit einer Konzentration 10 mg/mL des Lyophilisats gelöst.

Das sulfatierte Hirudinfragment (54-65; Bestellnummer: H-7425) von Bachem (Bubendorf, Schweiz) wurde ebenfalls in Kristallisationspuffer mit 2.5 mg/mL gelöst.

Im nächsten Schritt wurden 40 μ L der Lösung des Hirudinfragments mit 160 μ L der Thrombinlösung gemischt. Nach Inkubation von 1 h bei 4 °C wurde die Kristallisation bei 4 °C mittels der *Hanging-Drop*-Methode durchgeführt.

Die Hirudin/Thrombin-Lösung wurde 1:1 mit der Reservoirlösung (20 mM NaH₂PO₄, 27 % Polyethylenglykol 8000, pH 7.5) vermischt und anschließend wurden 2 µl dieser Lösung in der Mitte eines Deckglases platziert. Unmittelbar nach dem Mischen des Proteins und der Reservoirpufferlösung wurde ein *Microseeding* (Übertragen von Kristallisationskeimen) mit einem Pferdehaar durchgeführt. Die Vertiefungen der Kristallisationsplatten wurden mit 500 µL der Reservoirlösung gefüllt. Anschließend wurden mit den Deckgläsern die Vertiefungen abgedeckt und mit Silikon versiegelt. Kristalle von guter Streuqualität konnten nach 10 bis 14 Tagen beobachtet werden.

Zum Erreichen der endgültigen Inhibitorkonzentration (5 mM des Inhibitors und 10% DMSO) für den Kristall wurde eine DMSO-Lösung des Inhibitors (50 mM) im Verhältnis 1:10 mit einer 1:1 Lösung aus Kristallisations- und Reservoirpuffer verdünnt.

Ein mittelgroßer Kristall ohne sichtbare Mängel wurde ausgewählt und in die Inhibitor Lösung für 1h überführt.

6.5.3 Datensammlung und Prozessierung

Die Kristalle wurden für die Datensammlung unter Verwendung einer Kälteschutzmittellösung von 20 % Glycerin in Reservoirpuffer bei 110 K eingefroren.

Der Datensatz wurde im Institut mit Strahlung aus einer 1µS Mikrofokusröhre (*INCOATEC*) und einem Bildplattendetektor (*MAR345*) gesammelt. Die Datenprozessierung und Skalierung erfolgte unter Verwendung des *HKL2000* Programmpakets [3].

6.5.4 Strukturbestimmung und Verfeinerung

Die Koordinaten des humanen Thrombin (PDB-Code 1H8D) [4] wurden für den molekularen Ersatz mit *Phaser* des *CCP4* Programmpakets [5] verwendet. Für die anfängliche *Rigid-Body*-Verfeinerung des Proteinmoleküls, gefolgt von wiederholten Zyklen von Maximalwahrscheinlichkeitsenergieminimierung (*Simulated Annealing*) und B-Faktor-Verfeinerung ist das Programm *PHENIX* [6] verwendet worden. Zum Bestimmen der Temperaturfaktoren der Struktur wurde die TLS Verfeinerung angewendet. Die Definition der TLS -Gruppen wurde mit dem TLSMD Server [7,8] durchgeführt. 5% der Daten wurden

zufällig ausgewählt und zur Berechnung der R_{free} -Werte verwendet und wurden daher nicht für die Verfeinerung genutzt. Die Aminosäureseitenketten wurden in σ gewichtete $2F_o - F_c$ und $F_o - F_c$ Elektronendichtekarten mit *Coot* [9] modelliert. Nach dem ersten Verfeinerungszyklus wurden die Wassermoleküle und anschließend Ionen und Liganden in die Elektronendichte eingebettet und dem Modell hinzugefügt.

Multiple Seitenkettenkonformationen wurden modelliert, wenn während der Verfeinerung eine geeignete Elektronendichte beobachtet wurde und wenn die Besetzung der Seitenkette mindestens 20% zeigte.

Das endgültige Modell wurde mit in *PHENIX* implementierten Validierungsoptionen oder *MolProbity* [10] validiert. Die Ramachandran-Plots wurde mit *PROCHECK* [11] berechnet. Die Parameter der Datensammlung, der Elementarzelle und die Verfeinerungsstatistiken sind in Tabelle 1 ersichtlich. Die Analyse der Temperaturfaktoren wurde mit *Moleman* [12] durchgeführt. Die Benennung der Protein-Aminosäuren erfolgte nach Bode et al. [13]. Koordinaten und Strukturfaktoren wurden in der Proteindatenbank mit dem folgenden Code hinterlegt: 4E7R.

