



**Fragmentbasiertes Design von p97-Liganden:
Identifizierung von Startstrukturen zur Entwicklung von
Protein-Protein-Interaktionsinhibitoren für die
SHP-Bindestelle der AAA+ ATPase p97**

*Fragment-based design of p97-ligands:
Identification of starting points for the development of
protein-protein-interaction inhibitors targeting the
SHP-binding site of the AAA+ ATPase p97*

Dissertation zur Erlangung des naturwissenschaftlichen Doktorgrades
der Graduate School of Life Sciences,
Julius-Maximilians-Universität Würzburg,
Klasse Biomedizin

Vorgelegt von
Sebastian Helmut Bothe

aus
Mannheim
Würzburg **2021**

Eingereicht am:

Mitglieder des Promotionskomitees:

Vorsitzender: Prof. Dr. Thomas Dandekar

1. Betreuer: Prof. Dr. Christoph Sotriffer

2. Betreuer: Prof. Dr. Hermann Schindelin

3. Betreuer: Prof. Dr. Alexander Buchberger

Tag des Promotionskolloquiums:

Doktorurkunden ausgehändigt am:

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, die Dissertation „Fragmentbasiertes Design von p97-Liganden: Identifizierung von Startstrukturen zur Entwicklung von Protein-Protein-Interaktionsinhibitoren für die SHP-Bindestelle der AAA+ ATPase p97“, eigenständig, d.h. insbesondere selbständig und ohne Hilfe eines kommerziellen Promotionsberaters, angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

Würzburg, 12.2.2021

Unterschrift

Affidavit

I hereby confirm that my thesis entitled „Fragment-based design of p97-ligands: Identification of starting points for the development of protein-protein-interaction inhibitors targeting the SHP-binding site of the AAA+ ATPase p97“, is the result of my own work. I did not receive any help or support from commercial consultants. All sources and / or materials applied are listed and specified in the thesis.

Furthermore, I confirm that this thesis has not yet been submitted as part of another examination process neither in identical nor in similar form.

Würzburg, 12.2.2021

Signature

*„Die Wirklichkeit, von der wir sprechen können, ist nie die Wirklichkeit „an sich“,
sondern eine gewusste Wirklichkeit oder sogar in vielen Fällen
eine von uns gestaltete Wirklichkeit.“*

Werner Heisenberg

Abstract

The AAA+ ATPase p97 is an essential protein involved in numerous cellular processes and plays a key role in multiple aspects of protein homeostasis. Its functional diversity is mediated through the interaction with a large number of distinct cofactors binding to the N-domain of p97. Due to its significant role in regulating a variety of physiological responses, p97 has emerged as a potential therapeutic target. A small molecule inhibiting the cofactor binding would be important to dissect the molecular and cellular functions of p97 cofactors, thus helping to unravel their specific role in controlling p97 activity. Such compounds may also open routes to new cancer therapies.

In this work, a fragment-based approach was pursued for the identification of chemical starting points for the development of a protein-protein interaction inhibitor addressing the SHP binding site. Therefore, computer-assisted methods, such as virtual screenings and molecular dynamics simulations, as well as biophysical techniques including bilayer interferometry, X-ray crystallography, and ligand-based NMR techniques, were applied.

The computer-based design started with an analysis of the known p97 crystal structures in complex with the SHP motifs of cofactors UFD1 and Derlin-1. In addition, molecular dynamics simulations were used to analyze the water properties within the SHP binding site. Based on these results, pharmacophore models were developed and utilized in the subsequent virtual screening and docking process. With the help of molecular dynamics simulations, ten compounds were selected for experimental validation. Two of these were confirmed as ligands in STD-NMR and bilayer interferometry experiments.

In parallel, a biophysical fragment screening of over 650 compounds was performed using the bilayer interferometry method. This led to the identification of 22 compounds binding to the N-domain. Fifteen of these fragments were confirmed in an orthogonal STD-NMR assay. Five compounds showed affinities with K_D values below 500 μM and favourable ligand efficiencies for further optimization. Furthermore, the binding kinetics and affinity of xanthohumol, a natural product reported in the literature as a p97 inhibitor, were determined and binding to the N-domain was confirmed.

To identify possible binding sites of these five fragments, mixed solvent molecular dynamics simulations were performed. These revealed that all compounds address the SHP binding site in the N-domain. The regions coincide with hot spots of the cofactor binding and, thus, represent potential anchor points for a protein-protein interaction inhibitor. For two fragments, the postulated binding sites were confirmed by X-ray crystallography and STD-NMR measurements on p97 alanine mutants, respectively. The X-ray structure obtained is the first p97 structure showing a fragment bound to the N-domain.

Zusammenfassung

Die AAA+ ATPase p97 ist ein essenzielles Protein, das an einer Vielzahl zellulärer Prozesse beteiligt ist und eine Schlüsselrolle in der Protein-Homöostase spielt. Die funktionale Diversität von p97 beruht auf der Interaktion zahlreicher unterschiedlicher Kofaktoren, die vorwiegend an die N-Domäne von p97 binden. Aufgrund seiner Bedeutung in der Regulierung diverser physiologischer und pathologischer Prozesse stellt p97 eine interessante Zielstruktur für die Entwicklung neuer Wirkstoffe dar, die insbesondere in der Krebstherapie von Bedeutung sein könnte. Bekannte p97-Inhibitoren greifen vor allem die ATPase-Funktion des Proteins an. Ein neuer pharmakologischer Ansatz stellt die Inhibierung der Kofaktorbindung an die N-Domäne dar. Ein solcher Protein-Protein-Interaktionsinhibitor wäre nicht nur von therapeutischem Interesse, sondern hätte auch einen besonderen Nutzen für die Entschlüsselung molekularer und zellulärer Funktionen von p97-Kofaktoren.

In dieser Arbeit wurde ein fragmentbasierter Ansatz für die Identifizierung von chemischen Startstrukturen für die Entwicklung eines Protein-Protein-Interaktionsinhibitors verfolgt. Als Zielstruktur wurde die SHP-Bindestelle in der N-Domäne gewählt. Die Identifizierung von Liganden erfolgte sowohl durch computergestützte Methoden (insbesondere virtuelles Screening und Molekulardynamik-Simulationen) als auch experimentell durch biophysikalische Techniken (wie Biolayer-Interferometrie, Röntgenstrukturanalyse und ligandbasierte NMR-Techniken).

Die Grundlage des computerbasierten Designs stellte eine Analyse der bekannten Kristallstrukturen der p97-Komplexe mit den SHP-Motiven der Kofaktoren UFD1 und Derlin-1 dar. Darüber hinaus dienten Molekulardynamik-Simulationen der Analyse der Wassereigenschaften innerhalb der SHP-Bindestelle. Darauf aufbauend wurden verschiedene Pharmakophormodelle entwickelt, die die Grundlage des im Anschluss durchgeführten virtuellen Screenings und Dockings bildeten. Anhand der Ergebnisse von Molekulardynamik-Simulationen wurden zehn Verbindungen für die experimentelle Validierung ausgewählt. Hiervon konnten zwei Fragmente in STD-NMR- und Biolayer-Interferometrie-Experimenten als Liganden bestätigt werden.

In einem parallel durchgeführten biophysikalischen Fragmentenscreening mittels Biolayer-Interferometrie wurden unter mehr als 650 Verbindungen 22 identifiziert, die an die N-Domäne binden. 15 dieser Fragmente wurden durch einen orthogonalen STD-NMR-Assay bestätigt. Fünf dieser Verbindungen zeigten Affinitäten mit K_D -Werten kleiner 500 μM und günstigen Ligandeffizienzen. Des Weiteren konnte die Bindungskinetik und Affinität des in der Literatur als p97-Inhibitor berichteten Naturstoffes Xanthohumol bestimmt und eine Bindung an die N-Domäne bestätigt werden.

Zur Identifizierung möglicher Bindestellen dieser fünf Fragmente wurden *mixed-solvent* Molekulardynamik-Simulationen durchgeführt. Diese ergaben, dass alle Verbindungen die SHP-Bindestelle in der N-Domäne adressieren. Die Regionen fielen mit *hot spots* der Kofaktorwechselwirkungen zusammen und stellen somit mögliche Ankerpunkte für die Weiterentwicklung dar. Für zwei Fragmente konnten die postulierten Bindestellen mittels Röntgenstrukturanalyse bzw. STD-NMR-Messungen an p97-Alanin-Mutanten bestätigt werden. Die erhaltene Röntgenstruktur ist die erste p97-Struktur, die ein gebundenes Fragment an der N-Domäne zeigt.

Danksagung

Für das Gelingen dieser Arbeit sei **allen** Beteiligten herzlich gedankt.

Ich danke Prof. Dr. Christoph Sotriffer, Prof. Dr. Hermann Schindelin und Prof. Dr. Alexander Buchberger für die Aufnahme in das GRK 2243 und die Möglichkeit mein Wissen über die unterschiedlichen Disziplinen der Wirkstoffentwicklung und des Ubiquitin-Systems zu vertiefen. Des Weiteren bedanke ich mich für die hervorragende Betreuung während dieser Zeit, mit nötigen Hilfestellungen, Diskussionen und Denkanstößen, aber auch der großen Freiheit die mir bei der Arbeit an diesem Projekt gewährt wurde.

Dr. Petra Hänzelmann für ihre Einführung in biochemische Methoden, sowie für ihre stetige Hilfe, bei allen biochemischen, strukturellen und p97 relevanten Fragestellungen, die maßgeblich zum Gelingen dieser Arbeit beigetragen haben.

Prof. Dr. Ruth Brenk für die Möglichkeit und Unterstützung der Durchführung eines Fragmentscreenings in ihrem Arbeitskreis an der Universität Bergen (Norwegen), sowie für die herzliche Aufnahme während des dortigen Forschungsaufenthalts.

Prof. Dr. Thomas Dandekar für die Übernahme des Vorsitzes meiner Promotion.

Markus Zehe für die zahlreichen und langen Diskussionen zu dem Themenfeld des fragmentbasierten Wirkstoffdesigns und der Hilfestellung bei den STD-NMR Experimenten.

Stephan Böhler, der mich fast das ganze Projekt als Praktikant, Masterrand, bis zum Doktorand in diesem Projekt begleitet und unterstützt hat.

Darüber hinaus gilt mein Dank :

- AD Dr. Curd Schollmayer für die wissenschaftliche und technische Unterstützung bei den NMR-Messungen, sowie Prof. Dr. Ulrike Holzgrabe für die Bereitstellung des NMR-Spektrometers.
- Der TA Monika Kuhn für ihre große Hilfe bei der Expression und Aufreinigung von p97 Konstrukten, sowie der TA Khanh Kim Dao (Universität Bergen) für ihre Unterstützung während meines Aufenthaltes in Norwegen.

- Josef Kehrein für die Unterstützung bei der Durchführung der Simulationen mit GROMACS.
- Dr. Hendrik Wünsche (Sartorius) für die Ratschläge und Organisation eines K2 Octet Biolayer-Interferometrie Gerätes.
- Dr. Andrea Thorn und Dr. Florian Sauer für die Hilfestellung zu kristallographischen Fragen.
- Meinem Praktikant Thomas Bretschneider für das Zuarbeiten während des virtuellen Screenings.
- Matthias Diebold für die zahlreichen Erörterungen bezüglich computerbasierter Verfahren und für die hervorragende Systemadministration im Arbeitskreis von Prof. Sotriffer.
- Carolina Galgenmüller für die Zusammenarbeit am gemeinsamen Zielprotein p97
- Meinen Kollegen im Arbeitskreis Sotriffer Erik Endres, Natalia Yuan Chen, Gerald Keller, Carina Herbst und Maximilian Kuhn für den wissenschaftlichen Austausch und die schöne gemeinsame Zeit, sowie allen Kollegen in der Strukturbiologie am Rudolf-Virchow-Zentrum.

Zuletzt gilt mein besonderer Dank meiner Familie, meinen Eltern und Schwiegereltern für den großen Rückhalt während dieser Arbeit. Und hier im ganz besonderem Maße Laura, für ihre große Unterstützung, Geduld und Ausdauer, ohne die diese Arbeit nicht hätte gelingen können. Und nicht zuletzt Jona und Leander für ihre zahlreichen Aufmunterungen.

Inhaltsverzeichnis

Eidesstattliche Erklärung	iii
Abstract	ix
Zusammenfassung	xi
Danksagung	xiii
1 Einleitung	1
2 Theorie	3
2.1 Die AAA+ ATPase p97	3
2.1.1 Struktur und Aufbau von p97	4
2.1.2 Biologische Funktionen von p97 und Verknüpfung mit dem Ubiquitin-Proteasom-System	7
2.1.3 Interaktionen von p97 mit Kofaktoren	9
2.1.3.1 UBX- und UBXL-Domänen	12
2.1.3.2 Das VIM- und VBM-Motiv	12
2.1.3.3 Das SHP-Motiv	13
2.1.4 p97 als pharmakologischer Angriffspunkt (<i>Target</i>)	13
2.1.4.1 Erkrankungen, die mit p97-Mutationen in Verbindung stehen	13
2.1.4.2 Erkrankungen, die mit nativem p97 in Verbindung stehen	14
2.1.4.3 p97 und virale Infektionen	15
2.1.5 Derzeit bekannte p97-Inhibitoren	15
2.1.5.1 ATPase-Inhibitoren	15
2.1.5.2 Allosterische Inhibitoren	16
2.1.5.3 Inhibitoren mit unbekanntem Wirkmechanismus	17
2.1.5.4 Fragmentscreening an p97-N und p97-ND1	17
2.1.5.5 Das Chalkon Xanthohumol als Inhibitor von p97	18
2.1.6 Protein-Protein-Interaktionen	24
2.1.7 Die N-Domäne als mögliches Ziel für p97-Inhibitoren	25

2.2	Computergestützte Methoden in der Wirkstoffentwicklung	27
2.2.1	Molekulare Deskriptoren zur Charakterisierung von Molekülbibliotheken	27
2.2.2	Pharmakophormodelle	29
2.2.3	Molekulares Docking	30
2.2.3.1	Generierung von Bindemodi des Liganden	32
2.2.3.2	Scoringfunktionen	33
2.2.4	Molekulardynamik-Simulationen	35
2.2.4.1	Molekulare Kraftfelder	36
2.2.4.2	Theorie der MD-Simulation	39
2.3	Fragmentbasiertes Wirkstoffdesign (<i>Fragment based drug design</i> , FBDD)	43
2.3.1	Idee des FBDD	43
2.3.2	Aufbau und Ablauf eines FBDD	44
2.3.3	Fragmentbibliotheken	44
2.3.4	Biophysikalische Methoden im FBDD	47
2.3.5	Auswahl geeigneter Fragmente über Ligandeffizienz-Indizes	48
2.3.6	Methoden zur Optimierung der Fragmente	51
2.3.7	Computerbasierte Methoden im fragmentbasierten Kontext	52
2.4	Theoretische Hintergründe zu den eingesetzten biophysikalischen Methoden	54
2.4.1	Kernspinresonanz-Spektroskopie	54
2.4.1.1	Grundlagen	55
2.4.1.2	T_1 und T_2 -Relaxationszeit	56
2.4.1.3	Spindiffusion und Sättigung	56
2.4.1.4	Das Sättigungstransfer-Differenz (STD)-NMR- Experiment	57
2.4.2	Röntgenstrukturanalyse	58
2.4.2.1	Theoretische Grundlagen	58
2.4.2.2	Qualitätsbeurteilung von Datensätzen	62
2.4.3	Biolayer-Interferometrie (BLI)	63
2.4.3.1	Theoretischer Hintergrund der BLI-Technik	64
2.4.3.2	Gleichgewichts- und kinetische Analyse von Protein-Ligand-Interaktionen im Kontext der BLI .	68
2.4.3.3	1:1-Bindungsmodell	68
2.4.3.4	Komplexere Bindungsmodelle	70
3	Zielsetzung dieser Arbeit	73

4	Materialien, Geräte und Software	75
4.1	Software	75
4.2	Verwendete Puffer	76
4.3	Geräte	77
4.4	Chemikalien	78
4.5	Spezielle Verbrauchsmaterialien	80
5	Methoden	81
5.0.1	Abbildungen	81
5.1	Computerbasierte Methoden	81
5.1.1	Virtuelle Fragmentbibliothek	81
5.1.1.1	Auswahl der Fragmente für die Fragmentbibliothek	81
5.1.2	Aufbereitung der zusammengestellten Fragmentbibliothek	83
5.1.2.1	Charakterisierung der Fragmentbibliotheken . . .	83
5.1.2.2	Hauptkomponentenanalysen	83
5.1.3	Strukturanalysen der SHP-Bindestelle	84
5.1.3.1	Grundlegende Einstellungen in MOE	84
5.1.3.2	Strukturanalysen durch visuelle Inspektion in MOE	84
5.1.3.3	Interaktionpotential-Analysen in MOE	85
5.1.3.4	<i>Hot-spot</i> -Analyse unter Verwendung des FTMap-Servers	85
5.1.4	Pharmakophorbasiertes Screening	85
5.1.4.1	Erstellung der Pharmakophormodelle	85
5.1.4.2	Konformerbibliotheken und pharmakophorbasiertes Screening	85
5.1.5	Virtuelles Screening	86
5.1.5.1	Vorbereitung	87
5.1.5.2	Erster Dockinglauf	87
5.1.5.3	Auswahl der Fragmente für den zweiten Dockinglauf	88
5.1.5.4	Zweiter Dockinglauf	89
5.1.5.5	Rescoring der Bindeposen des zweiten Dockinglaufs	89
5.1.5.6	Auswahl der Fragmente für MD-Simulationen . .	90
5.1.6	MD-Simulationen	90
5.1.6.1	Präparation der verwendeten Liganden und Proteinstrukturen, sowie Erstellung der Startkonformation der Simulationen	90