Röntgenstruktur des Inhibitors MI432 in Thrombin zur Modellierung eines Matriptase-Komplexes

<i>PDB-Code</i>	4E7R
A. Data collection and processing	
Wavelength (Å)	1.54178
Space group	P1
<i>Unit cell parameters</i>	
a, b, c (Å)	50.2, 50.3, 71.8
α, β, γ (°)	97.6, 96.6, 90.6
Matthews coefficient (Å ³ /Da)	2.4
Solvent content (%)	49.3
Molecules in asymmetric unit	2
B. Diffraction data ^a	
Resolution range (Å)	30.00-2.25 (2.30-2.25)
Unique reflections	32 856
R (I) sym (%)	7.8 (34.8)
Completeness (%)	95.6 (93.1)
Redundancy	2.7 (2.6)
I/ σ (I)	13.4 (2.9)
C. Refinement	
Resolution range	29.3 – 2.25
<i>Reflections used in refinement</i>	
(work/free)	29.805 / 1.591
<i>Final R values for all reflections</i>	
(work/free) (%)	17.1 / 22.3
Protein residues (L chain/ H chain)	56 / 500
Sodium ions	4
Inhibitor atoms	80
Water molecules	265
<i>RMSD from ideality</i>	
Bond length (Å)	0.009
Bond angles(°)	1.2
<i>Ramachandran plot (PROCHECK)</i>	
Residues in most favoured regions (%)	85.9
Residues in additionally allowed regions (%)	13.8
Residues in generously allowed regions (%)	0.3
<i>Mean B-factor (Å²)</i>	
Protein (L + H chain)	30.3
Ligand	29.0
Sodium, phosphate, glycerol, NAG	66.5
Water molecules	32.2

^a Numbers in brackets are for the highest resolution shell.

Tabelle 6.1 Datensammlungs- und Verfeinerungsstatistiken für die Thrombin-Komplexstruktur

6.6 Referenzen

- [1] M. Hammami, E. Rühmann, E. Maurer, A. Heine, M. Gütschow, G. Klebe, T. Steinmetzer, New 3-amidinophenylalanine-derived inhibitors of matriptase, *Med. Chem. Commun.* 3 (2012) 807-813.
- [2] T. Steinmetzer, A. Schweinitz, A. Stürzebecher, D. Dönnecke, K. Uhland, O. Schuster, P. Steinmetzer, F. Müller, R. Friedrich, M.E. Than, W. Bode, J. Stürzebecher, Secondary amides of sulfonated 3-amidinophenylalanine. New potent and selective inhibitors of matriptase, *J. Med. Chem.* 49 (2006) 4116–4126.
- [3] Z. Otwinowski, W. Minor, [20] Processing of X-ray diffraction data collected in oscillation mode, in: *Methods in Enzymology*, vol. 276, *Macromolecular Crystallography Part A*, Elsevier, 1997, pp. 307–326.
- [4] E. Skordalakes, G.G. Dodson, D.S. Green, C.A. Goodwin, M.F. Scully, H.R. Hudson, V.V. Kakkar, J.J. Deadman, Inhibition of human alpha-thrombin by a phosphonate tripeptide proceeds via a metastable pentacoordinated phosphorus intermediate, *J. Mol. Biol.* 311 (2001) 549–555.
- [5] The CCP4 suite: programs for protein crystallography, *Acta Crystallogr. D Biol. Crystallogr.* 50 (1994) 760–763.
- [6] P.D. Adams, P.V. Afonine, G. Bunkóczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.-W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, P.H. Zwart, PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr. D Biol. Crystallogr.* 66 (2010) 213–221.
- [7] J. Painter, E.A. Merritt, Optimal description of a protein structure in terms of multiple groups undergoing TLS motion, *Acta Crystallogr. D Biol. Crystallogr.* 62 (2006) 439–450.
- [8] J. Painter, E.A. Merritt, TLSMD web server for the generation of multi-group TLS models, *J Appl Crystallogr* 39 (2006) 109–111.
- [9] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, *Acta Crystallogr. D Biol. Crystallogr.* 60 (2004) 2126–2132.
- [10] V.B. Chen, W.B. Arendall, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, D.C. Richardson, MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallogr. D Biol. Crystallogr.* 66 (2010) 12–21.
- [11] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: A program to check the stereochemical quality of protein structures, *J Appl Crystallogr* 26 (1993) 283–291.

- [12] G.J. Kleywegt, J.-Y. Zou, M. Kjeldgaard, T.A. Jones, Around O, in: M.G. Rossmann, E. Arnold, H. Fues, T. Hahn, H. Wondratschek, U. Müller, U. Shmueli, E. Prince, A. Authier, V. Kopský, D.B. Litvin, S. Hall, B. McMahon (Eds.), *International Tables for Crystallography*, F, *International Tables for Crystallography*, International Union of Crystallography, Chester, England, 2006, pp. 353–356.
- [13] W. Bode, I. Mayr, U. Baumann, R. Huber, S.R. Stone, J. Hofsteenge, The refined 1.9 Å crystal structure of human alpha-thrombin: interaction with D-Phe-Pro-Arg chloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion segment, *EMBO J.* 8 (1989) 3467–3475.

7 Zusammenfassung/Summary

7.1 Deutsche Fassung

Im zweiten Kapitel wurde gezeigt, dass die Anwendung von ITC-Verdrängungstitrationen und direkten ITC-Titrationen bei niedrigen c-Werten als eine zuverlässige Technik verwendet werden kann, um schwache Wechselwirkungen zu untersuchen, wie sie üblicherweise für die Bindung von Fragmenten an ein Zielprotein vorkommen.

Hierfür müssen jedoch einige Grundbedingungen erfüllt werden. Zuerst ist für ITC-Verdrängungstitrationen wichtig, dass das untersuchte Fragment mit einem Teil der Proteinbindetasche interagieren muss, der mit dem Bindungsbereich des verwendeten Referenzliganden überlappt. Zweitens muss der ausgewählte Referenzligand eine signifikant höhere oder niedrigere Bindungsenthalpie besitzen, so dass eine Wärmedifferenz zu dem untersuchten Fragment gemessen werden kann. Ferner ist zu beachten, dass bei jeder ITC-Verdrängungstitration, alle Fehler, die für die Referenzliganden ermittelt wurden, einen Einfluss auf die Genauigkeit der für das Fragment erhaltenen Parameter haben.

Schließlich kann, wenn kein geeigneter Referenzligand zur Verfügung steht und das Fragment enthalpisch genug bindet, eine direkte Titration bei niedrigen c-Werten als Alternative verwendet werden. In diesem Fall ist jedoch ein gewisses Vorwissen über die zu erwartende Bindungsaffinität des Fragments nötig, um die erforderliche Überschusskonzentration des Fragments am Ende der Titration abschätzen zu können.

Die Löslichkeit des Fragments und des Proteins sind ebenfalls wichtige Aspekte bei der Titration mit niedrigen c-Werten, da hohe Konzentrationen des Fragments in der Injektionsspritze erforderlich sind, um in der Probenzelle eine Konzentration des Fragments zu erreichen, die den K_D -Wert des Fragments überschreitet. Daher ist eine Messung des K_D -Werts des Fragments in einem unabhängigen Experiment sehr zu empfehlen, um die erforderliche Konzentration des Fragments abzuschätzen.

Letztlich ist es sehr wichtig, dass die unter den verschiedenen Versuchsbedingungen gemessenen Bindungsenthalpien nicht quantitativ verglichen werden sollten, sondern nur relativ zueinander durch Verwendung eines sorgfältig ausgewählten Messprotokolls.

Es ist bemerkenswert, dass mit verschiedenen Referenzliganden, die für die Verdrängung der Fragmente verwendet wurden, abweichende enthalpische Signale beobachtet wurden. Wahrscheinlich wird dies durch Unterschiede in der Wasserstruktur der verschiedenen Liganden in der Bindetasche verursacht. Die Betrachtung der relativen Enthalpiedifferenzen in den verschiedenen Verdrängungsexperimenten zeigte dagegen eine konsistente relative Differenz. Dies ermöglicht es, die untersuchten Fragmente relativ zueinander zu charakterisieren.

Die detaillierte Interpretation der enthalpischen Signaturen erfordert jedoch einen Vergleich der Kristallstrukturen der Fragmente. Dies wurde in Kapitel drei durchgeführt.

Dort wurden Röntgenkristallstrukturen von sechs Fragmenten (**1-6**), die die S1 Tasche von Thrombin adressieren und der Referenzliganden (**A** und **C**) bestimmt. Diese wurden in Bezug auf die im vorhergehenden Kapitel bestimmten thermodynamischen Bindungsprofile analysiert.