5.1.6.2	Generelle Einstellungen aller AMBER MD-Simulationen	91
5.1.6.3	Äquilibrierung der Systeme	91
5.1.6.4	Auswertung mittels der <i>root mean square deviation</i>	92
5.1.6.5	MD-Simulationen des virtuellen Screenings	92
5.1.6.6	MD-Simulation zur Abschätzung der Wassereigenschaften der SHP-Bindestelle	93
5.1.6.7	<i>mixed-solvent</i> Simulationen	95
5.2	Biochemische Methoden	97
5.2.1	p97-Konstrukte	97
5.2.2	Protein Charakterisierung	99
5.2.2.1	Ermittlung von Proteinkonzentrationen	99
5.2.2.2	SDS-PAGE Gele	99
5.2.3	Expression und Aufreinigung von rekombinanten p97-Spezies	99
5.2.3.1	Expression von p97-Spezies	99
5.2.3.2	Aufreinigung von p97 Spezies	100
5.2.4	Herstellung und Biotinylierung von AviTag-p97	102
5.2.4.1	Biotinylierung von AviTag-p97-Konstrukten unter Verwendung von BirA	102
5.3	Herstellung der Stammlösungen ausgewählter Fragmente	102
5.4	Kristallisation von p97-ND1	102
5.4.1	Initiales-Kristallisationsscreening und Optimierung der gefundenen Bedingung	103
5.4.2	Ko-Kristallisation und Soaking ausgewählter Fragmente	105
5.4.3	Datensammlung und Auswertung	106
5.5	Biolayer-Interferometrie (BLI)-Messungen	107
5.5.1	Allgemeine Vorgehensweise	107
5.5.2	Beladen der Sensoren	108
5.5.2.1	Beladen der Sensoren mit p97-ND1	108
5.5.2.2	Beladen der Sensoren mit p97-N	109
5.5.3	Validierung des BLI-Assays	109
5.5.4	Fragmentscreening	111
5.5.4.1	Initiale-Screenings	111
5.5.4.2	Auswertung der initialen-Screenings	113
5.5.4.3	Auswahl der Fragmente aus dem initialen-Screening für Konzentrationsabhängigkeitsmessungen	115
5.5.4.4	Konzentrationsabhängigkeitsmessungen ausgewählter Fragmente	115

5.5.4.5	Analyse der Plateauphasen	116
5.5.5	Messungen der Affinitäten der Bindungen von Xanthohumol (XN) an die N- und ND1-Domäne von p97 . .	118
5.6	STD-NMR-Messungen	119
5.6.1	Durchführung der STD-NMR-Messungen	119
5.6.1.1	Einstellungen des Spektrometers	119
5.6.1.2	Umpufferung des Proteins	119
5.6.1.3	Messungen ausgewählter Fragmente des virtuellen Screenings	119
5.6.1.4	Messungen an der ND1-Domäne	120
5.6.2	Auswertung der STD-NMR-Messungen	120
5.7	Fluoreszenz-anisotropie-Kompetitionsassay	121
6	Ergebnisse & Diskussion	123
6.1	Virtuelles Screening an der SHP-Bindestelle von p97	123
6.1.1	Charakterisierung der Fragmentbibliotheken anhand physikochemischer Eigenschaften	123
6.1.2	Strukturbasierte Analyse des SHP-Bindemotivs und der Bindestelle von UFD1 und Derlin-1	125
6.1.2.1	Analyse der Kristallstrukturen der SHP-Motive von UFD1 und Derlin-1 im Komplex mit p97-N	125
6.1.2.2	Interaktionspotential-Analysen für die SHP-I und SHP-II Bindestellen	127
6.1.2.3	Analyse mittels des FTMap-Servers	130
6.1.2.4	Einfluss von Wassermolekülen innerhalb der SHP-I Region	132
6.1.2.5	Zusammenfassung der Strukturanalysen	136
6.1.3	Virtuelles Screening an der SHP-I Bindestelle	137
6.1.3.1	Pharmakophormodelle und Pharmakophor-Screenings	137
6.1.3.2	Docking	139
6.1.3.3	MD-Simulationen	146
6.1.3.4	Auswahl von 10 Fragmenten für experimentelle Untersuchungen	151
6.1.3.5	Zusammenfassung und Diskussion des virtuellen Screenings an der SHP-I Bindestelle in der p97-N Subdomäne	155
6.2	Biophysikalische Messungen an p97	157
6.2.1	Validierung des Biolayer-Interferometrie Assays	157

6.2.1.1	Auswertung mittels des Langmuir-Modells	158
6.2.1.2	Kinetische Auswertung	159
6.2.1.3	Diskussion der Validierung	162
6.2.2	Der p97-Inhibitor Xanthohumol	165
6.2.2.1	Messungen an der isolierten N-Domäne	165
6.2.2.2	Auswertung des Gleichgewichtszustandes mittels des Langmuir-Modells	166
6.2.2.3	Kinetische Auswertung der Messungen	167
6.2.2.4	Messungen an der ND1-Domäne	169
6.2.2.5	Untersuchung einer konformationsbedingten Abhängigkeit der Bindung von XN an p97-ND1	175
6.2.2.6	Zusammenfassung und Diskussion	177
6.2.3	Fragmentscreenings	180
6.2.3.1	Initiales-Screening mit p97-ND1	180
6.2.3.2	Messung der Konzentrationsabhängigkeit ausgewählter Fragmente des initialen p97-ND1-Screenings	183
6.2.3.3	Zusammenfassung des initialen-Screenings mit p97-ND1	185
6.2.3.4	Initiales-Screening mit der isolierten N-Domäne von p97	186
6.2.3.5	Messung der Konzentrationsabhängigkeit ausgewählter Fragmente des initialen p97-N-Screenings	188
6.2.3.6	Vermessung gefundener Treffer des p97-N-Screenings mit p97-ND1	191
6.2.4	Analyse der Sensorgramme des Screenings mit der N-Domäne	194
6.2.4.1	Kurvenverläufe in den Sensorgrammen	194
6.2.4.2	Abhängigkeiten von einem 1:1-Bindemodell	198
6.2.4.3	Analyse der Plateauphasen der Sensorgramme	200
6.2.4.4	Zusammenfassung der Analysen der Sensor-gram- me des p97-N-Screenings	201
6.2.5	Ergebnisse der orthogonalen STD-NMR-Messungen	204
6.2.6	Zusammenfassung des Fragmentscreenings an p97	206
6.3	Biophysikalische Untersuchungen ausgewählter Fragmente aus dem virtuellen Screening	212
6.3.1	Etablierung des STD-NMR-Assays	212
6.3.2	STD-NMR-Messungen ausgewählter Fragmente aus dem virtuellen Screening	215

6.3.3	BLI-Messungen ausgewählter Fragmente aus dem virtuellen Screening	217
6.4	Identifizierung möglicher Bindestellen der gefundenen Fragmente	221
6.4.1	<i>mixed-solvent</i> Simulationen ausgewählter Fragmente	221
6.4.2	STD-NMR-Messungen an Punkt-Mutatanten von p97-ND1	226
6.4.3	Kristallstruktur des TROLL2 p97-ND1 Komplexes	229
6.4.3.1	Kristallisation von Apo-p97-ND1 ^{L198W}	229
6.4.3.2	Datensammlung und Strukturlösung von p97-ND1 ^{L198W} im Komplex mit Fragmenten	230
6.4.3.3	Struktur des Fragments TROLL2 im Komplex mit p97-ND1 ^{L198W}	232
6.5	Zusammenfassung und Diskussion des virtuellen Screenings und Fragmentscreenings an p97	240
6.5.1	Vergleich der Struktur motive	240
6.5.2	Dreidimensionale Strukturanalyse	242
6.5.3	Ligandeffizienz-Indizes	246
6.5.4	Relevanz der Fragmente für einen möglichen PPI-Inhibitor der SHP-Bindestelle	250
7	Zusammenfassung & Ausblick	253
7.1	Zusammenfassung	253
7.2	Ausblick	255
A	Verteilung unterschiedlicher Deskriptoren der eingesetzten Fragmentbibliotheken	257
B	Docking und MD-Posen der für die experimentelle Validierung ausgewählten Fragmente des virtuellen Screenings an der N-Domäne	261
C	Retrospektive Betrachtung der Dockingergebnisse	267
D	Sensorgramme der Xanthohumol-Messungen mit p97	269
D.1	Messungen mit der N-Domäne (SSA-Sensoren)	269
D.2	Messungen mit der ND1-Domäne (SSA-Sensoren)	272
D.3	Messungen mit der ND1-Domäne (SAX-Sensoren)	274
E	Sensorgramme und Langmuir-Modelle des p97-ND1-Screenings	277
F	Sensorgramme und Langmuir-Modelle des p97-N-Screenings	279

G	Sensorgramme und Langmuir-Modelle der Vermessung ausgewählter Treffer des p97-N-Screenings mit p97-ND1	287
H	Kinetische Auswertung der Fragmente anhand ihrer empirischen Geschwindigkeitskonstanten	291
I	Eingesetzte Filter und Einstellungen des <i>Conformational Imports</i> in MOE295	
I.1	Eingesetzte Filter in der Fragmentbibliothek Generierung	295
I.2	Einstellungen <i>Confomation Import</i>	296
J	Protokolle der AMBER MD-Simulationen	299
J.1	Energieminimierung in <i>sander</i>	299
J.2	Äquilibrierung	300
J.3	Produktionsläufe	301
	Literatur	303

Abbildungsverzeichnis

2.1	Aufbau der AAA+ ATPase p97	6
2.2	Funktionen von p97 in der Zelle	7
2.3	Übersicht der an die N-Domäne bindenden Motive und Domänen	11
2.4	Struktur von Xanthohumol	18
2.5	Strukturen der Inhibitoren CB-5083 und UPCDC30245 im Komplex mit p97	20
2.6	Fragmentoptimierung	52
2.7	NMR Vektormodell	56
2.8	Interferenz an plan-parallelen Ebenen	65
2.9	Aufbau und Funktionsweise der optischen Sensoren der Biolayer- Interferometrie	67
3.1	Zielsetzung der Arbeit	73
5.1	Schematischer Aufbau des virtuellen Screenings	86
5.2	Schematischer Ablauf des <i>Protein Ligand Interaction Fingerprints</i> in MOE	88
5.3	Schematischer Ablauf des Fragmentscreenings	111
5.4	Aufbau Screeningplatte	112
5.5	Abschätzen der Plateauphasen	117
5.6	Aufbau eines Fluoreszenzanisotropie-Experiments	122
6.1	PCA und chemischer Raum	124
6.2	Die SHP-Bindestelle	127
6.3	Interaktionspotentiale für die SHP-II Bindestelle	128
6.4	Interaktionspotentiale für die SHP-I Bindestelle	129
6.5	Mittels FTMap-Server vorhergesagte <i>hot-spots</i>	131
6.6	Gefundene Wechselwirkungen der chemischen Proben innerhalb der SHP-I Bindestelle	132
6.7	Wassermoleküle in bekannten p97-N Kristallstrukturen	135
6.8	Gefundene Wasserstrukturen auf Grundlage einer 100 ns MD-Simulation	135
6.9	Schematischer Aufbau der Pharmakophormodelle	137

6.10	Struktureller Aufbau der Pharmakophormodelle	138
6.11	Verteilung der $Score_{HA}$ Werte für das erste Docking	140
6.12	Verteilung der $Score_{HA}$ Werte für das zweite Docking	141
6.13	ChemPLP-Ergebnisse der PLIF-Analysen im Hinblick auf die ver- wendeten Pharmakophormodelle	143
6.14	ASP-Ergebnisse der PLIF-Analysen im Hinblick auf die verwen- deten Pharmakophormodelle	144
6.15	Ergebnisse der 25 ns MD-Simulationen	150
6.16	Ergebnisse der 50 ns-MD Simulationen	151
6.17	Docking und MD-Profil der als Ligand bestätigten Verbindung SHP9153	
6.18	Docking und MD-Profil der als Ligand bestätigten Verbindung SHP3154	
6.19	Sensorgramme für ADP	158
6.20	1:1-Langmuir-Modell ADP	159
6.21	Ergebnisse der kinetischen Auswertung unter Annahme eines 1:1-Bindungsmodells	161
6.22	Kinetische Auswertung mittels k_{obs}	162
6.23	Abhängigkeit des Messsignals von der Konformation von p97	164
6.24	Sensorgramme zweier Xanthohumol (XN)-Messungen	165
6.25	Langmuir-Modell für die Auswertung der XN-Messungen	166
6.26	Ergebnisse der kinetischen Auswertung für einen Sensor der XN-Messungen	168
6.27	Kinetische Auswertung mittels der empirischer Geschwindigkeits- konstanten für XN	169
6.28	Sensorgramme zweier XN-Messung unter Verwendung von SAX-Sensoren	170
6.29	Sensorgramme zweier XN-Messungen unter Verwendung von SSA-Sensoren	171
6.30	Langmuir-Modell für XN-Messungen an p97-ND1 (SSA)	171
6.31	Langmuir-Modell für XN-Messungen an p97-ND1 (SAX)	172
6.32	Ergebnisse der kinetischen Auswertung von XN an der ND1-Domäne für einen jeweils einen SAX- und SSA-Sensor	173
6.33	Kinetische Auswertung mittels der empirischer Geschwindigkeits- konstanten für XN an die ND1-Domäne	175
6.34	Ergebnisse der Untersuchung einer konformationsbedingten Abhängigkeit der Bindung von XN an p97	176
6.35	Fluoreszenzanisotropie Kompetitionsassay	179
6.36	Z-Faktoren des initialen-Screenings mit p97-ND1	182
6.37	Gemessene Verschiebungen der Fragmente des initialen-Screening mit p97-ND1	182

6.38	Normierung der Signale	183
6.39	Übersicht des Fragmentscreenings mit p97-ND1	184
6.40	Vier Sensorgramme von Fragmenten, die eine Konzentrations- abhängigkeit der Verschiebungen im p97-ND1 Screening zeigten .	186
6.41	Doppelreferenzierte Daten aller Fragmente des initialen-Screenings mit p97-N	187
6.42	Übersicht des Fragmentscreenings mit p97-N	187
6.43	Vier Sensorgramme von Fragmenten die eine Konzentrations- abhängigkeit der Verschiebungen im p97-N Screening zeigten . . .	189
6.44	Langmuir-Modelle für die vier definierten Kategorien von Fragmenten	190
6.45	Gegenüberstellung der erhaltenen Affinitäten mit p97-N und p97-ND1	193
6.46	Globales 1:1-Modell mit Massentransport-Effekt	195
6.47	Abhängigkeit der Kurvenverläufe von einem 1:1- bzw. 2:1-Bindungsmodell	196
6.48	Gegenüberstellung der kinetischen Affinitäten (basierend auf k_{obs}) und der im Langmuir-Modell erhaltenen Affinitäten	199
6.49	Auswertung der Plateauphasen	201
6.50	Zusammenfassung der Auswertung der Sensorgramme	203
6.51	STD-NMR-Messungen aller Treffer des N-Domänen-Screenings . .	206
6.52	STD-NMR Messungen an ID5	213
6.53	Relative std_n^{rel} -Signale des Fragments ID5	215
6.54	STD-NMR-Messungen der Fragmente des virtuellen Screenings . .	217
6.55	BLI Einpunktmessungen der Fragmente des virtuellen Screenings .	218
6.56	Sensorgramme der Fragmente des virtuellen Screenings	219
6.57	Gemessene Affinitäten der Fragmente SHP3 und SHP9	220
6.58	Kinetische Auswertung der Daten von SHP9 (p97-N) mittels k_{obs} .	220
6.59	Übersicht über die gefundenen Bindestellen anhand der Besetzungs- dichten der Dummy-Atome	224
6.60	Besetzungsdichten für unterschiedliche Schweratome	225
6.61	p97-ND1-Alanin-Mutanten	228
6.62	STD-NMR-Messungen an p97-ND1 Alanin-Punkt-Mutanten	228
6.63	Abhängigkeit des STD-Signals von der Proteinkonzentration	229
6.64	Proteinkristalle der p97-ND1 Domäne	230
6.65	Kristallpackung der erhaltenen p97-ND1 ^{L198W} Kristalle in der Raumgruppe-P622	232
6.66	Vergleich der p97-ND1-TROLL2 Struktur mit PDB: 5DYG	235
6.67	Anomales Signal des Bromatoms von TROLL2	236

6.68	Bindetasche und Elektronendichten für TROLL2	236
6.69	Interaktionen von TROLL2 mit der p97-N-Domäne	238
6.70	Strukturmotive der Treffer des p97-N-Screenings	241
6.71	PCA der an p97-N bindenden Fragmente und deren Lage im Bezug zu benutzten Fragmentbibliotheken	244
6.72	Ligandbasierte Pharmakophore Modelle	245
6.73	Bindestellen der Fragmente in Bezug zu der Bindung des SHP-Motivs von UFD1	251
7.1	Grafische Zusammenfassung der Arbeit	254
A.1	Verteilung verschiedener Moleküldeskriptoren (Teil.1)	257
A.2	Verteilung verschiedener Moleküldeskriptoren (Teil.2)	258
A.3	Verteilung verschiedener Moleküldeskriptoren (Teil.3)	259
B.1	Docking und MD-Profil der Verbindung SHP1	261
B.2	Docking und MD-Profil der Verbindung SHP2	262
B.3	Docking und MD-Profil der Verbindung SHP4	262
B.4	Docking und MD-Profil der Verbindung SHP5	263
B.5	Docking und MD-Profil der Verbindung SHP6	263
B.6	Docking und MD-Profil der Verbindung SHP10	264
B.7	Docking und MD-Profil der Verbindung SHP11	264
B.8	Docking und MD-Profil der Verbindung SHP12	265
C.1	Beste Treffer Consensus-Score	267
C.2	Besten Treffer Consensus-Score	268
D.1	Sensorgramme aller sechs Messungen, der 1:1-Verdünnungsreihen von 20 bis 0.625 μ M XN mit p97-N.	269
D.2	Überlagerung der simulierten Kurven basierend auf einer globalen Regression eines 1:1-Bindungsmodells von XN-Messungen mit p97-N.	270
D.3	Überlagerung der simulierten Kurven basierend auf einer globalen Regression eines 1:1-Bindungsmodells XN-Messungen mit p97-N.	271
D.4	Sensorgramme aller vier Messungen, der 1:1-Verdünnungsreihen von 20 bis 0.625 μ M XN-Messungen mit p97-ND1.	272
D.5	Überlagerung der simulierten Kurven basierend auf einer globalen Regression eines 1:1-Bindungsmodells von XN mit p97-ND1 (SSA-Sensoren).	273
D.6	Sensorgramme aller vier Messungen, der 1:1-Verdünnungsreihen von 20 bis 0.625 μ M XN mit p97-ND1.	274

D.7 Überlagerung der simulierten Kurven basierend auf einer globalen Regression eines 1:1-Bindungsmodells von XN-Messungen mit p97-ND1 (SAX-Sensoren).	275
D.8 Überlagerung der simulierten Kurven basierend auf einer globalen Regression eines 1:1-Bindungsmodells von XN-Messungen mit p97-ND1 (SAX-Sensoren).	276