Die zwei eher schwach bindende Fragmente **5** und **6** zeigten nur wenige Wechselwirkungen zum Protein. Infolgedessen besitzen sie signifikante Restmobilität und ihre schwache Potenz machte eine detaillierte thermodynamische Charakterisierung unmöglich. Ein Chlor-Thiophen-Fragment (**2**), das in ähnlicher Weise in der S1-Tasche erkannt wird, zeigte zusätzliche starke Wechselwirkung oberhalb der S1-Tasche und ein umfangreiches Wassernetz deckte die Bindungsposition des Fragments ab. Dies machte dieses Fragment zum enthalpischsten Binder der Reihe und zu einem besseren Liganden im Vergleich zu den schwachen Chlor-Fragmenten wie m-Chlorbenzamid (**5**). Jedoch wird ein Teil des Enthalpievorteils dieses Fragments durch einen ungünstigen positiven Entropiebeitrag kompensiert.

Ein viertes Fragment (**1**), mit einer aromatischen Methoxy-P1-Kopfgruppe bindet mit geringer Affinität an das Protein jedoch wird auch dessen Bindung durch Enthalpie

dominiert. Zusätzlich zu seiner Bindung in der S1 Tasche, verfügt dieses Fragment über mehrere polare Kontakte mit dem Protein oberhalb der S1 Tasche.

Schließlich wurden die beiden am stärksten bindenden Fragmente mit Amidinogruppen, Benzamidin (**3**) und Amidinopiperidin (**4**), untersucht. Letzteres Fragment stellte sich als der beste Binder der Reihe heraus, und zeigte einen enthalpischen und entropischen Vorteil gegenüber Benzamidin. Beide Fragmente liegen unter den angewandten Messbedingungen geladen vor und bezahlen damit einen hohen Preis für ihre Desolvatisierung in Bezug auf die Enthalpie, der viel größer als für die anderen vier Fragmente sein sollte. Dieser Preis muss durch die starke Salzbrücke, die zu Asp189 gebildet wird, kompensiert werden. Darüber hinaus zeigen beide Fragmente ohne weitere polare Kontakte zum Protein ein perfektes Wassernetzwerk, welches die S1-Tasche abdeckt.

Bemerkenswert ist, dass die Fragmente, die ein ziemlich perfektes Solvatationsmuster im gebundenen Zustand zeigen auch die höchste Potenz aufweisen. Wahrscheinlich ist im Falle von Fragmentbindungen, der Einfluss von Qualität und Vollständigkeit des Wassernetzwerkes des gebundenen Fragments auf die thermodynamische Signatur der Ligandbindung größer, als bei Liganden mit Arzneistoffgröße.

In diesem Kapitel wurde auch die Seitenkette des Glu192 Restes, hinsichtlich ihrer großen konformativen Flexibilität in den Thrombin-Kristallstrukturen analysiert, wo er verschiedene Konfigurationen in den sechs Fragmentkomplexen aufweist. Teilweise ist dessen Carboxylatgruppe in Wechselwirkungen mit den Fragmenten eingebunden und kann die Solvatationsstruktur und somit auch die thermodynamische Signatur der Fragmente beeinflussen. In künftigen Arbeiten muss daher ein größeres Augenmerk auf diese Seitenkette für die Interpretation von Thrombin-Kristallstrukturen gelegt werden.

Die Untersuchung des Einflusses von anwesendem Hirudinfragment während ITC Messungen ergab, dass dies keinen Einfluss auf die thermodynamischen Profile der gemessenen Liganden hat und auch in Zukunft auf den Einsatz dieses teuren Hirudinfragments verzichtet werden kann.

Abschließend kann aus dieser Studie geschlossen werden, dass zumindest eine semiquantitative Abschätzung bezüglich der thermodynamischen Signatur von

Fragmenten möglich ist. Dafür ist eine genaue strukturelle Untersuchung der beteiligten Aminosäureseitenketten, insbesondere in Bezug auf deren Solvatationsstruktur, nötig. Nur dann macht es Sinn ein Fragment als enthalpischen oder entropischen Startpunkt für die nachfolgende Leitstrukturoptimierung zu verwenden. Bereits kleine Veränderungen der Solvatatstruktur solcher schwachen Binder können leicht die thermodynamische Signatur von "stärker enthalpisch" zu "stärker entropisch" verschieben.

Aus dem vierten Kapitel können ebenfalls einige wertvolle Schlüsse gezogen werden. Die Bindungsaffinität eines Liganden wird nicht nur durch die Eigenschaften des gebildeten Protein-Ligand-Komplexes bestimmt, sondern auch durch die Unterschiede von Liganden in wässriger Lösung vor Proteinbindung. Konformativ eingeschränkte Liganden können einen signifikanten Bindungsvorteil gegenüber flexiblen Liganden haben, wenn in einem rigiden Ligandgrundgerüst eine Geometrie bereits präorganisiert ist, die dem Liganden in gebundenem Zustand ähnlich ist.

Es ist auch bemerkenswert, dass die Methylierung eines polaren Amidstickstoffs in Ligand E nicht zwangsläufig zu einem Abfall der Affinität führt, obwohl die freie NH-Gruppe des entsprechenden Liganden in einer H-Bindung an das Protein beteiligt war. Diese Interaktion wurde durch ein Wassermolekül zu Tyr60AOH vermittelt, das im Fall des N-Me-Derivats geringfügig an eine neue Position verschoben wurde, wo es in eine Art Wassergürtel integriert ist, der in der Folge den Liganden in seiner Bindung fixieren kann.

Offensichtlich hat eine neu gebildete Solvatationshülle, die oberhalb des Protein-Komplexes mit seinem gebundenen Liganden angeordnet ist, einen Einfluss auf die Bindung dieses Liganden. Die Abdeckung von Liganden durch ein gut etabliertes Wassernetzwerk reduziert offensichtlich die Restbeweglichkeit des gebundenen Liganden, wie durch die B-Faktor-Analyse gezeigt wurde.

Insgesamt legt die verwendete Ligand-Serie nahe, dass selbst kleine Veränderungen, wie die Addition oder Verschiebung von Methylgruppen an einem ansonsten identischen Grundgerüst bemerkenswerten Einfluss auf die Affinität und die Aufspaltung von Enthalpie/ Entropie der Ligandbindung haben.

Auch in Kapitel fünf wurden Untersuchungen durchgeführt deren Ergebnisse den großen Einfluss von Präorganisationseffekten auf die Bindungsaffinität und Thermodynamik der Ligand Bindung nahelegen.

Die hohe Affinität des Aminomethylen-Inhibitors MI001 resultiert wahrscheinlich aus optimalen intra- und intermolekularen Wasserstoffbrücken im zentralen Teil des Komplexes. Sollte diese Konformation des Inhibitors derjenigen des Inhibitors in Lösung entsprechen, könnte dieser nur bei diesem Inhibitor mögliche Präorganisationseffekt eine Erklärung für die hohe Bindungsaffinität liefern.

Diese Vermutung wurde auch durch die Enzymassays einer zweiten Serie gestützt, wo die Unterschiede in der Bindungsaffinität ohne den P4 Substituenten, der für einen Teil der intramolekularen Fixierung essentiell ist, deutlich geringer ausfallen.

Auch für die Liganden dieser zweiten Serie sollten noch hochaufgelöste Röntgenstrukturen und thermodynamische Daten gesammelt werden, um die Ergebnisse zu bestätigen. Die Struktur von MI002 sollte wegen ihrer zweifelhaften zweiten Konformation neu erstellt oder neu verfeinert werden.

In Kapitel sechs wurde ein neuer vom 3-Amidinophenylalanin abgeleiteter Matriptase Inhibitor in Thrombin kristallisiert und die Röntgenkristallstruktur eines repräsentativen, verwandten Inhibortyps in Matriptase mit dieser überlagert.

Aus der Überlagerung dieser Strukturen konnten Informationen hinsichtlich der in dieser Studie beobachteten Selektivitätsprofile zu anderen Serinproteasen gezogen werden. Durch die große Homologie der mit dem Inhibitor interagierenden Aminosäuren in der Bindetasche von Thrombin und Matriptase können zudem zuverlässige Vorhersagen über den möglichen Bindungsmodus in Matriptase getroffen werden.

Es wurde darüber hinaus beobachtet, dass die Verwendung von sterisch anspruchsvollen Inhibitoren zu einer unerwarteten Veränderung der Raumgruppe eines Proteinkristalls durch das Eindiffundieren in den Zwischenräumen der Kristallpackung zur Bindetasche führen kann. Während des Soakings von anderen sterisch anspruchsvollen Matriptaseinhibitoren kam es sogar zu einer Zerstörung der Proteinkristalle.

7.2 English Version

In chapter two we showed that the application of ITC displacement titrations and direct ITC titrations at low c -value can be used as a reliable technique to study low-affinity interactions as usually given for the binding of a fragment to a target protein.

Important aspects can be concluded from this study. First of all, for ITC displacement titrations the fragment must bind to a part of the protein binding pocket that overlaps with the binding pose region of the applied reference ligand. Secondly, the selected reference ligand must have a significant higher or lower binding enthalpy so that a heat difference signal can be recorded for the fragment.