Tabellenverzeichnis

2.1	Übersicht der in der Literatur bekannten Inhibitoren von p97	21
2.2	Kriterien der Ro3	46
2.3	Chemischen Sonden des FTMap-Algorithmus	54
4.1	In dieser Arbeit verwendete Software	75
4.2	Verwendete Puffer	76
4.3	Verwendete Geräte	77
4.4	Chemikalien	78
4.5	Screening Verbindungen	79
4.6	Labeling-Kits, Enzyme und Sonstiges	80
4.7	Sonstiges Verbrauchsmaterial	80
5.1	Fragmentbibliotheken	82
5.2	Berechnete Deskriptoren der Fragmentbibliotheken	84
5.3	Behandlung der Wassermoleküle in GOLD basierend auf den vier Pharmakophormodellen	87
5.4	Konstrukte von p97 und deren Verwendung	98
5.5	Screens des initialen-Kristallisationsscreenings	104
5.6	Qualität eines Assays anhand des Z-Faktors	113
6.1	Ergebnisse der Pharmakophor-Screenings	138
6.2	Mittelwerte und Grenzwerte basierend auf den $Score_{HA}$ Werten des ersten Dockinglaufs	140
6.3	Anzahl Moleküle für den zweiten Dockinglauf nach Selektion durch die PLIF-Analyse.	141
6.4	Selektionskriterien für die PLIF-Analyse	146
6.5	Ausgewählte Fragmente (ChemPLP-Scoringfunktion) für die MD-Simulationen (25 ns)	148
6.6	Ausgewählte Fragmente (ASP-Scoringfunktion) für die MD-Simulationen (25 ns)	149
6.7	Übersicht der für die experimentellen Tests ausgewählten Frag- mente.	152
6.8	Kinetische Auswertung der ADP-Messungen	160

6.9	Ergebnisse der kinetischen Analyse mittels globaler exponentieller Regression der XN-Messungen	167
6.10	R^2 Werte der einzelnen Regressionen für ein 1:1-Modell	168
6.11	Ergebnisse der kinetischen Analyse mittels globaler exponentieller Regression für XN an der ND1-Domäne (Messungen mit SSA-Sensoren)	174
6.12	Ergebnisse der kinetischen Analyse mittels globaler exponentieller Regression für XN an der ND1-Domäne (Messungen mit SAX-Sensoren)	174
6.13	Zusammenfassung aller XN-Messungen an der N- und ND1-Domäne	177
6.14	Priorisierung der Fragmente anhand des $Score_{BLI}$	202
6.15	STD-NMR-Messungen aller Fragmente mit p97-ND1	205
6.16	Übersicht Fragmentscreening	207
6.17	Übersicht aller an p97 bindenden Fragmente	208
6.18	STD-NMR-Messungen ausgewählter Fragmente des virtuellen Screenings	216
6.19	Kristallisationsbedingung für p97-ND1	230
6.20	Statistiken der Datensammlungen von mit TROLL2 ko-kristallisierten und gesoakten p97-ND1 ^{L198W} Kristallen	233
6.21	Statistiken der Verfeinerung von Datensatz 1	234
6.22	Interaktionen von TROLL2 mit p97	238
6.23	Vergleich der Mittelwerte verschiedener LE-Indizes der ausgewählten Fragmente mit in der Literatur berichteten Fragment-Screenings	247
6.24	Indizes ausgewählter Fragmente des N-Domänen Screenings . . .	248
6.25	Indizes der bei <i>Chimenti et al.</i> berichteten Fragmenten, sowie von XN	249
C.1	Übersicht Fragmente Consensus-Score	268
E.1	Sensorgramme und Langmuir-Modelle aller identifizierten Treffer .	277
F.1	Sensorgramme aller im N-Domänen Screening identifizierten Treffer	279
G.1	Sensorgramme und Langmuir-Modelle der Vermessung ausgewählter Treffer des p97-N Screenings mit p97-ND1	287
G.2	Sensorgramme aller Fragmente diese nicht mit dem Langmuir-Modell modelliert werden konnten	290
H.1	Lineare Regressionen der aus der kinetischen Auswertung erhaltenen k_{obs} in Abhängigkeit von der Konzentration	291

Abkürzungsverzeichnis

ADMET	Absorption, Distribution, Metabolismus und Elimination
ADP	Adenosindiphosphat
ATP	Adenosintriphosphat
BLI	Biolayer-Interferometrie
CAD	Chromatin-assoziierten Abbau (<i>Chromatin Associated protein Degradation</i>)
CADD	Computerbasiertes Wirkstoffdesign (<i>Computer-Aided Drug Design</i>)
Cdc48	<i>Cell division cycle protein 48</i>
DUB	Deubiquitinasen
ER	Endoplasmatisches Retikulum
ERAD	Endoplasmatischer Retikulum-assoziiertes Abbau (<i>Endoplasmic Reticulum Associated protein Degradation</i>)
FBDD	Fragmentbasiertes Wirkstoffdesign (<i>Fragment Based Drug Design</i>)
GA	Genetischer Algorithmus
GAFF	<i>General Amber Force Field</i>
HTS	Hochdurchsatzverfahren (<i>High-Throughput Screening</i>)
ITC	Isotherme Titrationskalorimetrie (<i>Isothermal Titration Calorimetry</i>)
LE	Ligandeffizienz (<i>Ligand Efficiency</i>)
LELP	<i>Lipophilicity corrected Ligand Efficiency</i>
LLE	<i>Lipophilic Ligand Efficiency</i>
$\log P$	n-Octanol/Wasser Koeffizient
MAD	Mitochondrien-assoziiertes Abbau (<i>Mitochondria Associated protein Degradation</i>)
MD	Molekulardynamik
MR	molekularer Ersatz (<i>Molecular Replacement</i>)
MS	Massenspektrometrie
MST	<i>Microscale Thermophoresis</i>
NMR	Kernspinresonanz (<i>Nuclear Magnetic Resonance</i>)
NOE	Kern-Overhauser-Effekt (<i>Nuclear Overhauser Effect</i>)

NPT	isotherm-isobares Ensemble mit konstanter Anzahl an Teilchen N , einer konstanten Temperatur T und konstantem Druck P
NTP	Nukleosidtriphosphat
NVE	mikrokanonisches System mit konstanter Anzahl an Teilchen N , einem konstanten Volumen V und der konstanten Energie E
NVT	kanonische Ensemble mit konstanter Anzahl an Teilchen N , einer konstanten Temperatur T und konstantem Volumen V
PCA	Hauptkomponenten-Analyse (<i>Principal Component Analysis</i>)
PPI	Protein-Protein-Interaktionen
RF	Radiofrequenz
Ro3	<i>Rule of 3</i>
Ro4	<i>Rule of 4</i>
Ro5	<i>Rule of 5</i>
RQC	Ribosom-assoziierte Qualitätskontrolle (<i>Ribosome Quality Control</i>)
SAR	Struktur-Wirkungsbeziehungen (<i>Structure/Activity Relationship</i>)
SAXS	Röntgenkleinwinkelstreuung (<i>Small Angle X-ray Scattering</i>)
SILE	Größenunabhängiger Ligandeffizienz-Index (<i>Size-Independent Ligand Efficiency</i>)
SPR	Oberflächenplasmonenresonanz-Spektroskopie (<i>Surface-Plasmon-Resonance Spectroscopy</i>)
STD	Sättigungstransfer-Differenz
TPSA	topologische polare Oberfläche (<i>Topological Polar Surface Area</i>)
TSA	<i>Thermal-Shift-Assays</i>
UBX	<i>Ubiquitin Regulatory X</i>
UBXL	<i>ubiquitin regulatory X like</i>
UPR	<i>Unfolded Protein Response</i>
UPS	Ubiquitin-Proteasom-Systems
VBM	<i>VCP-binding motif</i>
VCP	<i>Valosin-containing protein</i>
VIM	<i>VCP- interacting motif</i>
XN	Xanthohumol

Konstanten und Einheiten

Ångstrom (Einheit)	$1 \text{ \AA} = 1.0 \times 10^{-10} \text{ m}$
Kalorie (Einheit)	$1 \text{ cal} = 4.1868 \text{ J}$
Boltzmann-Konstante ^a	$k_B = 1.380\,649 \times 10^{-23} \text{ J K}^{-1}$
Avogadro-Konstante ^a	$N_A = 6.022\,140\,76 \times 10^{23} \text{ mol}^{-1}$
Gaskonstante ^a	$R = 8.314\,462\,618 \text{ J mol}^{-1} \text{ K}$
Dezimaltrennzeichen	= .
Tausendertrennzeichen	= ,

^a Quelle: CODATA 2018

für Laura Sophia Koschmal

Kapitel 1

Einleitung

Die Geschichte der Menschheit ist untrennbar mit der Suche nach Arzneien gegen Krankheiten verknüpft. Die Zeugnisse dieser Suche reichen weit in die Wiege der Zivilisation zwischen Euphrat und Tigris zurück. Dort wurden die bisher ältesten überlieferten Rezepturen von Arzneien auf sumerischen Keilschrifttafeln nahe der Stadt Nippur gefunden, mit einer Vielzahl an Zubereitungen von verschiedensten Arzneipflanzen gegen unterschiedliche Krankheiten [1]. In der heutigen Zeit ist die Suche nach Pharmaka ein komplexes, multidisziplinäres Feld, in dem die verschiedensten naturwissenschaftlichen Disziplinen zusammenarbeiten, um den Weg von der Identifizierung einer Zielstruktur, der Findung möglicher Wirkstoffe, bis hin in zu deren klinischer Phase zu begleiten.

Das Verständnis von molekularen Mechanismen von Krankheiten bildet für den Wirkstofffindungsprozess eine essentielle Grundlage. Vor diesem Hintergrund ist das Graduiertenkolleg 2243 „Ubiquitinierung verstehen: Von molekularen Mechanismen zu Krankheiten“ entstanden, in dessen Rahmen diese Arbeit durchgeführt wurde. Die Ubiquitinierung ist ein komplexer Prozess, bei dem Zielproteine durch eine enzymatische Kaskade mit der Übertragung des Proteins Ubiquitin modifiziert werden, welches vielfach zum Abbau des Zielproteins führt.

Die Forschungsschwerpunkte des Graduiertenkollegs 2243 sollen ein tieferes Verständnis ermöglichen: der Mechanismen und Regulation von ubiquitinierenden Enzymen und der Substrat-Erkennung von de-ubiquitinierenden Enzymen; der transkriptionellen Kontrolle von (de)-ubiquitinierenden Enzymen; und der eng mit der Ubiquitinierung verknüpften AAA+ ATPase p97 für Gesundheit und Krankheit eines Organismus.

Ein Langzeitziel des Graduiertenkollegs 2243 ist auch die Entwicklung von niedermolekularen Inhibitoren, die krankheitsrelevante Zielstrukturen im Ubiquitinierungssystem angreifen sollen. In diesem Kontext liegt der Schwerpunkt der hier vorliegenden Arbeit: Startstrukturen für mögliche Protein-Protein-Interaktionsinhibitoren der AAA+ ATPase p97 zu identifizieren.

Kapitel 2

Theorie

2.1 Die AAA+ ATPase p97

Die Hydrolyse von Adenosintriphosphat (ATP) durch die Enzymklasse der ATPasen und die dadurch freiwerdende Energie, die für verschiedenste Reaktionen zur Verfügung steht, ist eine grundlegende Reaktion in allen bekannten Lebensformen. Eine große Gruppe innerhalb der existierenden ATPasen stellen die sogenannten Walker-Typ NTP (Nukleosidtriphosphatasen) dar (benannt nach Sir John Ernest Walker), in deren Gruppe sich die Superfamilie der AAA+ ATPasen (*ATPases associated with diverse cellular activities*) befindet. In den etwa 200 bis 250 Aminosäuren großen AAA+ Kassetten, findet man neben dem Walker A (oder P-Schleife) und dem Walker B Motiv, die für Bindung und Hydrolyse von ATP verantwortlich sind, weitere Motive, die sie von den klassischen P-Schleifen NTPasen unterscheiden. Angehörige der AAA+ Proteinfamilie vollführen zumeist mechanochemische Arbeit, d.h. sie wandeln durch Konformationsänderung die bei der Hydrolyse von ATP freiwerdende Energie in mechanische Bewegung um. Durch diese mechanochemische Arbeit werden eine Vielzahl an Aufgaben innerhalb der Zelle, wie Translokation von Proteinen oder Nukleinsäuren, Remodellierung (Strukturänderungen) von Protein-Komplexen oder Entfaltung von Proteinen zum Abbau vollführt. Der typische Aufbau eines solchen AAA+ Motor-Proteins besteht aus einem ringförmigen Hexamer, wobei ein axialer Kanal zwischen den sechs Monomeren entsteht. 1982 wurde in einem genetischen Screening im Labor von David Botstein die zu dieser Superfamilie gehörende ATPase p97, auch bekannt unter den Namen *Valosin-containing protein* (VCP) oder *Cell division cycle protein 48* (Cdc48) in *Saccharomyces cerevisiae*, identifiziert [2-5]. p97 und Homologe sind ubiquitär in *Eukarya*, *Archaea*, wie auch in einigen *Eubacteria* zu finden, womit p97 in allen drei Domänen des Lebens vertreten ist [6]. Es spielt in einer Vielzahl an verschiedensten zellulären Prozessen eine Rolle und macht einen Anteil von bis zu 1% der Masse aller cytosolischen Proteine aus [7].

2.1.1 Struktur und Aufbau von p97

Die dreidimensionale Struktur von p97 wurde mittels Elektronenmikroskopie, Röntgenkleinwinkelstreuung (*Small Angle X-ray Scattering*, SAXS) und Röntgenstrukturanalyse Anfang der 2000er Jahre aufgeklärt. Die funktionale Form von p97 bildet ein Homohexamer, gebildet aus sechs Monomeren von je etwa 90 kDa. Zu den Proteinen der AAA+ Superfamilie gehörend, lässt sich p97 hier zu der Unterfamilie der „N-Ethylmaleimid-Sensitiven Fusion (NSF)/Cdc48/Pex“-Proteine zählen. Proteine dieser Unterfamilie beinhalten je Monomer eine N-terminale Domäne (N-Domäne), welche zur Kofaktor- und/oder Substraterkennung nötig ist, gefolgt von zwei AAA+ Kassetten (D1- und D2-Domäne). Als zusätzliches Strukturelement besitzt p97 einen unstrukturierten C-terminalen Bereich, der ebenfalls in die Bindung von Kofaktoren beteiligt ist, eine Phosphorylierungsstelle beinhaltet und die ATPase Aktivität der D2-Domäne beeinflusst. Den zentralen Bereich von p97 bilden die D1- und D2-Domänen der sechs Monomeren, die zwei gestapelte Ringe bilden, der von einem axialen Kanal durchzogen wird. Die N-Domänen liegen im Adenosindiphosphat (ADP)-gebundenen Zustand in der Ebene der D1-Domänen und zeigen von der sechs-zähligen Drehachse weg (Abbildung 2.1 (a)) [8–10].

Die N-Domäne (Aminosäuren 1-187 im humanem p97) lässt sich wiederum in zwei Subdomänen einteilen: einer doppelten $\Psi\beta$ -Fass Struktur (Nn, Aminosäuren 21-105) und einer Vier- β -Strang-Fass Struktur (Nc, Aminosäuren 112-187). Beide Subdomänen werden durch eine sechs Aminosäuren lange Schleife verbunden. Die beiden AAA+ Kassetten D1 (Aminosäuren 209-460) und D2 (Aminosäuren 481-763) bestehen jeweils aus einer α/β P-Schleifen NTPase Domäne, die eine sogenannte Rossmann-Faltung einnimmt. In der Nukleotidbindestelle findet man die klassischen Walker A (P-Schleife; Konsensusmotiv: G(x)₄GKT; x, beliebige Aminosäure) und Walker B (Konsensusmotiv: hhhhDE; h, hydrophobe Aminosäure) Motiv. Ein konserviertes Asparagin (N348 in D1, N624 in D2) fungieren jeweils als „Sensor 1“ und helfen bei der Koordination des angreifenden Wassermoleküls. Der sogenannte Arginin-Finger (R359 in D1, R635 in D2) stabilisiert die neu entstehende negative Ladung während der ATP-Hydrolyse. Die N-Domäne ist mit der D1-Domäne über eine Schleife (Aminosäuren 188–208), wie auch die D1-Domäne mit der D2-Domäne über eine Schleife (Aminosäuren 461–480) verbunden. Beide Schleifen stehen in unmittelbarem Kontakt mit den Nukleotidbindestellen der D1- bzw. D2-Domäne. Der durch die sechs Monomere geformte Kanal weist eine Engstelle in der D1-Domäne auf, die durch sechs Histidine (Aminosäuren H317) gebildet wird und auch als „His Gate“ bezeichnet wird.

Zwischen den D2-Domänen weist der Kanal einen deutlich größeren Durchmesser auf und beinhaltet zwei postulierte Substratbindeschleifen. Die D2-Domäne enthält eine kleinere Schleife 1 in α -Helix 2 (Aminosäuren 549-553) und eine längere Schleife 2 in α -Helix 3 (Aminosäuren 585-595) mit zwei hoch-konservierten Argininen (R586 und R599), die zwei positiv geladene Ringe am Eingang der Kanal Pore bilden [10–12].

p97 liegt abhängig von der Besetzung der Nukleotidbindestellen in drei Konformationen vor, wobei für die Betrachtung der N-Domäne nur Konformation-I (beide ATPase Domänen sind mit ADP besetzt) und Konformation-III (beide ATPase Domänen sind mit ATP besetzt) von Bedeutung sind. Konformation-II stellt einen Übergangszustand dar, bei dem nur die D2-Domänen mit ATP besetzt sind, während die D1-Domänen mit ADP besetzt sind [13]. Bei Konformation-I liegen die N-Domänen in der Ebene der D1-Domänen, wohingegen in der Konformation-III die N-Domänen nach oben aus der Ebene heraus zeigen. Bei beiden Konformationen werden unterschiedliche Bereiche der N-Domäne sterisch abgeschirmt (siehe Abbildung 2.1 **(b)** und **(c)**).

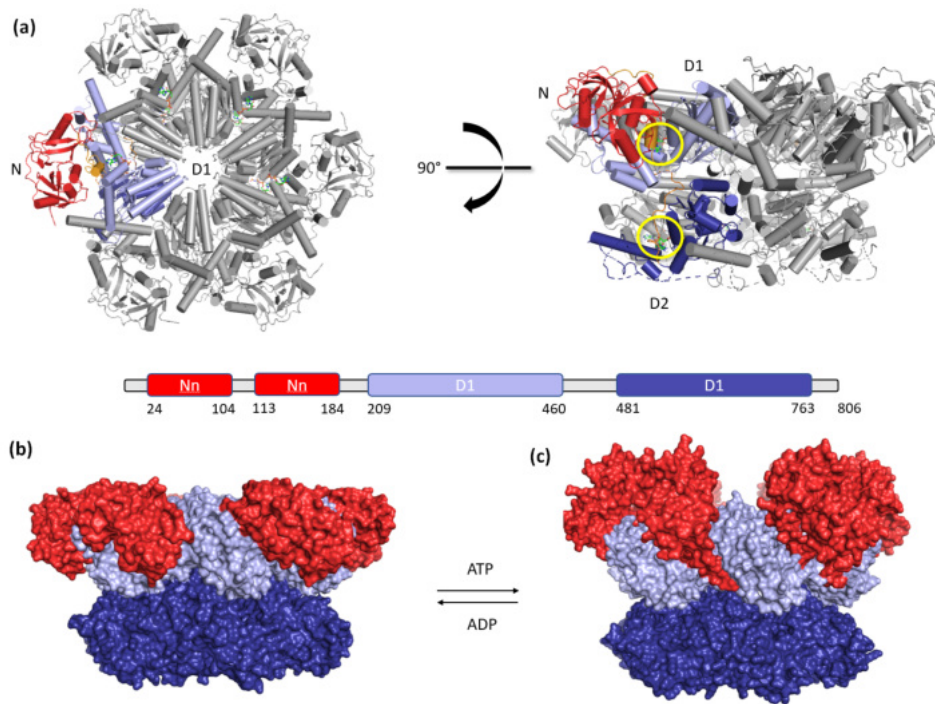


Abbildung 2.1: Aufbau der AAA+ ATPase p97

(a) p97 Monomer mit der N- (rot), D1- (hellblau) und D2-Domäne (dunkelblau). Schleifen zwischen N- und D1-Domäne bzw. D1- und D2-Domäne sind orange dargestellt. In Gelb sind die beiden Nukleotidbindestellen in der D1- und D2-Domäne gezeigt. In der Mitte ist der axiale Kanal durch das Hexamer zu sehen. Darunter ist die Domänen-Architektur eines p97 Monomers gezeigt.

(b) Oberflächendarstellung der ADP-gebundenen Konformation von p97 mit der N-Domäne in der Ringebene, sowie (c) der ATP-gebundenen Konformation mit nach oben zeigenden N-Domänen. (Die Abbildung wurden unter Verwendung der Strukturen aus der Protein Datenbank (*Protein Data Bank*, PDB [14]): 5FTK und 5FTN erstellt.)

2.1.2 Biologische Funktionen von p97 und Verknüpfung mit dem Ubiquitin-Proteasom-System

p97 nimmt als Bestandteil des Ubiquitin-Proteasom-Systems (UPS) eine zentrale Funktion beim Schutz einer Zelle vor cytotoxischen (proteotoxischen) Effekten ein, die durch fehl gefaltete oder beschädigte Proteine hervorgerufen werden. p97 ist hierbei in eine Vielzahl von Qualitätskontrollpfaden an unterschiedlichen Orten in der Zelle involviert und kontrolliert im Zusammenspiel mit zahlreichen Kofaktoren die Aufrechterhaltung der Proteinhomöostase, insbesondere auch unter Stressbedingungen. Konkret ist p97 in den Endoplasmatischen Retikulum-assoziierten Abbau (*endoplasmic reticulum associated protein degradation*, ERAD), den Mitochondrien-assoziierten Abbau (*mitochondria associated protein degradation*, MAD), den Chromatin-assoziierten Abbau (*chromatin associated protein degradation*, CAD), den cytosolischen Abbau und die Ribosom-assoziierte Qualitätskontrolle (*ribosome quality control*, RQC) involviert (Abbildung 2.2) [15–17].

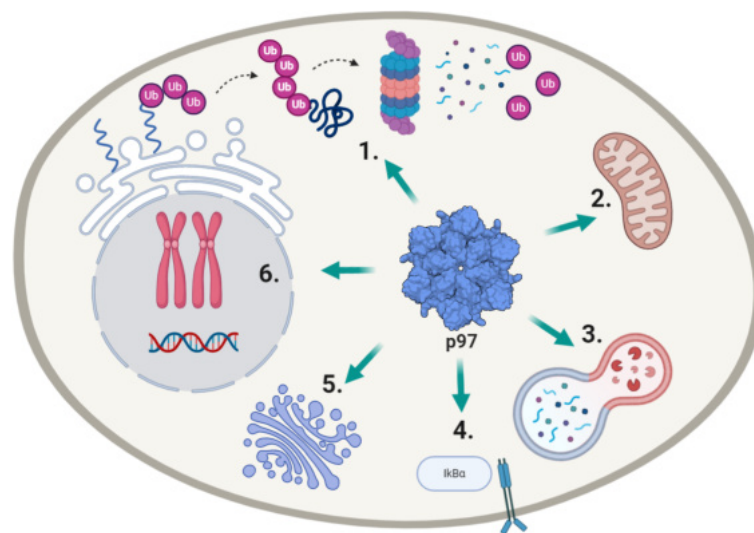


Abbildung 2.2: Funktionen von p97 in der Zelle

1. Endoplasmatisches Retikulum-assoziiertes Proteinabbau (ERAD), 2. Mitochondrien-assoziiertes Proteinabbau (MAD), 3. und 5. Membranfusionsprozesse (wie u.a. Autophagosome, lysosomale Prozesse, Reassemblierung des Golgi Apparats), 4. NF- κ B Aktivierung, 6. Chromatin-assoziiertes Proteinabbau (CAD). *Abbildung wurde mit Biorender (BioRender.com) erstellt.*

Vereinfacht dargestellt werden defekte oder fehl gefaltete Proteine mit dem Protein Ubiquitin zum Abbau markiert. Hierbei werden mehrere Ubiquitin Proteine zu einer Kette konjugiert, wobei die Verknüpfungsart entscheidend für die Signalweiterleitung ist. Im Falle des UPS werden Ubiquitinketten gebildet, die typischerweise über Lysin 48 verknüpft sind. Die Modifizierung der Substrate

läuft hierbei über eine enzymatische Kaskade bestehend aus E1 (Ubiquitin aktivierendes Enzym), E2 (Ubiquitin konjugierendes Enzym) und E3 (Ubiquitin Ligase) unter ATP-Hydrolyse ab [18]. Diese Polyubiquitin-K48-Ketten werden von p97 Kofaktoren erkannt und p97-vermittelt an das Proteasomsystem zum Abbau weitergegeben. Hierbei stellt p97 über die Hydrolyse von ATP die nötige Energie bereit, um die zum Abbau markierten Proteine aus ihrer Umgebung (Membranen, Proteinkomplexe, Chromosomen) herauszulösen und in das Cytosol zu überführen, in dem der anschließende Abbau stattfindet. Für diese Abbauprozesse spielt das Heterodimer UFD1-NPL4 (*ubiquitin fusion degradation protein 1 - nuclear protein localization protein 4 homolog*) als Kofaktor eine entscheidende Rolle; es ist in alle oben genannten Abbauwegen beteiligt [19].

Der ERAD ist sicherlich der am besten untersuchte Abbauweg, in den p97 involviert ist. p97 spielt dabei eine zentrale Rolle in der Retrotranslokation fehl gefalteter glykosylierter Proteine aus dem ER Lumen in das Cytosol für den nachfolgenden proteasomalen Abbau. Die Ubiquitin-Ligase HRD1 (*Hmg2-regulated degradation*) bildet dabei den protein-translozierenden Kanal in der ER-Membran und katalysiert die Ubiquitinierung fehl gefalteter Proteine auf der cytosolischen Seite. Die markierten Proteine werden anschließend vom p97-UFD1-NPL4 Komplex gebunden und unter ATP-Hydrolyse mechanisch aus der ER-Membran extrahiert [20]. Im Weiteren spielen p97-Kofaktoren wie z.B. die UBX-Proteine oder Deubiquitinasen (DUBs) eine Rolle. Hierdurch ist p97 ein zentrales Protein im ERAD-System und somit wichtig für die Proteinhomeostase der Zelle. Die Störung dieses Gleichgewichts durch fehl gefaltete Proteine führt schnell zum Auslösen der sogenannten *Unfolded Protein Response* (UPR), einem komplexen Signalweg, der schlussendlich zur Einleitung der Apoptose führt [21].

Ähnliches gilt für den MAD Weg, bei dem p97 ubiquitinierte Proteine aus der äußeren mitochondrialen Membran extrahiert und an das Proteasom-System übergibt. Störungen führen auch hier schnell zur Apoptose. Neben diesen Membranassoziierten Abbau Prozessen, spielt p97 auch eine entscheidende Rolle im Zellkern, bei der Extraktion und beim Abbau ubiquitiniertes Proteinen aus dem Chromatin. Zusätzlich unterschützt p97 die Reparatur von DNA-Doppelstrangbrüchen, ein Mechanismus, der mit einer Vielzahl von posttranslationalen Modifikation, unter anderem der Ubiquitinierung, verknüpft ist. Hierbei sind die Kofaktoren FAF1 und UFD1-NPL4 wahrscheinlich von Bedeutung [22]. Daneben ist p97 am Abbau von Proteinkomplexen im Cytosol im Zusammenspiel mit der Ubiquitin Ligase HUWE1 beteiligt [23].

Neben seiner Schlüsselfunktion im UPS ist p97 auch an der Aktivierung von NF- κ B (durch Abbau seines Inhibitors I κ B α), an Membranfusionsprozessen (lysosomalen Prozessen und Autophagie), und an Membranprozessen im Golgiapparat und ER beteiligt. Für nähere Details hierzu sei auf die Übersichtsartikel [16, 17, 19] verwiesen.

2.1.3 Interaktionen von p97 mit Kofaktoren

Wie bereits erwähnt ist p97 an einer Vielzahl von zellulären und regulatorischen Prozessen involviert. Diese funktionale Diversität von p97, die die enzymatische Reaktion der ATP-Hydrolyse in mechanische Energie umwandelt und einer Vielzahl an Prozessen zur Verfügung stellt, beruht auf dem Zusammenspiel von verschiedenen Kofaktoren, die an p97 binden und von denen bislang mehr als 30 bekannt sind. Typischerweise bestehen die Kofaktoren aus einem spezifischen p97 Bindemodul und zusätzlichen funktionalen Domänen, z.B. für die Erkennung von ubiquitinierten Proteinen, für katalytische Prozesse oder für Membraninteraktionen (zusammengefasst in [11, 16]). Nach *Hänzelmann et al.* [24] lassen sich die Kofaktoren von p97 aufgrund ihrer Funktion wie folgt einteilen:

1. Substrat-rekrutierende Kofaktoren, wie z.B. UBA-UBX Proteine oder UFD1-NPL4, welche neben einem spezifischen p97-Motiv ein Bindungsmotiv für Ubiquitin besitzen, um ubiquitinierte Substrate mit p97 verknüpfen;
2. Substrat-prozessierende Kofaktoren, wie Ubiquitin Ligasen (E3), Deubiquitinasen (DUBs) und N-Glykanasen (PNGasen), welche ubiquitinierte oder N-glykosylierte Proteine prozessieren;
3. Regulatorische Kofaktoren, wie z.B. UBXD4, ASPL oder SVIP.

Anhand ihrer Bindungsstellen lassen sich die Kofaktoren in zwei Gruppen einteilen: in solche, die an die N-Domäne binden und in solche, die an den C-terminalen Teil bindenden. Im Vergleich zu der Vielzahl an unterschiedlichen Kofaktoren, existiert nur eine kleine Anzahl an Bindemotiven/Domänen. Hierzu zählen für die Bindung an die N-Domäne: Die UBX-(*ubiquitin regulatory X*) und UBXL-(*ubiquitin regulatory X like*) Domänen, die zwei linearen Bindemotive VIM-(*VCP-interacting motif*) und VBM-(*VCP-binding motif*), sowie das SHP-Motiv. An den C-terminalen Teil binden Kofaktoren mit dem PUB (PNGase/UBA oder UBX Proteine; auch bekannt als PUG) und der PUL (PLAP, Ufd3p, and Lub1p) Domäne (zusammengefasst in [11, 16]). Die weitere Beschreibung beschränkt sich auf die Interaktionen an der N-Domäne.

Die genannten Bindemotive/Domänen mit Bindung an die N-Domäne interagieren nur mit zwei unterschiedliche p97 Bereichen, wobei Kofaktoren mit der UBX/UBXL Domäne-, sowie dem VIM- oder dem VBM-Motiv an die gleiche Region zwischen der Nn- und Nc-Untereinheit binden. Die Bindestelle für das SHP-Motiv befindet sich in der Nc-Untereinheit. Eine Übersicht über die unterschiedlichen Interaktionen ist in Abbildung 2.3 gezeigt.

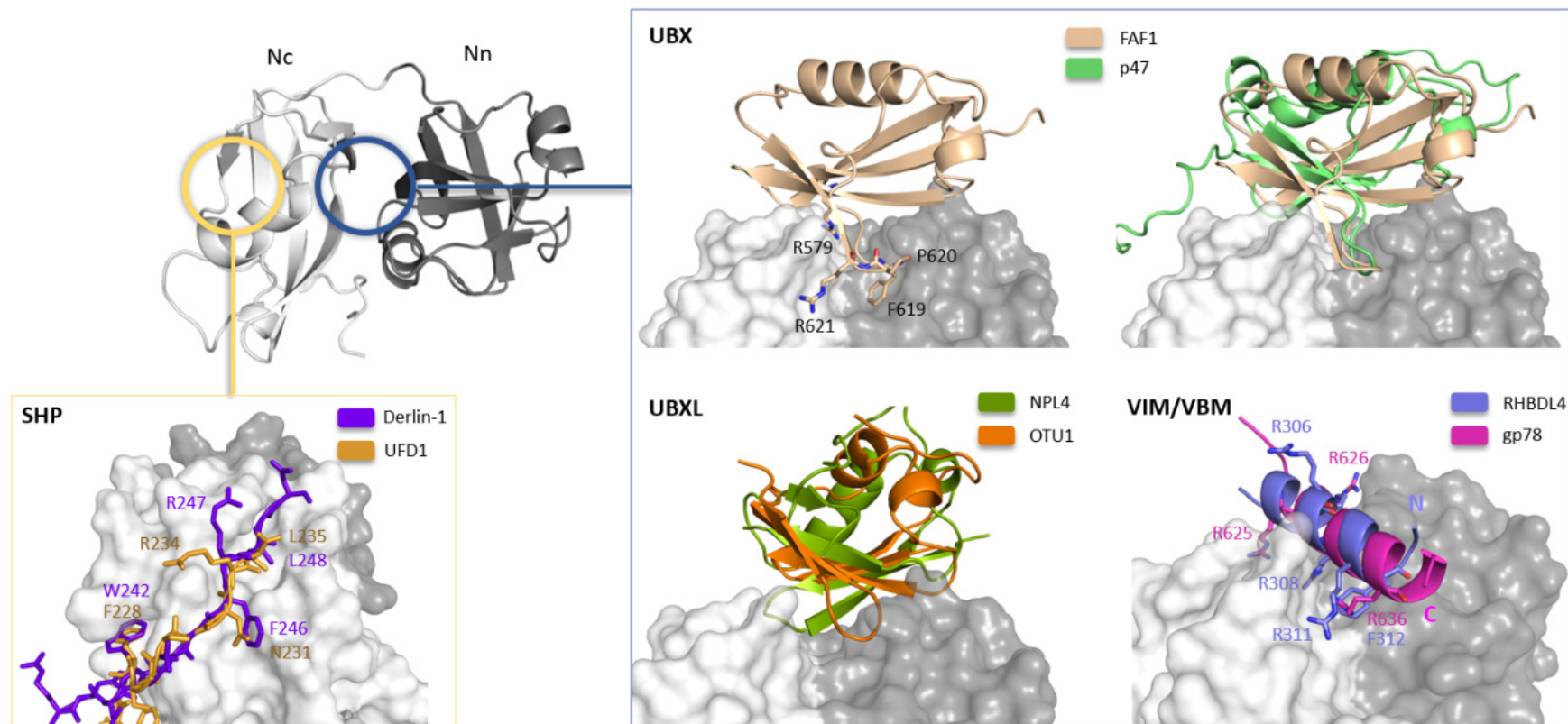


Abbildung 2.3: Übersicht der an die N-Domäne bindenden Motive und Domänen

Die Bindestellen für UBX/UBXL, VIM und VBM (blau) befindet sich zwischen der Nn- und Nc-Untereinheit, für das SHP-Motiv in der Nc-Subdomäne der N-Domäne (gelb). Eine Überlagerung der SHP-Motive von UFD1 (PDB: 5B6C) und Derlin-1 (PDB: 5GLF) ist dargestellt. Überlagert sind die UBX-Domänen der Kofaktoren FAF1 (PDB: 3QQ8) und p47 (PDB: 1S3S). Beispielfhaft ist für FAF1 die typische *FcisP touch-turn* Konformation, die an der Bindung beteiligt ist, zu sehen. UBXL Domänen binden über ein ähnliches FPP-Motiv in der gleichen Region (OTU1, PDB: 4KDI). NPL4 bindet trotz der ähnlichen Domänen-Architektur (UBXL) in einer anderen Konformation an die N-Domäne. VIM- und VBM-Motiv enthaltende Kofaktoren binden an die gleiche Stelle zwischen den Subdomänen, mit jeweils entgegengesetzter Orientierung. Zu sehen ist das für die Bindung wichtigste C-terminal liegende R636 für gp78 (PDB:3TIW). RHBDL4 bindet ähnlich den UBX-Domänen enthaltenden Kofaktoren, an dieser Stelle über ein Phenylalanin (F312) (PDB: 5EPP).

2.1.3.1 UBX- und UBXL-Domänen

UBX- und UBXL-Domänen bestehen aus etwa 80 Aminosäuren und besitzen eine dem Ubiquitin ähnliche β -„Grasp“-Faltung in Form einer β - β - α - β - β - α - β -Sekundärstruktur [25]. UBX-Kofaktoren können in zwei Gruppen eingeteilt werden: in solche, die eine UBX-Domäne und eine UBA (*ubiquitin-associated*)-Domäne besitzen, welche sie in die Lage versetzt ubiquitinierte Substrate zu binden und solche die nur eine UBX-Domäne besitzen [24].

Durch Kristallstrukturen der FAF1-UBX [26–28], UBXD7-UBX [29] und p47-UBX [30] Domänen im Komplex mit der p97 N-Domäne sind die molekularen Wechselwirkungen bekannt. Als konserviertes Bindemotiv (Signaturmotiv) besitzen alle in der Schleife zwischen β 3 und β 4 ein R...FPR-Motiv. Das Phenylalanin zeigt hierbei in eine hydrophobe Tasche, gebildet zwischen der Nc- und Nn-Subdomäne der N-Domäne. Das Motiv besitzt eine seltene *FcisP touch-turn*-Konformation. Diese cis-Peptidkonformation des Prolins mit ungewöhnlichen Ψ - und Φ -Winkeln, ermöglicht erst die Wechselwirkung des Phenylalanins mit p97. UBXL-Domänen wie z.B. in der DUB YOD1 binden ähnlich über ein FPP-Motiv [31]. Die UBXL-Domäne im Kofaktor NPL4 besitzt zwar die gleiche erwähnte β -„Grasp“-Faltung, interessanterweise fehlt hier jedoch das *touch-turn*-Motiv, so dass NPL4 in einem etwas eher zur N-D1 Schleife orientierten Bindungsmodus an p97 bindet [32].