Furthermore, it has to be regarded that, as for any displacement titration, all errors affecting the parameter determinations of the reference ligand will add to the accuracy of the parameters obtained for the fragment. Finally, if no suitable reference ligand is available and the fragment binds enthalpically enough, low c -value titrations can be applied as an alternative. In this case, however, some anticipated knowledge about the expected binding affinity of the fragment must be available to estimate the required excess concentration of the fragment at the end of the titration.

The solubility of the fragment and the protein are the most crucial issues in low c -value titrations because the required high concentrations of the fragment to be studied in the injection syringe should exceed the final sample cell concentrations with respect to the fragment's K_D value. In order to estimate the required concentration of the fragment, a crude estimate of the fragment's K_D obtained from an independent experiment is highly recommended.

It is very important that measured binding enthalpies should not be compared quantitatively across different experimental conditions, but only relative to each other by using one suitable measurement protocol. Remarkably, different reference ligands used for the displacement of pre-incubated fragments reveal deviating enthalpic signals. Most likely these differences are caused by differences in the water structure and thus in the residual

solvation pattern at the binding site of the different ligands used for titration. The consideration of the relative enthalpy differences across the different displacement experiments showed a consistent relative difference. This allows to characterize the studied fragments relative to each other.

The detailed interpretation of the enthalpic signature requires a comparison of the corresponding crystal structures in order to select fragments as superior enthalpy-dominated candidates for further development. This has been performed in chapter three.

To achieve this goal, high-resolution crystal structures of six fragments binding to the S1 pocket of thrombin were determined and analyzed with respect to the thermodynamic binding profiles observed for the bound fragments determined in the previous chapter. Two rather weak binding fragments (5 and 6) do not succeed to establish strong interactions with the enzyme. In consequence, they show significant residual mobility and their weak potency made a detailed thermodynamic characterization impossible. A chlorothiophenamide fragment (2) which is similarly recognized in the S1 pocket shows additional strong interactions beyond the S1 pocket and an extensive water network capping the fragment's binding position turn this fragment into the most enthalpic binder of the series and makes it a much better binder compared to the weak fragments such as the analog m-chlorobenzamide. However, a part of the enthalpic advantage of this fragment is compensated by an unfavorable positive entropic contribution. A fourth fragment (1) exhibiting a methoxy aromatic P1 head group binds with less affinity to the protein, however its binding is also dominated by enthalpy. In addition to its S1 binding, this fragment experiences several polar contacts with the protein beyond the S1 pocket.

Finally, the two most potent fragments studied are of the amidino-type molecules benzamidine (3) and amidinopiperidine (4). The latter fragment turns out to be the best binder of the series, showing an enthalpic and entropic advantage over benzamidine. Both fragments are charged under used conditions and therefore a high enthalpic desolvation price has to be paid which should be much larger than for the other four fragments. This cost has to be overcome by the strong salt bridge formed to Asp189. In addition, both

fragments allow for a perfect water network capping the S1 pocket even though they are not involved in any further polar contacts.

Remarkably, the fragments exhibiting a rather perfect solvation pattern in the adopted binding pose also experience the highest potency. Likely in the case of fragment binding the influence of the quality and completeness of the water network wrapping around the bound fragment is even more determining for the thermodynamic signature than in case of drug-size ligand binding.

In this chapter also the side chain of Glu192 which exhibits large conformational flexibility across the thrombin crystal structures where it adopts different configurations in the six fragment complexes has been analyzed. Partially its carboxylate group is involved in interactions with the fragments and this residue can influence the residual solvation structure taking an impact on the thermodynamic signature of the fragments. In future work there has to be a focus on this residue for the interpretation of thrombin crystal structures.

The investigation of the influence of present hirudin during ITC measurements revealed that there is no effect on the thermodynamic profile of the measured ligand and can therefore be neglected in the future.

Overall for this study it can be concluded that at least a semi-quantitative estimate on the thermodynamic signature of the fragments is possible. Therefore, a detailed control of the involved structural changes is mandatory, particular with respect to the residual solvation structure. Only then the assignment of a fragment to be a more enthalpic or entropic starting point for subsequent lead optimization makes sense. A modulation of the solvation structure for such weak binders can easily shift the thermodynamic signature from “more enthalpic” to “more entropic”.

From the fourth chapter also several messages can be learnt. Binding affinity is not only determined by the properties of the formed protein-ligand complex but some differences can already discriminate ligands in aqueous solution prior to any protein binding.

Conformationally restricted ligands can experience a significant binding advantage over more flexible ligands if the correct protein-bound geometry is well preorganized in a rigidified skeleton.