2.1.3.2 Das VIM- und VBM-Motiv

Mit der Konsensussequenz $Rx_2h_3A_2x_2Rh$ (x , beliebige Aminosäure; h , hydrophobe Aminosäure) handelt es sich beim VIM-Motiv um eine Arginin-reiche α -Helix mit zwei konservierten Argininen, wobei das C-terminal liegende Arginin für die Bindung am wichtigsten ist. Das Motiv findet sich unter anderem in den Kofaktoren gp78, ANKZF1 und SVIP. Die Kristallstruktur des gp78 VIM-Motivs im Komplex mit der N-Domäne von p97 zeigt, dass sich die α -Helix in den Spalt zwischen den Nn- und Nc-Domänen legt [12]. Dieser Spalt wird zu einer Seite durch die Schleife zwischen den beiden Subdomänen, auf der anderen Seite durch eine Salzbrücke zwischen D35 und R144 begrenzt.

Bei dem VBM-Bindemotiv handelt es sich, ähnlich dem VIM-Motiv, ebenfalls um eine Arginin-reiche α -Helix mit der Konsensussequenz EhR_4Lxh_2 . Die Kristallstruktur des VBM-Motivs der Protease RHBDL4 mit der N-Domäne von p97

zeigt, dass sich die α -Helix ähnlich dem VIM-Motiv in die gleiche Spalte zwischen die beiden Subdomänen legt. Interessanterweise bindet das VBM-Motiv jedoch in entgegengesetzter Orientierung im Vergleich zum VIM-Motiv [33].

2.1.3.3 Das SHP-Motiv

Während die UBX- und UBXL-Domänen, sowie das VIM- und VBM-Motiv mit der gleichen Region von p97 interagieren, bindet das kurze lineare SHP-Motiv an eine weitere Region der Nc-Subdomäne von p97. Es beinhaltet die Konsensussequenz $h(x)_{1-2}F/W(x)_{0-1}GxGx_2L$ und findet sich als zusätzliches Bindemotiv in zahlreichen UBX Kofaktoren, sowie in UFD1 und Derlin-1.

Das Motiv ist unter anderem für die Bindung des Heterodimers UFD1-NPL4 an p97 wichtig und spielt daher eine essenzielle Rolle in allen Abbauprozessen, an denen UFD1-NPL4 beteiligt ist [16]. Die SHP-Binderegion in der Nc-Subdomäne von p97 ist die in dieser Arbeit primär adressierte Binderegion für die Entwicklung möglicher p97 Protein-Protein-Interaktionsinhibitoren. Auf die Beschaffenheit und Interaktionen dieses Motivs wird detailliert in der strukturellen Analyse im Ergebnisteil eingegangen. Zusammenfassend lässt sich hier festhalten, dass alle eben genannten Bindemotive, mit Ausnahme des SHP-Motivs, in die gleiche Region von p97 binden was zur Hypothese führt, dass die Konkurrenz der Kofaktoren untereinander eine weitere Ebene der Regulation darstellen könnte.

2.1.4 p97 als pharmakologischer Angriffspunkt (*Target*)

Aufgrund der Beteiligung von p97 an einer Vielzahl unterschiedlicher zellulärer Prozesse, ist es nicht verwunderlich, dass p97 bei mehreren Krankheiten eine Rolle zu spielen scheint. Hierbei unterscheidet man zwischen Erkrankungen, die mit p97-Mutationen in Verbindung stehen (neurodegenerative Erbkrankheiten) und solchen, bei denen das native p97 eine Rolle zu spielen scheint (Krebserkrankungen und virale Infektionen).

2.1.4.1 Erkrankungen, die mit p97-Mutationen in Verbindung stehen

Die seltene autosomal dominant vererbte Krankheit IBMPFD (*Inclusion Body Myopathy with Paget's Disease of the Bone and Frontotemporal Dementia*, auch als Multisystem Proteinopathie 1 (MSP1) bezeichnet), einer Symptom-Trias aus Einschlusskörpermyopathie (einer im mittleren Erwachsenenalter auftretenden

Muskelschwäche), Morbus Paget (Auftreten von verdickten und verformten Knochen) und einer frontotemporalen Demenz (Nervenzelluntergang im Frontalhirn bzw. den Schläfenlappen), steht im Zusammenhang mit Mutationen in p97. Aktuell sind ca. 20 Mutationen bekannt, die alle in der N- und D1-Domäne lokalisiert sind [34]. Aktuelle Diskussionen betreffen hier die Frage, ob die Mutationen zu einer Überaktivierung, oder einer Inaktivierung von p97 führen. Therapeutische Ansätze wären im ersten Fall eine Inaktivierung, im zweiten Fall hingegen eine Aktivierung von p97.

Die amyotrophe Lateralsklerose (ALS, auch als Lou-Gehrig-Syndrom bezeichnet) ist eine degenerative Erkrankung des motorischen Nervensystems, die durch den britischen Astrophysiker Stephan Hawking als prominentesten Patienten öffentliche Aufmerksamkeit erhielt [35]. Hierbei gehen Motorneuronen im Gehirn und im Rückenmark zugrunde, mit der Folge einer immer weiter fortschreitenden Funktionseinschränkung der Muskulatur, welche letztlich zum Tod führt. Bei 10% dieser Krankheit spricht man von familiärer ALS (FALS), bei der mehrere Mitglieder einer Familie von dieser Krankheit betroffen sind. 18 unterschiedliche genetische Mutationen wurden mit FALS Erkrankungen in Verbindung gebracht, von denen 1-2% aller Fälle auf p97-Mutationen zurückzuführen sind [36].

Auch die Charcot-Marie-Tooth-Erkrankung Typ 2Y (CMP2Y), eine hereditäre Neuropathie, bei der die Unterschenkelmuskulatur schwindet, steht in Zusammenhang mit p97. Eine Subpopulation der Patienten zeigt hier ebenfalls Mutationen in p97 [36, 37].

2.1.4.2 Erkrankungen, die mit nativem p97 in Verbindung stehen

Das UPS steht schon lange im Fokus im Zusammenhang mit Krebserkrankungen. Krebszellen zeigen, durch genetische oder epigenetische Veränderungen eine unausgeglichene Proteinsynthese und eine Anreicherung an fehl gefalteten Proteinen. Diese Zellen sind daher massiv auf ein funktionierendes UPS angewiesen, um proteotoxischen Stress zu vermeiden. Zahlreiche pharmakologische Angriffspunkte innerhalb des UPS stehen daher im Fokus für neue therapeutischer Ansätze gegen Krebszellen [38]. p97 als Teil des UPS stellt hier ein neues mögliches Target für die Entwicklung neuer Therapieansätze dar. Die pharmakologische Inhibierung von p97 führte *in cellulo* zum Zelltod unter anderem von soliden Tumoren von HCT116-Zellen (Kolonkarzinom) und A549-Zellen (Lungenkarzinom) [39], wie auch in B-ALL Zelllinien (B-Zell Akute lymphatische Leukämie) [40]. Des Weiteren würde p97 einen höchst interessanten Ansatz zur Therapie von

Krebserkrankungen bieten, die Resistenzen gegen Proteasom-Inhibitoren aufweisen [41].

2.1.4.3 p97 und virale Infektionen

Schlussendlich wird p97 auch als mögliches Target bei viralen Infektionen in Erwägung gezogen, was vor dem Hintergrund der 2020 ausgebrochene Covid-19 Pandemie von besonderem Interesse sein könnte. Auf Basis von siRNA-Screens wurde p97 als Wirts-Protein für die Replikation von Polioviren, humanen Cytomegalieviren [36], wie auch von Coronaviren [42] (zu denen auch SARS-CoV-2 zählt) identifiziert. Eine Inhibierung von p97 in Wirtszellen könnte somit ein interessanter Ansatz sein, um in den Infektionsverlauf dieser Viren einzugreifen.

2.1.5 Derzeit bekannte p97-Inhibitoren

Die beim Verfassen dieser Arbeit Ende 2020 bekannten p97-Inhibitoren lassen sich anhand ihres Wirkmechanismus, in die folgenden drei Kategorien einteilen:

- ATPase-Inhibitoren, die entweder selektiv eine oder beide ATPasen der D1- und/oder D2-Domäne kompetitiv inhibieren;
- allosterische Inhibitoren, die die ATPase-Funktion mit einem nicht-kompetitiven Mechanismus inhibieren;
- und Inhibitoren mit unbekanntem Wirkmechanismus.

Im Folgenden sollen die drei Klassen kurz dargestellt werden, wobei das Hauptaugenmerk auf deren Angriffspunkten an p97 liegen und gleichzeitig eine Zusammenfassung aller in der Literatur bekannten Verbindungen gegeben werden soll. Eine detaillierte Darstellung der einzelnen Inhibitoren, mit ihren *in vitro* und *in cellulo* hervorgerufenen Effekten, würde den Rahmen sprengen. Hierzu sei auf die Übersichtsartikel [36, 43, 44] verwiesen, die jeweils einen Teil der hier aufgelisteten Moleküle abdecken. Eine Übersicht aller Inhibitoren und deren Strukturformen ist in Tabelle 2.1 gegeben.

2.1.5.1 ATPase-Inhibitoren

Die ATPase-Inhibitoren stellen die größte Klasse der bekannten Inhibitoren dar. Die 2010 in einem Hochdurchsatz-Screening (*High-Troughput Screening*, HTS) gefundenen, ein 2-Anilino-4-Aryl-1,3-Thiazol als Leitstruktur enthaltenden ATPase-Inhibitoren stellen die ersten bekannten p97-Inhibitoren dar, mit IC₅₀

Werten bis zu 110 nM [45]. Kurz darauf wurde mit dem Chinazolinderivat N²,N⁴-dibenzylchinazolin-2,4-diamin (DBeQ) ein weiterer, die D1- und D2-ATPase-Funktion kompetitiv angreifender Inhibitor gefunden [46]. DBeQ zeigte eine Inhibition des ERAD Systems, eine Aktivierung des UPR sowie eine Beeinträchtigung des Proteinabbaus über Autophagie. Von DBeQ abgeleitet wurden die beiden an der D2-ATPase angreifenden Inhibitoren ML240 und ML241 entwickelt, die eine bessere Spezifität und Affinität gegenüber p97 aufweisen [47]. Auf ML240 und ML241 aufbauend wurde der D2-ATPase-Inhibitor CB-5083 entwickelt (siehe auch Abbildung 2.5). Dieser zeigte eine beeindruckende antitumorale Wirkung gegen Lungenkarzinome, kolorektale Karzinome und andere Krebsarten und ging daraufhin als erster p97-Inhibitor in die klinische Phase-I, in der er allerdings aufgrund nicht ausreichender Selektivität, durch Hemmung der Phosphodiesterase PDE6 (deren Hemmung zu Blindheit führt) scheiterte [39, 41, 44]. Aktuell führt das US-Unternehmen Cleave Therapeutics, das sich auf die Entwicklung von p97 Inhibitoren fokussierte hat, die klinische Phase-I des auf Grundlage von CB-5083 weiterentwickelten Inhibitors CB-5339 durch, der gegenüber PDE6 eine 40-fach höhere Selektivität für p97 besitzt [48].

Neben den eben vorgestellten nicht-kovalenten ATPase-Inhibitoren wurden auch solche mit einem kovalenten Mechanismus gefunden. Dabei nutzen alle die Nähe des reaktiven Cysteins C522 zur Nukleotid Bindestelle in der D2-Domäne aus. Syk Inhibitor III (3,4-Methylendioxy- β -nitrostyren) bindet kovalent an C522 und inhibiert somit die Funktion der ATPase. Gefunden wurde Syk Inhibitor III in einem zellbasierten Proteinabbau-Assay, bei dem 160 Proteinkinase-Inhibitoren auf deren Einfluss auf das UPS getestet wurden [49]. Die kovalenten Inhibitoren NMS-859 [50], Withaferin A [50], Δ (10,11)-Dehydrocurvularin [51] und LC-1028 [52] inhibieren ebenfalls die ATPase-Funktion der D2-Domäne durch kovalente Modifikation des C522. Hierbei sollte LC-1028 insofern ein besonderes Interesse geschenkt werden, als dieser (nachgewiesen in LC-MS/MS Analysen) ebenfalls das Cystein C105 in der N-Domäne kovalent modifizierte.

2.1.5.2 Allosterische Inhibitoren

Alle berichteten allosterischen Inhibitoren hemmen die ATPase-Funktion der D1- und/oder der D2-Domäne und führen aufgrund dessen zu einem Funktionsverlust von p97. Eine interessante allosterische Bindestelle liegt hierbei im Bereich zwischen die D1- und D2-Domäne. An diese Bindestelle greifen das Triazolderivat NMS-873 (und Analoga) [50, 53], das Phenylindol UPCDC30245

(und Analoga) [54–56], sowie das Tetrahydrocarbazol MSC1094308 [57] (und Analoga) an und führen zu einer nicht kompetitiven Hemmung der p97-ATPase-Funktion. Von UPCDC30245 konnte von *Banerjee et al.* eine hochaufgelöste (bis zu 2.3 Å) Cryo-EM Struktur erhalten werden, die den Inhibitor im Komplex mit p97 zeigt [13] (siehe Abbildung 2.5). Eine weitere allosterische Bindestelle liegt in der D2-Domäne und ist Angriffspunkt einer Serie von Indolamid-Inhibitoren, die die ATPase-Funktion der D2-Domäne nicht-kompetitiv hemmen [58].

2.1.5.3 Inhibitoren mit unbekanntem Wirkmechanismus

Eeyarestatin I zeigt eine p97 inhibierende Wirkung, die nicht auf der Hemmung der ATPase-Funktion zu beruhen scheint, sondern wahrscheinlich einen Einfluss auf die Konformation von p97 hat. Die Eeyarestatin I Bindestelle liegt in der D1-Domäne von p97 und die Bindung beruht wahrscheinlich auf einem kovalenten Mechanismus [59].

Ähnliches gilt für die beiden Naturstoffe 1-Hydroxydehydroherbarin und Phomapyrrolidon A, deren Wirkmechanismus nicht ATP-sensitiv ist und deren Bindestellen nicht weiter bekannt sind [43]. Das Azol-Antimykotikum Clotrimazol ist ebenfalls ein schwacher Inhibitor von p97, die genaue Bindestelle ist unbekannt, wird aber zwischen der D1- und D2-Domäne vermutet [60]. Des Weiteren zeigen Oxaspirol und 5'-I-Fuligocandin B inhibitorische Effekte auf p97 mit unbekanntem Wirkmechanismus [36]. In diese Klasse gehört auch der Naturstoff Xanthohumol, für den als einzigen eine mögliche Bindestelle in der N-Domäne von p97 berichtet wird und der somit für diese Arbeit von besonderem Interesse ist. Auf diesen soll später noch detaillierter eingegangen werden.

2.1.5.4 Fragmentscreening an p97-N und p97-ND1

Neben den eben vorgestellten p97-Inhibitoren, wird in der Literatur bei *Chimenti et al.* [61] ein mit p97 durchgeführtes Fragmentscreening berichtet. 2485 Fragmente wurden dort mit einem Oberflächenplasmonenresonanz-Spektroskopie Assay auf ihre Bindung an die N- bzw. ND1-Domäne hin untersucht. Der Fokus lag auf der Nukleotidbindestelle in der D1-Domäne. Für die isolierte N-Domäne wurden sechs Fragmente identifiziert, die eine Bindung mit Affinitäten zwischen 241 und 1960 μM zeigten.

2.1.5.5 Das Chalkon Xanthohumol als Inhibitor von p97

Der Naturstoff Xanthohumol (XN) findet sich in den Zapfen weiblicher *Homulus lupulus* (Hopfen) Pflanzen. Chemisch gesehen handelt es sich bei XN um ein offenkettiges prenyliertes Flavonoid, dass zu den Chalkonen gezählt wird. Die beiden aromatischen Ringe A und B sind durch eine α , β -ungesättigte Carbonylfunktion getrennt (siehe Abbildung 2.4). XN stellt mit 90% die Hauptverbindung prenylierter Chalkone im Hopfen dar und ist mit einem Massenanteil von 0.1 bis 1.0% in getrocknetem Hopfen zu finden [62].

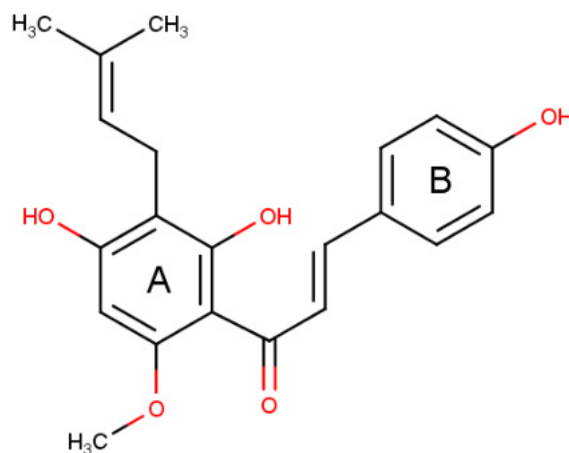


Abbildung 2.4: Struktur von Xanthohumol

Es wird fast ausschließlich durch Isolierung aus Hopfen gewonnen. Die Wasserlöslichkeit ist mit nur 0.0013 g/L schlecht. Durch das Michael-System zyklisiert XN, begünstigt durch Hitze sowie pH-Werte unter 6 oder über 8, meist zu Isoxanthohumol [63]. Mit einem Molekulargewicht von 354 Dalton, fünf Wasserstoffbrückenakzeptoren, drei Wasserstoffbrückendonatoren und einem $\log D^{pH 7.4}$ von 4.65 entspricht XN der *Rule of 5* von Lipsinki (Definition siehe Tabelle 2.2) [64].

Ein breites pharmakologisches Wirkspektrum wird XN zugeschrieben, dem neben leberprotektiven, antimikrobielle und antiinflammatorische Effekte, vor allem anticancerogene Aktivitäten attestiert werden [65, 66]. In *in vitro* Studien wurden die Effekte von XN in Krebszellen untersucht. Diese zeigten eine inhibierende Wirkung auf die Proliferation, Migration und Invasion von Krebszellen, sowie Apoptose und Zellzyklus-Stillstand induzierende Effekte. Daneben soll es positive synergistische Effekte auf Chemotherapien und Bestrahlungstherapien besitzen. Eine Vielzahl an möglichen zellulären Zielproteinen und biochemischen Stoffwechselwegen werden in diesem Kontext diskutiert [67]. In der

ChEMBL Datenbank für bioaktive Moleküle mit wirkstoffartigen Eigenschaften [68] werden 16 Zielproteine für XN berichtet. Hierzu zählen unter anderem die mitogen-aktivierenden Protein-Kinasen ERK1/ERK2, die NAD(P)H-Dehydrogenase, die MAP Kinase p38 und die Caspase-3, zu denen teilweise Affinitätsdaten vorliegen. p97 wird hier nicht als Zielprotein aufgeführt.

Sasazawa et al. [69] postulieren, dass XN an die N-Domäne von p97 bindet und in den Prozess der Autophagosomen-Reifung eingreift. Es konnte gezeigt werden, dass die Fusion von Autophagosomen mit Lysosomen durch XN inhibiert wird. Durch Inkubation von Zell-Lysat mit an Agarosepartikeln immobilisiertem XN konnte durch anschließende Analyse mittels MALDI-TOF-MS bzw. LC-MS/MS p97 als Zielprotein identifiziert werden. Rekombinantes p97 zeigte in einem Kompetitionsassay mit immobilisiertem und freiem XN einen Dosis-abhängigen Effekt in Westernblots.

Zur Eingrenzung der Bindestelle wurden drei Konstrukte von p97 hergestellt: p97 Δ N, p97 Δ D1 und p97 Δ D2. Hierbei zeigte nur das p97 Δ N-Konstrukt eine deutlich schwächere Bande im Westernblot, wodurch die Autoren zu der Schlussfolgerung kamen, dass die Bindestelle von XN auf der N-Domäne von p97 zu finden sein muss [69]. Wie oben beschrieben spielt p97 eine entscheidende Rolle im ERAD-System der Zelle. Ein Funktionsverlust oder Inhibierung der Bindung von ERAD-assoziierten Kofaktoren an p97 durch XN würde somit zu einem Anstieg an ungefalteten Proteinen und zur Induktion von ER-Stress in der Zelle führen. *Sasazawa et al.* konnten durch den signifikanten Anstieg der ER-Stress assoziierten Proteinmarker CHOP und GRP78 genau dieses für XN nachweisen. Bis zum Zeitpunkt der vorliegenden Arbeit ist keine auf diese Ergebnisse aufbauende Publikation erschienen. Da XN mittels der in [70] beschriebenen Methode an Agarose-Beads immobilisiert wurde, liegt XN ohne definierte Orientierung auf den Agarose-Beads vor. Anhand der Immobilisierung kann daher nicht auf ein Bindeepitop von XN geschlossen werden. Die Frage nach der genauen Bindestelle und Affinität von XN bleibt unbeantwortet.

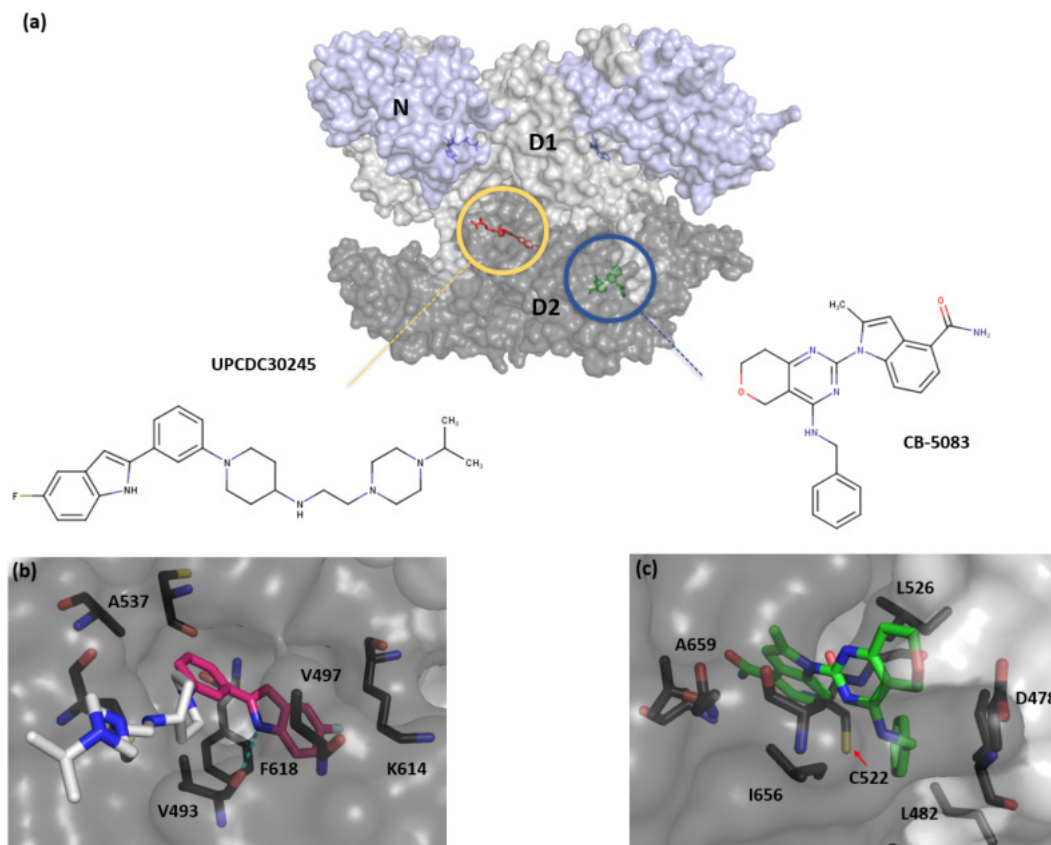


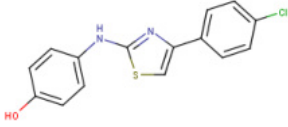
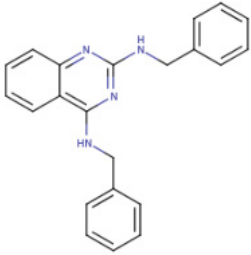
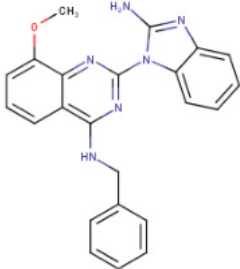
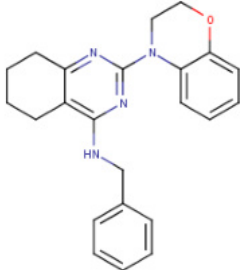
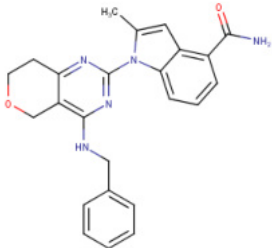
Abbildung 2.5: Strukturen der Inhibitoren CB-5083 und UPCDC30245 im Komplex mit p97

(a) Zu sehen sind die Bindestellen von UPCDC30245 (rot) zwischen der D1- und D2-Domäne, sowie von CB-5083 (grün) in der ATP-Bindestelle der D2-Domäne.

(b) In die Bindung von UPCDC30245 involvierte Aminosäuren. Der weiße Bereich des Inhibitors ist in der Cryo-EM Struktur (PDB: 5FTJ) nicht aufgelöst. Die Bindung von UPCDC30245 beruht im rot markierten Bereich größtenteils auf hydrophoben Interaktionen. In Blau ist eine Wasserstoffbrücke zwischen dem Indolring und der Hauptkette von V493 zusehen.

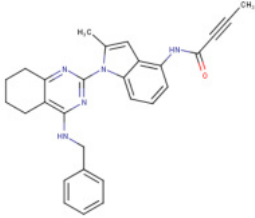
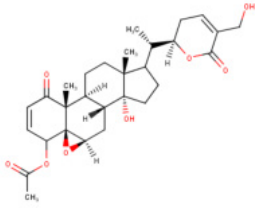
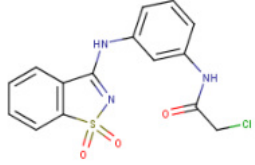
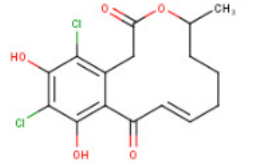
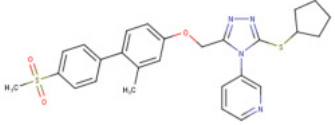
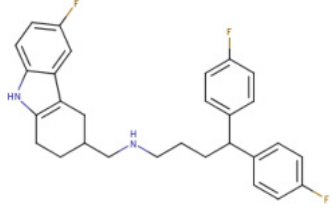
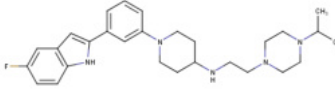
(c) Detailansicht von CB-5083 in der ATP-Bindestelle der D2-Domäne (PDB: 6MCK). CB-5083 war der erste p97-Inhibitor der bis in die klinische Phase-I gelang, dort allerdings scheiterte. Im Hintergrund ist C522 zu sehen, an dem eine Reihe von kovalenten Inhibitoren angreifen.

Tabelle 2.1: Übersicht der in der Literatur bekannten Inhibitoren von p97

Struktur	ID	IC ₅₀	Mechanismus
	4-[4-(para-Chlorophenyl)-1,3-thiazol-2-]aminophenol [45]	70 nM	ATP kompetitiv, via D1 und D2
	DBeQ [46]	<10 µM	ATP kompetitiv, via D1 und D2
	ML240 [47]	110 nM	ATP kompetitiv, via D2
	ML241 [47]	110 nM	ATP kompetitiv, via D2
	CB-5083 [44]	11 nM	ATP kompetitiv, via D2

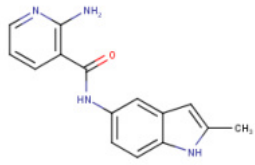
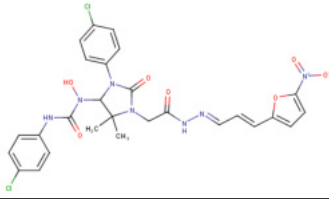
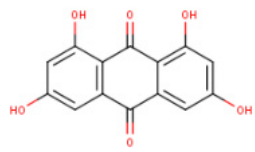
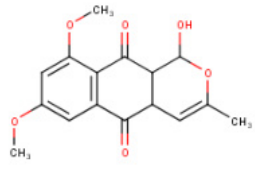
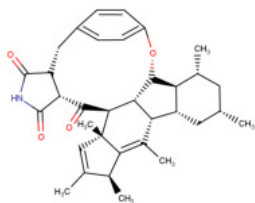
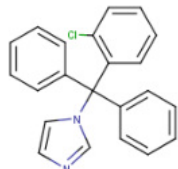
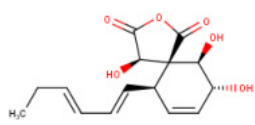
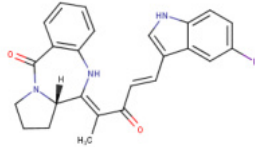
Fortsetzung auf nächster Seite

Tabelle 2.1 – Fortsetzung von letzter Seite

Struktur	Name, Bezeichnung	IC ₅₀	Mechanismus
	LC-1028 [52]	33 nM	ATP kompetitiv, via D2 (kovalent an C522)
	Withaferin A [50]	22 µM	ATP kompetitiv, via D2 (kovalent an C522)
	NMS859 [50]	370 nM	ATP kompetitiv, via D2 (kovalent an C522)
	Δ(10,11)-Dehydrocurvularin [51]	15 µM	ATP kompetitiv, via D2 (kovalent an C522)
	NMS873 [53]	24 nM	Allosterisch D1-D2
	MSC1094308 [57]	7.2 µM	Allosterisch D1-D2
	UPCDC30245 [54–56]	20 nM	Allosterisch D1-D2

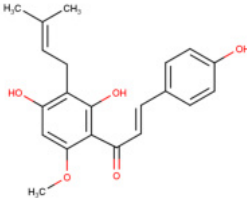
Fortsetzung auf nächster Seite

Tabelle 2.1 – Fortsetzung von letzter Seite

Struktur	Name, Bezeichnung	IC ₅₀	Mechanismus
	2-Amino-N-(2-methyl-indol-5-yl)pyridine-3-carboxamid [58]	0.5 µM	Allosterisch D2
	Eeyarestatin I [59]	5-10 µM	nicht bekannt
	Rheoemodin [43]	39.8 µM	ATP sensitiv, D1 selektiv
	1-Hydroxydehydroherbarin [43]	21.7 µM	Allosterisch
	Phomapyrrolidon A [43]	6.6 µM	Allosterisch
	Clotrimazol [60]	12 µM	nicht bekannt
	Oxaspirol B [36]	31 µM	nicht bekannt
	5'-I-Fuligocandin [36]	7 µM	nicht bekannt

Fortsetzung auf nächster Seite

Tabelle 2.1 – Fortsetzung von letzter Seite

Struktur	Name, Bezeichnung	IC ₅₀	Mechanismus
	Xanthohumol [69]	-	nicht bekannt

2.1.6 Protein-Protein-Interaktionen

Wechselwirkungen zwischen Proteinen spielen in den unterschiedlichsten Funktionen der Zelle eine entscheidende Rolle. Zahlreiche solcher Protein-Protein-Interaktionen (PPI) wurden als Zielstrukturen für unterschiedliche Krankheiten identifiziert. Dennoch spielen niedermolekulare Verbindungen, die die direkte Bindung zwischen Proteinen beeinflussen, bisher nur eine untergeordnete Rolle bei bekannten Wirkstoffen. PPI gelten als eine schwierig zu adressierende Zielstruktur im Wirkstoffdesign und werden teilweise als für Wirkstoffe nicht zugänglich (*undruggable*) klassifiziert [71]. Zum Verständnis der Interaktionen zwischen Proteinen hat sich die *hot-spot*-Theorie etabliert [72]. Die für die Interaktion wichtigen Aminosäuren sind hier nicht gleichmäßig über die bindende Region verteilt, sondern lokalisieren sich in „heißen“ Bereichen [73]. Die Aminosäuren solcher *hot-spots* sind hauptsächlich an der Bindung der Proteine beteiligt. Experimentell können solche *hot-spot*-Aminosäuren über sogenannte Alanin-Scanning-Mutagenesen identifiziert werden [72]. Hierbei werden systematisch die an der Bindung beteiligten Aminosäuren durch Alanin oder Glycin ausgetauscht, um deren Beitrag zur Bindung zu bestimmen. Es wurde beobachtet, dass sich die Zusammensetzung solcher *hot-spots* sich von „normalen“ Lösungsmittel exponierten Stellen unterscheidet. So findet man in diesen Bereichen eine höhere Anzahl der Aminosäuren Tryptophan, Arginin, Tyrosin, Histidin und teilweise auch Phenylalanin und Methionin [71, 74]. PPI werden neben Wechselwirkungen von polaren Aminosäuren vor allem durch hydrophobe Interaktionen bestimmt. Die an PPI beteiligten Proteinoberflächen besitzen typischerweise eine Fläche in einem Größenbereich von 1000 bis 6000 Å² [75]. Anhand ihrer Strukturmerkmale lassen sich PPI in drei groben Klassen einteilen [76]:

- Interaktionen zwischen globulären Proteinen;
- einzelne Peptidketten, die mit globulären Proteinen wechselwirken;

- Interaktionen zwischen einzelnen Peptidketten.

Hierbei wird die Klasse, bei der eine Peptidkette mit einem globulären Protein wechselwirkt, als die für das Wirkstoffdesign interessanteste Klasse angesehen [76]. Im Kontext der Interaktionen an der N-Domäne von p97 fällt die Bindung des SHP-Motivs an die Nc-Subdomäne in diese Strukturklasse. Zur Entwicklung von PPI-Inhibitoren werden unterschiedliche Strategien in der Literatur berichtet [71]: konventionelles HTS (das sich aber als nicht zielführend zur Entwicklung von PPI-Inhibitoren herausgestellt hat), Peptid-Mimetika (z.B. zur Imitation von α - oder β -Strukturen) und Fragmentbasiertes Wirkstoffdesign, das sich als äußerst effektiv in der Entwicklung von PPI-Inhibitoren erwiesen hat.

2.1.7 Die N-Domäne als mögliches Ziel für p97-Inhibitoren

Sieht man von dem an die N-Domäne bindenden Naturstoff Xanthumol ab, so adressieren alle bekannten Inhibitoren die D1- bzw. D2-Domäne, mit einem Fokus auf die ATPase-Funktion von p97 (siehe 2.1.5). Ein essentielles Problem der direkt kompetitiven ATPase-Inhibitoren ist eine entsprechende Selektivität für p97 zu erreichen. Durch das ubiquitäre Vorkommen von ATPasen ist stets das Problem gegeben, dass neben p97 auch weitere ATPasen gehemmt werden können.

Auch wenn die allosterisch angreifenden Inhibitoren diese Problematik nicht in gleichem Ausmaß aufweisen, so inhibieren sie die Funktion von p97 als Ganzes. Vor dem Hintergrund der vielen Prozesse in dem p97 eine Rolle spielt, kann dies ebenfalls schnell zu nicht gewollten Wirkungen in der Zelle führen.

Da alle Funktionen von p97 über die Bindung verschiedener Kofaktoren vermittelt werden, wäre ein selektiverer Ansatz sicherlich die Inhibierung der Bindung einzelner Kofaktoren. Somit könnten im günstigsten Fall einzelne Signalwege von p97 „ausgeschaltet“ werden, während andere Funktionen von p97 erhalten bleiben. Ein solcher Ansatz wäre nicht nur hinsichtlich neuer „*tool compounds*“ für die weitere Aufklärung von p97-Funktionen interessant, sondern möglicherweise auch ein neuer Ansatz um selektivere, in ihrem Wirkungsspektrum begrenztere neue Wirkstoffkandidaten zu erhalten. Hier ist sicherlich die N-Domäne von p97 mit ihren zwei Binderegionen und der Vielzahl an bindenden Kofaktoren die erste Wahl für solchen PPI-Inhibitoren. Aber auch hier besteht die Problematik der Selektivität. Die oben bereits erwähnte Vielzahl an Kofaktoren binden an die gleichen Stellen in der N-Domäne, sodass eine spezifische Selektivität für einen Kofaktor sich sicherlich nur bedingt erreichen lässt.

Trotz allem wären PPI-Inhibitoren, die selektiv eine der beiden Binderegionen der N-Domäne blockieren, eine bedeutende Neuerung im Vergleich zu den bekannten p97-Inhibitoren. Solche PPI-Inhibitoren wurden bisher in der Literatur nicht berichtet und würden ein *Novum* darstellen. Als Zielstruktur bietet sich hier sicherlich die Bindestelle des SHP-Motivs an, die zur der strukturellen PPI-Klasse gehört, die am interessantesten für die Entwicklung von wirkstoffähnlichen Molekülen angesehen wird und die größten Kavitäten in der N-Domäne zeigt.