It is also remarkable that the methylation of a polar amide nitrogen does not necessarily result in a drop of affinity, even though the free NH group of the corresponding ligand E has been involved in an H-bond to the protein. An interstitial water molecule used to mediate this hydrogen bond to Tyr60A_{OH} is slightly shifted in the N-Me derivative to a new position where it can establish a kind of interaction belt wrapping around and in consequence fixing the ligand in its binding pose.

Obviously the newly formed solvation layer arranged across the protein complex with its bound ligand takes an impact on binding. The coverage of ligands by a well-established water layer obviously reduces the residual mobility of the bound ligand as indicated by the B-factor analysis.

Overall, the used ligand series suggests that even small changes involving the addition or shifting of methyl groups at a given parent scaffold can take remarkable influence on the affinity and even more on the enthalpy/entropy partitioning in ligand binding.

The results of the studies conducted in chapter five also suggest, the major impact of preorganization effects on the binding affinity and thermodynamics of ligand binding.

The high affinity of the amino methylene inhibitor likely results from optimal intra- and intermolecular hydrogen bonds in the central part of the complex. If this conformation of the inhibitor corresponds to that of the inhibitor in solution, these potential preorganization effects, observed only for this inhibitor, could provide a possible explanation for the high binding affinity.

This assumption is supported by the measured significantly lower binding affinities of the second series, lacking the P4 substituent, which is essential for some of the intramolecular fixation.

Also for the ligands of this second series high-resolution X-ray structures and thermodynamic data should be collected to confirm the results.

In chapter six, a new derivative of a 3-amidinophenylalanine matriptase inhibitor (MI432) was crystallized with thrombin and the resulting complex was superimposed with a representative X-ray crystal structure of matriptase bound to a related inhibitor.

From the superimposition of these structures information could be extracted regarding the selectivity profiles towards serine proteases observed in this study.

Due to the large homology of amino acids in the binding pocket of thrombin and matriptase, reliable predictions about the possible binding mode of the inhibitor, can be made for matriptase.

It was also observed that the use of bulky inhibitors may lead to an unexpected change in the space group of a protein crystal by diffusing them into the unoccupied interstices of the crystal packing to the binding pocket.

During the soaking of other bulky matriptase inhibitors even a decomposition of the protein crystals has been observed.

8 Anhang/Annex

8.1 Abkürzungen/Abbreviations

°C	Grad Celsius
ACB	2-(aminomethyl) -5-chlorobenzylamide
Ala	Alanin
AMBA	4-amidinobenzylamide
AMC	Aminomethylcoumarin
Arg	Arginin
ASA	solvent accessible surface area
Asn	Asparagin
Asp	Asparaginsäure
B-factor	Temperaturfactor/Debye–Waller factor
Bzls	Benzylsulfonyl
Cha	Cyclohexylalanine
Cys	Cystein
DMSO	Dimethylsulfoxid
D _{sat}	Degree of saturation
FBDD	Fragment-based drug discovery
FBLD	Fragment-based lead discovery
F _c	Calculated structure factor amplitude
F _o	Observed structure factor amplitude
G ⁰	Freie Gibbs Energie
Gln	Glutamin
Glu	Glutaminsäure
Gly	Glycin
H ⁰	Enthalpie
H-bond	Hydrogen bond
HCB	2-Hydroxymethyl-5-chlorobenzylamid
Hepes	2-(4-(2-Hydroxyethyl) - 1-piperazinyl) -ethansulfonsäure
His	Histidin
HTS	High-throughput screening
Ile	Isoleucin
ion.	Ionisation
ITC	Isothermal titration calorimetry
J	Joule
k	Kilo
K	Kelvin
K _a	Assoziationskonstante
K _D	Dissoziationskonstante

K _i	Inhibitionskonstante
Leu	Leucin
Lys	Lysin
MD	Molecular dynamics
NMR	Nuclear magnetic resonance
PDB	Protein data bank
PEG	Polyethylene glycol
Phe	Phenylalanin
Pro	Prolin
R	Ideale Gaskonstante
rmsd	Root mean square deviation
S ⁰	Entropie
Ser	Serin
SPR	Surface plasmon resonance
STD	Saturation Transfer Difference
T	Absolute Temperatur
Thr	Threonin
THR	Thrombin
TLS	Translation/libration/screw
Tos	Tosyl
tPA	Tissue plasminogen activator
Tricine	N-(2-hydroxy-1,1-bis(hydroxymethyl)ethyl) glycine
Tris	Tris(hydroxymethyl)-aminomethan
Trp	Tryptophan
TSA	Thermal Shift Assay
TSPP	Tetrasodium pyrophosphate
Tyr	Tyrosin
Val	Valin
VSA	van der Waals surface area
σ	Sigma

8.2 Beiträge auf Konferenzen/Conference contributions

Poster:

Thermodynamic signatures of fragment binding: Validation of direct versus displacement ITC titrations.