2.2 Computergestützte Methoden in der Wirkstoffentwicklung

Dem Computer-basierten Wirkstoffdesign (*Computer-aided drug design*, CADD) fällt eine immer größere Rolle bei der Findung neuer Wirkstoffe in der pharmazeutischen Industrie und im akademischen Bereich zu. Vor allem im Frühstadium der Suche nach neuen Wirkstoffen (wie z.B. der Identifizierung erster Treffer durch virtuelle Screenings oder der Optimierung von Leitstrukturen) sind die Methoden zu einem unersetzbaren Werkzeug geworden. Mit dem CADD ist es möglich, Millionen von Molekülen auf mögliche Liganden hin zu untersuchen und diese schnell auf einer rationalen Basis zu optimieren [77]. CADD-Methoden spielen nicht nur in der Suche nach neuen Leitstrukturen eine Rolle, sondern werden zunehmend auch in vielen anderen Phasen der Wirkstoffentwicklung eingesetzt, wie in der Vorhersage pharmakologischer Profile im Hinblick auf deren Toxizitäten, ADME (Absorption, Distribution, Metabolismus und Elimination) Eigenschaften oder Selektivität (z.B. Polypharmazie) [78]. Hierbei bietet das CADD die Möglichkeit der Beschleunigung, Kostenreduzierung und Minimierung von Fehlschlägen in der Wirkstoffentwicklung [79]. Die zunehmende Größe von biologischen und chemischen Datenbanken, sowie immer leistungsstärkere Rechner und Grafikkarten beschleunigen diesen Prozess. Nicht zuletzt erhält das CADD durch den Einzug an Methoden aus dem Bereich der Künstlichen Intelligenz aktuell einen neuen Auftrieb [80]. Im Folgenden werden die für diese Arbeit wesentlichen Aspekte und Methoden behandelt.

2.2.1 Molekulare Deskriptoren zur Charakterisierung von Molekülbibliotheken

Der Begriff Chemoinformatik wurde 1996 erstmals von Frank Brown benutzt, und bezeichnet einen (nicht immer scharf abgrenzbaren) Bereich der computerbasierten Methoden, mit denen es möglich ist, eine große Anzahl an Molekülen schnell zu charakterisieren [81]. Neben der Verarbeitung dreidimensionaler Molekülstrukturen fallen hierunter auch die Analyse und Verarbeitung von ein- und zweidimensionalen Moleküldeskriptoren. Gerade lineare Notationen (wie der SMILES-Code) und die Darstellung von Molekülen in sogenannten *bit strings* oder *fingerprints* machen es möglich, Bibliotheken von chemischen Verbindungen schnell auf verschiedene Eigenschaften hin zu untersuchen (wie beispielsweise in Ähnlichkeitsanalysen oder Substruktursuchen). Daneben bietet die Chemoinformatik die Möglichkeit, Moleküle anhand von sogenannten Deskriptoren

zu charakterisieren. Molekulare Deskriptoren wandeln hierbei die auf der Struktur basierende chemische Information in numerische Werte um. Hierzu gehören einfache beschreibende Deskriptoren, wie das Molekulargewicht, die Anzahl an Schweratomen oder die Anzahl rotierbarer Bindungen. Daneben existieren physikochemische Deskriptoren, wie beispielsweise der n-Octanol/Wasser Verteilungskoeffizient $\log P$, welcher die Lipophilie eines Moleküls beschreibt, oder die topologische polare Oberfläche (*topological polar surface area*, TPSA¹). Zur Vorhersage des $\log P$, ist der bei *Rekker et al.* [83] berichtete Ansatz der bekannteste. Bei dieser Methode wird ein Molekül in Fragmente zerlegt und diese Fragmente experimentell bekannten $\log P$ -Werten zugeordnet. Die Summe der $\log P$ -Werte der einzelnen Fragmente bildet darauf den berechneten $\log P$ ($clogP$). Ein bekanntes Ähnlichkeitsmaß zweier Moleküle ist der Tanimoto-Koeffizient. Hierbei werden für jedes Molekül sogenannte „Fingerabdrücke“ erzeugt, indem die Struktur der Moleküle in Bitfolgen überführt und verglichen wird. Der Koeffizient gibt die Größe der gemeinsamen Schnittmenge beider Moleküle an und kann Werte zwischen 0 und 1 annehmen (0: Moleküle besitzen keine Schnittmenge; 1: die Bitfolgen sind identisch) [84]. Eine Übersicht über weitere Deskriptoren ist in [85] zu finden. Basierend auf molekularen Deskriptoren, lässt sich ein *chemischer Raum* definieren. Hierbei besteht mitunter die Problematik einen n-dimensionalen Raum (wobei n die Anzahl verwendeter Deskriptoren ist) auf einen visuell anschaulichen Raum zu reduzieren [81]. Zur Reduktion der Dimensionalität bedient man sich häufig der Hauptkomponenten-Analyse (*Principal Component Analysis*, PCA), wenn eine Korrelation einzelner Variablen vermutet werden kann. Hierbei wird angenommen, dass ein Großteil der Variabilität eines Datensatzes mit Hilfe von wenigen, linear-kombinierten Variablen beschrieben werden kann. Allgemein lässt sich ein mehrdimensionale Datensatz durch einen Satz von Linearkombinationen darstellen:

$$\begin{aligned}
 PC_1 &= c_{1,1}x_1 + c_{1,2}x_2 + c_{1,p}x_p + \dots \\
 PC_2 &= c_{2,1}x_1 + c_{2,2}x_2 + c_{2,p}x_p + \dots \\
 &\dots \\
 PC_i &= c_{i,1}x_1 + c_{i,2}x_2 + c_{i,p}x_p + \dots
 \end{aligned} \tag{2.1}$$

Hierbei ist c der Koeffizient des Deskriptors x und p die Anzahl an Deskriptoren. Die Hauptkomponenten werden nun durch die Varianz-Covarianz-Matrix CC^T

¹Bei der TPSA wird nicht wie bei der PSA (*polar surface area*) die polare Fläche über eine generierte dreidimensionale Darstellung des Moleküls errechnet, sondern die Summe einzelner bereits berechneter PSA Anteile polarer Fragmente des Moleküls gebildet [82]

errechnet, mit n Zeilen (entsprechend der Anzahl an Molekülen) und m Spalten für die Anzahl an Deskriptoren, durch $n \times m$. Die Eigenvektoren dieser Matrix sind hierbei die Koeffizienten c , wobei die erste Hauptkomponente die meiste Varianz im Datensatz abdeckt, die zweite Hauptkomponente die zweithöchste Varianz usw. [86].

2.2.2 Pharmakophormodelle

Monty Kier veröffentlichte 1967 bis 1971 Studien der Wechselwirkungen von Muscarinrezeptor-Agonisten, in denen der Weg von Pharmakophormodellen bereitet wurde [87]. Die heutige Definition eines Pharmakors lautet [88]: *„Ein Pharmakophor ist ein Ensemble von sterischen und elektrostatischen Eigenschaften, die für eine optimale supramolekulare Interaktion nötig sind, um bei einer spezifischen biologischen Zielstruktur eine biologische Antwort hervorzurufen.“* Die Darstellung von Protein-Ligand-Wechselwirkungen auf Basis der dreidimensionalen Anordnung von gefundenen funktionalen Gruppen (wie Wasserstoffbrückendonatoren bzw. -akzeptoren, hydrophobe Gruppen etc.), die für die Interaktion wichtig sind, bietet die Möglichkeit Molekülbibliotheken effektiv auf Verbindungen mit ähnlichen Eigenschaften hin zu durchsuchen. 3D Pharmakophormodelle kommen bei einer Vielzahl von Programmen zum Einsatz und werden bevorzugt im frühen Stadium der Wirkstoffentwicklung eingesetzt [89]. Generell unterscheidet man zwischen struktur- und ligandbasierten Pharmakophormodellen.

Strukturbasierte Modelle bauen auf der Analyse von Kristallstrukturen auf. Im günstigsten Fall liegt ein Protein-Ligand-Komplex vor, aus dem die wesentlichen Interaktionen abgeleitet und in ein Modell überführt werden können. Für die Adressierung neuer Zielstrukturen ohne einen bekannten Protein-Ligand-Komplex ist die Generierung von Pharmakophormodellen deutlich schwieriger. Eine genaue Analyse der physikochemischen Eigenschaft der gewünschten Binderegion bildet hier die Grundlage des Modells. Hierbei kommt es auf die Identifizierung der wesentlichen Interaktionspunkte an, mit denen organische Moleküle interagieren könnte. Eine der gängigsten Methoden geht auf das von Peter Goodford entwickelte Programm GRID mit einem gitterbasierten Ansatz zurück [90]. An jedem Gitterpunkt wird mittels einer chemischen Sonde die Interaktionsenergie mit dem Protein berechnet. Die Berechnung der Energie erfolgt hierbei durch einen Kraftfeldansatz. Die in dieser Arbeit in MOE erstellten Interaktionspotentiale für verschiedene chemische Sonden greifen ebenfalls auf diese Methode zurück. An jedem Punkt x wird die Interaktionsenergie für eine gewählte

Sonde durch die folgende Gleichung berechnet:

$$E_x = \sum E_{ij} + \sum E_{ele} + \sum E_{hb} \quad (2.2)$$

wobei $\sum E_{ij}$ die Summe aller van-der-Waals Interaktionen, $\sum E_{ele}$ die Summe aller elektrostatischen Interaktionen und $\sum E_{hb}$ die Summe aller Wasserstoffbrücken Interaktionen, bedeuten [91]. Die so erhaltenen Potentiale für unterschiedliche Sonden liefern die Basis zur Erstellung eines Pharmakophormodells.

Ligandbasierte Pharmakophormodelle werden eingesetzt, wenn keine Kristallstruktur eines Protein-Ligand-Komplexes vorliegt, sondern nur eine Reihe von bioaktiven Molekülen bekannt ist. Hierbei wird versucht, durch Überlagerung der Moleküle diejenigen funktionellen Gruppen zu ermitteln, die allen bioaktiven Molekülen gemein sind und für die Bindung relevant zu sein scheinen. In dieser Arbeit wurde hierzu das *flexible Alignment* in MOE eingesetzt, um aus identifizierten Fragmenten mögliche gemeinsame Strukturelemente zu ermitteln. Hierbei werden funktionellen Gruppen, wie Wasserstoffdonatoren/-akzeptoren, aromatischen Systemen, hydrophobe Eigenschaften, durch dreidimensionale Gauß-Sphären dargestellt. Durch eine Konformationsuche wird daraufhin versucht, die vorher definierten funktionellen Gruppen zur maximalen Deckung zu bringen [92].

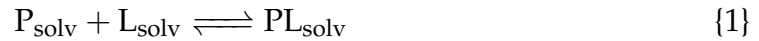
Mit den so erstellten Pharmakophormodellen können Molekülbibliotheken schnell auf Moleküle mit ähnlichen Eigenschaften durchsucht werden und neue Bibliotheken mit potentiell angereicherten bioaktiven Molekülen generiert werden [93].

2.2.3 Molekulares Docking

Das Verfahren des Dockings versucht zwei fundamentale Fragen im Wirkstoffdesign zu beantworten: Was ist der Bindemodus eines Ligand an ein Protein? Und wie ist die gefundene Pose zu bewerten bzw. was ist deren Bindungsenergie?

Ein Dockingverfahren besteht daher aus einer Suche nach unterschiedlichen Bindemodi eines Liganden und einem sich anschließenden Ranking der gefundenen Bindeposen auf Grundlage einer Scoringfunktion [94, 95]. Seit den ersten Anfängen Mitte der 1970er Jahre, wurden die Methode des Dockings stetig weiterentwickelt (nicht zuletzt durch die immer größere Rechenleistung von Computersystemen), so dass heute eine Vielzahl an verschiedensten Algorithmen und Methoden hierzu existieren [77]. Der Prozess der Bindung eines Liganden an dessen

Zielstruktur lässt sich aufgrund der Reaktionsgleichung {1} anhand der dazugehörigen Assoziations- bzw. Dissoziationskonstanten und der damit verbundenen Änderung der freien Enthalpie ΔG vereinfacht, wie folgt beschreiben:



$$K_A = 1/K_D = \frac{[PL]}{[P][L]} \quad (2.3)$$

$$\Delta G^0 = -RT \ln K_A = \Delta H^0 - T \Delta S^0 \quad (2.4)$$

Hierbei stellen $[P]$, $[L]$ und $[PL]$ die jeweiligen Konzentrationen des Proteins, des Liganden und des Komplexes dar. Die Assoziationskonstante K_A steht im direkten Zusammenhang mit der freien Enthalpieänderung ΔG^0 unter Standardbedingungen. Die freie Enthalpie ΔG^0 setzt sich nach der Gibbs-Helmholtz-Gleichung aus einem enthalpischen Term ΔH^0 und einem entropischen Term ΔS^0 zusammen (R Gaskonstante, T Temperatur). ΔH^0 stellt die Enthalpieänderung bei Bindung des Liganden dar, und setzt sich aus allen (bei Annahme eines nicht-kovalenten Mechanismus) nicht-kovalenten Wechselwirkungen und Konformationsänderungen zusammen. Der Entropieterm ΔS^0 beinhaltet die folgenden Terme:

$$\Delta S^0 = \Delta S_{\text{Solv}}^0 + \Delta S_{\text{Konf}}^0 + \Delta S_{r/t}^0 \quad (2.5)$$

ΔS_{Solv}^0 stellt den Entropieterm dar, der durch Solvatisierungs- bzw. Desolvatisierungsvorgänge, wie das Freisetzen von Wassermolekülen aus der Bindetasche, bestimmt wird. Der konformationsbedingte Entropieterm ΔS_{Konf}^0 beruht auf der Änderung der Konformationsfreiheitsgrade beider Bindungspartner. Der letzte Term $\Delta S_{r/t}^0$ ergibt sich aus der Änderung der Rotations- und Translationsfreiheitsgrade bei der Komplexbildung [96]. Erst wenn alle enthalpischen und entropischen Terme sich günstig aufsummieren, geht ein Ligand eine Bindung mit einem Protein ein.

Diese kurze thermodynamische Betrachtung zeigt, wie kompliziert und vielschichtig der Prozess der Ligandenbindung ist. Das Verfahren des molekularen Dockings stellt diesen Prozess in einer vereinfachten Weise dar, mit der Anforderung eines maximalen Durchsatzes an Molekülen, durch zeitsparende Rechenoperationen, wie auch einer möglichst genauen Beschreibung der Bindungsenergie. So besitzt im Screening nur der Ligand seine Flexibilität (in Form der generierten Bindeposen), während die Zielstruktur (meist ein Protein) als starr angenommen wird.

Um der biologischen Realität besser Rechnung zu tragen, wurde für die Seite des Proteins versucht, durch die Flexibilität einzelner Seitenketten [94] oder des Dockings mit Protein-Konformationsensembles [97], näher an der Realität liegende Modelle zu erstellen. Einflüsse eines expliziten Solvens, in dem sich die Reaktion normalerweise abspielt, fehlen dem Verfahren größtenteils ebenfalls. Auch hier versuchen verschiedene Ansätze, den nicht zu unterschätzenden Einfluss des Wassers durch das Einbeziehen einzelner Wassermoleküle [98] oder spezieller Wasserterme in den Scoringfunktionen [99] Rechnung zu tragen. Zusammenfassend lässt sich sagen, dass enthalpische Effekte durch die unterschiedlichen Scoringfunktionen leichter abgeschätzt werden können, als die entropischen Effekte der Ligandbindung.

2.2.3.1 Generierung von Bindemodi des Liganden

Wie oben erwähnt, bedarf es zunächst der Generierung von unterschiedlichen Bindeposen eines Liganden. Hierzu gibt es unterschiedliche Ansätze, wie die Generierung von externen Ligand-Konformerbibliotheken (z.B. OMEGA), die für *rigid* (starres) Docking herangezogen werden, eine Fragmentierung des Liganden mit anschließendem Wiederaufbau (FlexX), oder stochastische Methoden. Bei letzteren werden vor allem Monte Carlo oder genetische Algorithmen (GA) benutzt [95]. Das in dieser Arbeit verwendete Docking-Programm GOLD [100] bedient sich eines GA, auf den hier näher eingegangen werden soll.

Das Prinzip des GA lässt sich mit der Aussage des Sozialphilosophen Herbert Spencer: „*Survival of the fittest*“ beschreiben. Übertragen bedeutet dies, dass diejenigen Bindeposen eines Liganden weiterverfolgt werden, die am günstigsten bewertet werden. Inspiriert von der Evolutionstheorie wird am Anfang eine „Population“ von „Chromosomen“ (unterschiedliche Bindeposen) gebildet und auf sogenannten „Inseln“ verteilt. In diesen Chromosomen ist in „Genen“ die Information über einen Bindemodus codiert. Die Population wird anschließend einem evolutionärem Prozess unterzogen, um neue und bessere Bindeposen zu generieren. Im Detail enthält jedes Chromosom hierbei vier Bit-Strings auf denen unterschiedliche Informationen des Liganden bzw. des Proteins (Rotationswinkel für Hydroxyl- bzw. Amingruppen) gespeichert sind. Hierzu gehören auch Informationen über die Torsionswinkel des Liganden. Des Weiteren enthalten die Strings Kodierungen über Wasserstoffbrückenbindungen zwischen dem Liganden und dem Protein. Jede Population entwickelt sich nun unabhängig voneinander, wobei folgende Operationen auf die Chromosomen angewendet werden: zufällige Mutation bei einem String in einem Chromosom; Rekombination (*Cross-*

over), wobei ein Teil eines Strings zufällig mit einem String eines anderen Chromosoms vertauscht wird; und Migration, bei der ein ganzes Chromosom auf eine andere Insel übertragen wird. Die Anwendung dieser Operationen erfolgt zufällig. Der GA wird daraufhin für eine vom Benutzer festgelegte, oder anhand der rotierbaren Bindungen (Ligand) abgeleitete Anzahl an Generationen durchgeführt. Der vom Benutzer festgelegte „Selektionsdruck“ bildet zusammen mit dem ermittelten Wert der Scoringfunktion, die Wahrscheinlichkeit, mit der eine Bindepose in den nächsten Zyklus übernommen wird. Chromosomen, die einen hohen Score haben, haben somit eine höhere Wahrscheinlichkeit, in den nächsten Zyklus zu kommen, wo sie schlechtere Chromosomen ersetzen [101].