Rühmann E., Betz M., Fricke M., Heine A., Schäfer M., Klebe G.

Developments in Protein Interaction Analysis, San Diego, CA, USA, 2014

8.3 Danksagung/Acknowledgments

Diese Dissertation konnte nur mit Hilfe und Unterstützung zahlreicher Menschen entstehen. Besonders danken möchte ich:

Herrn Prof. Dr. Gerhard Klebe dafür diese Dissertation in seiner Arbeitsgruppe anfertigen zu können. Für seine kontinuierliche Unterstützung und stete Gesprächs- und Diskussionsbereitschaft. Seine Zuversicht und die gewährten Freiheiten haben eine sehr konstruktive und gute Arbeitsatmosphäre ermöglicht.

Herrn Prof. Dr. Andreas Heine für seine ausgezeichnete Einführung in die Proteinkristallographie, für das Korrekturlesen dieser Arbeit, für seine Hilfestellung bei allen möglichen kristallographischen Problemen und die Bereitschaft diese Arbeit als Zweitgutachter zu bewerten.

Herrn Prof. Dr. Steinmetzer für seine Ratschläge zur Reaktionskinetik und seiner Arbeitsgruppe für die Synthese der in Kapitel 5 und 6 untersuchten Thrombininhibitoren.

Dr. Martina Schäfer und *Marie Fricke* von Bayer Health Care danke ich für die Kollaboration, die den Anstoß für weitere Untersuchungen an Thrombin Fragmenten gegeben hat.

Herrn Prof. Dr. David Hangauer und seinem Doktoranden *Maan Khayat* für die Synthese einiger Inhibitoren.

Frau *Lydia Hartleben* für die tolle und zuverlässige Organisation der bürokratischen Arbeitsabläufe.

Herrn Hans-Dieter Gerber danke für die kompetente Beantwortung zahlreicher Fragen zur Ligand Synthese und Analytik.

Herrn Christian Sohn für seine Hilfe bei allen Arbeiten an den Röntgengeräten und vor allem für seine Unterstützung bei der photometrisch kinetischen Untersuchung der Thrombininhibitoren.

Florian Braun für eine zügige Nachsynthese einer dringend benötigten Verbindung.

Melinda Rupp für die Durchführung der MD-Simulationen in Kapitel vier.

Dr. Adam Biela für die umfangreiche und geduldige Einführung in die verschiedenen Arbeitstechniken.

Dr. Chris Rechlin für viele Diskussionen rund um das ITC-Gerät und dessen Reparaturen.

Der gesamten AG Klebe für die tolle Arbeitsatmosphäre, insbesondere *Frederik Ehrmann*, *Phong Nguyen*, *Martin Stieler* und meinen Bürokollegen *Stefan Krimmer*, *Michael Betz* und *Christof Siefker*.

Dem „*Elektro-Betz-Team*“ für eine zuverlässige Administration der IT-Systeme.

Dem European Research Council für die Finanzierung dieser Arbeit.

Dem Personal des BESSY und DESY für die freundliche und stets hilfreiche Unterstützung bei den kristallographischen Experimenten an dem jeweiligen Synchrotron.

Meiner Familie, insbesondere meinen Eltern *Dr. Eggert Rühmann* und *Rita Rühmann*, danke ich für die immer vorhandene Unterstützung während des Studiums und der Promotion.

Besonders bedanken möchte ich mich bei *Nedyalka Radeva* für ihr Verständnis, Motivation und ihre Unterstützung in jeder Hinsicht, während der gesamten gemeinsamen Zeit.

8.4 Lebenslauf/Curriculum Vitae

Aus Gründen des Persönlichkeitsschutzes wird von der elektronischen Veröffentlichung des Lebenslaufes abgesehen.

E R K L Ä R U N G (gem. § 10 Promotionsordnung)

Ich versichere, dass ich meine Dissertation

„Strukturelle und thermodynamische Charakterisierung der Fragment und Liganden
Bindung an Thrombin“

selbständig ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen bedient habe. Alle vollständig oder sinngemäß übernommenen Zitate sind als solche gekennzeichnet.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den 03.11.2015

.....

Eggert Rühmann