2.2.3.2 Scoringfunktionen

Eine Scoringfunktion übernimmt zwei Aufgaben: Zum einen wird sie beim Docking (wie im vorangegangenen Abschnitt zum GA beschrieben) zur Auswahl der besten Bindeposen eines Liganden gebraucht, zum anderen hat sie im Rahmen eines Screenings die Aufgabe, die Posen verschiedener Liganden untereinander zu beurteilen und zu reihen. Die Beurteilung einer Pose kann über deren Bindungsenergien beurteilt werden und sind über freie Energieberechnungen quantitativ zugänglich. Diese Verfahren sind allerdings viel zu rechenintensiv, um sie im Kontext eines virtuellen Screenings, bei dem tausende Protein-Ligand-Komplexe berechnet werden müssen, einzusetzen. Scoringfunktionen müssen also den oben beschriebenen komplexen Sachverhalt der Bindung eines Liganden massiv vereinfachen und durch rechnerisch besser zugängliche Modelle nachbilden.

Generell lassen sich Scoringfunktionen in drei verschiedene Kategorien einteilen: in empirische Scoringfunktionen, in wissensbasierte Scoringfunktionen und in kraftfeldbasierte Scoringfunktionen. Die Idee von Kraftfeldern wird im Zusammenhang mit Molekulardynamik Simulationen in Kapitel 2.2.4.1 beschrieben, da in dieser Arbeit nur Scoringfunktionen aus den ersten beiden Kategorien verwendet wurden [95, 102].

Der Idee von empirischen Scoringfunktionen liegt die Annahme zugrunde, dass sich die Bindungsaffinität eines Liganden durch die Summe von einzelnen individuellen Beiträgen (Terme) darstellen lässt [103]. Diese Terme werden anhand von Regressionsanalysen bekannter Protein-Ligand-Komplexe erhalten, wobei strukturelle Daten aus Röntgenstrukturanalysen, sowie experimentell bestimmte Affinitäten in die Modelle mit einfließen. Anhand eines so erstellten Modells ist

es möglich, die Bindungsaffinität einer unbekanntem Verbindung vorher zusa-
gen. Die empirische Funktion des Modells ist dabei stark von den verwendeten
Trainingsdatensätzen abhängig und wird nur solche Interaktionen ausreichend
beschreiben, die ähnlich denen sind, an denen sie trainiert wurde. In diese Ka-
tegorie fällt die in dieser Arbeit verwendete Scoringfunktion ChemPLP [104]. In
die Scoringfunktion ChemPLP fließen Terme der beiden Scoringfunktionen PLP
(*Piecewise Linear Potential*) [105] und ChemScore [106] ein. Die PLP-Terme be-
schreiben die sterische Komplementarität zwischen Protein und Ligand. Die Ter-
me der ChemScore-Funktion enthalten Beschreibungen für die winkelabhängi-
gen Wasserstoffbrückenbindungen und Bindungen an Metallatome. Zusätzlich
kommen Terme für das Torsionspotential aus dem Tripos-Kraftfeld [107] hin-
zu, um intramolekulare Wechselwirkungen des Liganden zu beschreiben. Jeder
Term bildet eine eigene Funktion mit unterschiedlichen Parametern, auf diese
hier nicht näher im Detail eingegangen wird.

Wissensbasierte Scoringfunktionen verfolgen dagegen einen anderen Ansatz und
beruhen ausschließlich auf den Strukturdaten von Protein-Ligand-Komplexen.
Hierbei werden statistische Potentiale (*mean force potentials*) erzeugt, die auf einer
statistischen Analyse der in Strukturen vorgefundenen Protein-Ligand-
Interaktionen beruhen. Beruhend auf dem Boltzmann-Gesetz lässt sich sagen,
dass Wechselwirkungen, welche häufig in Protein-Ligand Komplexen beobachtet
werden energetisch günstiger sind, als solche, die nicht häufig beobachtet wer-
den. Die Energie $E(i)$ lässt sich demnach über das inverse Boltzmann Gesetz
(Gleichung 2.6) berechnen, wobei k_B die Boltzmannkonstante, T die Temperatur
und $\rho(i)$ und ρ_{ref} , die vorgefundene bzw. Referenz-Verteilung einer Atompaar-
Interaktion i darstellen:

$$\Delta E(i) = -k_B T \ln \left(\frac{\rho(i)}{\rho_{ref}} \right) \quad (2.6)$$

Da es sich bei den berechneten Werten um keine wirklichen Energien handelt,
kann der lineare Term $k_B T$ weggelassen werden, sodass sich die allgemeine
Gleichung einer wissensbasierten Scoringfunktion für ein Atom i wie folgt ergibt:

$$Score(i) = -\ln \left(\frac{\rho(i)_{obs}}{\rho_{ref}} \right) \quad (2.7)$$

Der Gesamtscore einer solchen paarweisen distanzabhängigen Funktion ist somit
die Summe aller Einzelscores mit allen Ligandatomen l , Proteinatomen p und
deren Abständen r :

$$Score_{total} = \sum score(p, l, r) \quad (2.8)$$

Die in dieser Arbeit verwendete ASP-(*Astex Statistical Potential*)-Scoringfunktion [108] und DSX-Scoringfunktion [109] gehören dieser Kategorie an. Die der ASP-Scoringfunktion zugrunde gelegten Potentiale basieren auf den in der Protein Datenbank (*Protein Data Bank*, PDB [14]) hinterlegten Protein-Ligand-Komplexen, während bei der DSX-Scoringfunktion für die Atompaar-Potentiale zwischen zwei Potentialfunktionen gewählt werden kann: eine abgeleitet aus den Strukturen der Cambridge Strukturdatenbank (*Cambridge Structural Database*, CSD) [110] und eine abgeleitet aus den Strukturen der PDB. Potentiale für die Torsionswinkel des Liganden, sowie Potentiale für den Anteil der solvenszugänglichen Oberfläche eines Atoms, wurden bei der DSX-Scoringfunktion aus den Strukturen der PDB abgeleitet.

2.2.4 Molekulardynamik-Simulationen

„If we were to name the most powerful assumption of all, which leads one on and on in an attempt to understand life, it is that all things are made of atoms, and that everything that living things do can be understood in terms of the jiggings and wiggings of atoms.“ (Richard P. Feynman, *The Feynman Lectures on Physics* [111]). Dieses „Wackeln“ der Atome ist die Kernfrage, welche die Molekulardynamik (MD)-Simulation versucht zu beschreiben. Denn die reale Welt besteht nicht aus unbeweglichen Systemen, wie sie etwa in Docking-Verfahren zur Vereinfachung dargestellt werden, sondern aus sich zeitlich entwickelnden und gegenseitig beeinflussenden Prozessen. Die Bindung eines Liganden stellt keinen statischen Prozess wie sie im „Schlüssel-Schloss“-Prinzips [112] gedacht wird da, sondern ein dynamischer, in dem das Zusammenspiel von Konformationsänderungen des Liganden, des Proteins und der Wechselwirkung mit Wassermolekülen involviert sind. Die bestmögliche Beschreibung solcher auf atomarer Ebene ablaufenden dynamischen Prozesse der Bindung eines Wirkstoffes an sein Zielprotein würde eine quantenmechanische Beschreibung der sich zeitlich entwickelnden Bewegungen und Reaktionen von Atomen liefern. Solche Berechnungen liegen allerdings selbst mit den größten heutigen Supercomputern noch außerhalb dessen, was in einem sinnvollen Zeitrahmen auszuführen möglich wäre [113].

MD-Simulationen nutzen daher eine vereinfachte Beschreibung der Vorgänge auf Basis der Newtonschen Gesetze der Physik, um eine gute Näherung der real ablaufenden Prozesse zu liefern. Seit den Anfängen Mitte der 1970er Jahre mit der ersten Simulation eines Proteins hat sich die MD-Simulation zu einem riesigen Forschungsbereich in den unterschiedlichsten Teildisziplinen entwickelt und wurde mit dem Nobelpreis für Chemie im Jahr 2013 gewürdigt (Martin Karplus, Michael Levitt und Arieh Warshel für *"The development of multiscale models for complex chemical systems."* [114]).

Die theoretische Grundlage der MD-Simulation bilden die drei Newtonschen Axiome [111]:

1. Ein Körper bewegt sich so lange auf einer geraden Bahn mit konstanter Geschwindigkeit bis eine Kraft auf ihn wirkt.
2. Wirkt auf einen Körper eine Kraft, so wird er in Richtung der Kraft beschleunigt. Die Beschleunigung ist dabei direkt proportional zur Kraft und indirekt proportional zur Masse des Körpers.
3. Besteht zwischen zwei Körpern eine Kraftwirkung, so ist die Kraft, die Körper 1 auf Körper 2 ausübt, gleich der Kraft, die Körper 2 auf Körper 1 ausübt.

Die Trajektorie einer MD-Simulation ist hierbei die Lösung der Differentialgleichung des zweiten Axioms für diskrete Zeitschritte Δt (für eine vereinfachte Darstellung wird hier nur ein eindimensionaler Raum angenommen und nicht ein realer dreidimensionaler Raum):

$$\frac{d^2x_i}{dt^2} = \frac{F_{x_i}}{m_i} \quad (2.9)$$

Mithilfe dieser Gleichung lässt sich die Bewegung eines Atoms i mit der Masse m_i entlang einer Richtung x_i beschreiben, wenn auf dieses die Kraft F_{x_i} wirkt. Während x_i und m_i durch den Startpunkt bzw. den Atomtyp definiert sind, soll im nächsten Schritt die Berechnung der Kraft F_{x_i} , die auf das Atom wirkt, beschrieben werden.

2.2.4.1 Molekulare Kraftfelder

Die Kraft F , die auf ein Atom wirkt, wird mittels molekularer Kraftfelder berechnet. Ein Atom wird dabei als ganzes auf eine Kugel reduziert (Der Atomkern oder die Elektronen werden nicht explizit betrachtet). Welche Kräfte wirken nun auf dieses Atom? Ist das Atom Bestandteil eines Moleküls, so sind es zunächst die Kräfte der Bindungen zu anderen Atomen des Moleküls. Diese kovalenten

Bindungen werden in Kraftfeldern als „Federn“ beschrieben, die die Atome eines Moleküls verbinden und in den einfachsten Kraftfeldern dem Gesetz eines Hookeschen Oszillators gehorchen. Mit diesem Oszillator können das Stauchen und Dehnen der Bindung entlang ihrer Verbindungsachse, sowie das Biegen eines Bindungswinkels beschrieben werden. Daneben bedarf es der Beschreibung der potentiellen Energien der Torsionenwinkel einer Bindung. Neben intramolekularen, wirken intermolekulare Kräfte in Form von elektrostatischen- und van-der-Waals-Kräften auf das Atom. Die so erhaltene potentielle Energie eines Systems ist dabei nicht der *realen* potentiellen Energie gleichzusetzen, unterliegt jedoch, wie reale Prozesse der Boltzmann-Verteilung. Daraus lassen sich reale Energien und energetische Verteilungen von Systemen ableiten. Zusammenfassend lässt sich die potentielle Energie eines solchen Systems mit der folgenden allgemeinen Gleichung eines Kraftfeldes darstellen [115]:

$$E_{pot} = \sum_{\text{Bindungen}} \frac{k_i^l}{2} (l_i - l_{i0})^2 + \sum_{\text{Bindungswinkel}} \frac{k_i^\theta}{2} (\theta_i - \theta_{i0})^2 + \sum_{\text{Torsionswinkel}} \frac{V_i}{2} (1 + \cos(n\omega - \gamma)) + \quad (2.10)$$

$$\sum_{i=1}^N \sum_{j=i+1}^N \left(4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right) \quad (2.11)$$

Hierbei bezieht sich 2.10 der Kraftfeldgleichung auf die kovalenten Beiträge der Bindungen und 2.11 auf die nicht-kovalenten Beiträge. Blickt man auf den kovalenten Teil der Gleichung, findet man die Konstanten k_i^l, k_i^θ, V_n , die Referenzwerte l_{i0} für die Bindungslänge l_i , die Referenzwerte θ_{i0} für den Bindungswinkel θ_i , sowie die Multiplizität n , Diederwinkel ω und die Phasenverschiebung γ . Diese Konstanten und Referenzwerte werden als Parameter eines Kraftfeldes bezeichnet. Sie werden empirisch für das jeweilige Anwendungsgebiet eines Kraftfeldes bestimmt. Hierbei gibt es für ein Element mehrere sogenannte Atomtypen, die je nach Hybridisierungszustand oder Umgebung des Atoms unterschiedliche Parameter erhalten. Der zweite Teil der Kraftfeldgleichung (2.11) beinhaltet die Beschreibung der elektrostatischen Interaktionen gemäß dem Coulomb-Gesetzes 2.12:

$$E_{Ele} = \sum_{i=1}^N \sum_{j=i+1}^N \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (2.12)$$

Da in einem Kraftfeld-Modell die Elektronen eines Systems nicht betrachtet werden, werden die elektrostatischen Interaktionen und Eigenschaften klassisch durch Partialladungen ausgedrückt, die den Atomen zugewiesen werden. Die Summe aller paarweisen Interaktionen ergibt die elektrostatische potentielle Energie des Systems. Für große Systeme mit vielen Teilchen werden diese paarweisen

Berechnungen sehr schnell rechenintensiv, sodass bei heutigen MD-Simulations-Algorithmen, wie den in dieser Arbeit benutzen AMBER- und GROMACS-Umgebungen, für Simulationen in periodischen Umgebungen (Erläuterung siehe nächsten Abschnitt: Theorie der MD-Simulation) die sogenannte *Particle Mesh Ewald* (PME)-Methodik angewandt wird [116]. Hierbei werden die elektrostatischen Berechnungen in einen *nahen* und *fernen* Teil aufgespalten. Innerhalb der gesetzten Grenzen des *nahen* Teils werden die Interaktionen wie eben beschrieben berechnet. Im *fernen* Teil werden die Punktladungen auf ein Gitter übertragen und die Interaktionen anhand der dadurch erhaltenen Ladungsdichten mittels Fourier-Transformationen berechnet. Der zweite Teil der Kraftfeldgleichung (2.11) beinhaltet zudem die Beschreibung der van-der-Waals-Interaktionen, die durch ein Lennard-Jones 12-6 Potential beschrieben werden 2.13:

$$E_{van-der-Waals}(r) = \sum_{i=1}^N \sum_{j=i+1}^N 4\epsilon_{ij} \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \quad (2.13)$$

Die in dieser Arbeit benutzten Kraftfelder sind das zur Beschreibung für Proteine entwickelte AMBER ff14SB [117] Kraftfeld, sowie für kleine Moleküle das *General AMBER Force Field* (GAFF) [118]. AMBER ff14SB ist hierbei das aktuellste einer Generation von Kraftfeldern (ff94,ff99,ff99SB,ff12SB), die kontinuierlich weiterentwickelt werden, um die Strukturen von Proteinen adäquat darzustellen. Die Parameter des ff14SB wurden mittels NMR-Daten und quantenmechanischer Berechnungen bestimmt. Hierdurch konnte eine bessere Darstellung von α -Helices oder Rotamer-Verteilungen von Seitenketten erreicht werden [117]. GAFF, wie auch das in dieser Arbeit verwendete MMff94s-Kraftfeld [119] wurden speziell für die Simulation von organischen Molekülen entwickelt und wurden beide durch quantenmechanische *ab initio* Rechnungen, sowie empirische Daten (Röntgen- und Neutronenstrukturanalysen) parametrisiert. Neben den Kraftfeldern für Proteine und organische Moleküle spielt die Darstellung des Lösungsmittels (in biologischen Systemen Wasser) eine zentrale Rolle. Das in dieser Arbeit verwendete Modell ist das TIP3P [120]. TIP3P gehört zu der Gruppe der 3-Punkt-Modelle, bei der die Interaktionen des Wassers durch drei Interaktionspunkte dargestellt werden. Zur Vereinfachung wird hierbei die Geometrie des Wassers als starr angenommen. Daneben kam das OPLS-AA (*Optimized Potentials for Liquid Simulations - all atom*)-Kraftfeld [121, 122] für die Durchführung von sogenannten *mixed-solvent* MD-Simulationen zum Einsatz.

Neben der Verwendung in MD-Simulationen werden Kraftfelder auch als Scoringfunktionen im molekularen Docking eingesetzt. Die dort eingesetzten Kraftfelder beschränken sich zumeist auf die nicht-kovalenten Terme, um die polaren und nicht-polaren Interaktionen eines Protein-Ligand-Komplexes zu berechnen [102].

2.2.4.2 Theorie der MD-Simulation

Nachdem allgemein die potentielle Energie $V(x_i)$ eines Atoms i für seine Position x über das Kraftfeld bekannt ist, kann die neue Kraft F_i die auf das Atom nach einer Positionsänderung wirkt, berechnet werden. Zur Veranschaulichung wird ein System mit nur zwei Atomen angenommen. Die Kraft $F(r)$, die zwischen den beiden Atomen mit dem Abstand r nach einer Positionsänderung wirkt, ist über die Ableitung der potentiellen Energie nach r zugänglich:

$$F(r) = \frac{dV(r)}{dr} \quad (2.14)$$

Für ein Mehrteilchensystem setzt sich die auf ein Atom wirkende Kraft somit aus der Summe aller paarweise errechneten Kräfte zusammen. Mit der so erhaltenen Kraft F_i kann die Differentialgleichung 2.9 integriert werden. Dies erfolgt numerisch nach der Finite-Differenzen-Methode. Die Entwicklung des Systems wird für kleine Zeitintervalle Δt betrachtet. Für jeden Zeitschritt Δt wird der Kraftvektor, der auf ein Atom wirkt, als Summe aller paarweisen Interaktionen, basierend auf den neuen Positionen der Atome, berechnet. Dadurch wird die Beschleunigung a , die auf das Atom wirkt, zugänglich. Mit der Beschleunigung a kann nun anhand der aktuellen Position des Atoms und seiner Geschwindigkeit die Position zum Zeitpunkt $t + \Delta t$ vorausgesagt werden. Führt man diesen Zyklus iterativ aus, so erhält man die dynamische Entwicklung des Atoms in den Zeitschritten Δt [123].

Verschiedene Algorithmen wurden für die Lösung dieses Problems entwickelt, wobei das in dieser Arbeit benutzte Programm Packet AMBER mit den darin enthaltenen Programmen *sander* und *pmemd* auf dem Verlet-Algorithmus basieren [124]. Der Verlet-Algorithmus bestimmt die Position $x(t - \Delta t)$ eines Teilchens

zum Zeitpunkt $t + \Delta t$, anhand der Beschleunigung \mathbf{a} und Position $\mathbf{x}(t)$ zum Zeitpunkt t und der Position $\mathbf{x}(t - \Delta t)$ zum Zeitpunkt $t - \Delta t$:

$$\mathbf{x}(t + \Delta t) = \mathbf{x}(t) + \Delta \mathbf{v}(t) + \frac{1}{2} \Delta t^2 \mathbf{a}(t) \quad (2.15)$$

$$\mathbf{x}(t - \Delta t) = \mathbf{x}(t) - \Delta \mathbf{v}(t) + \frac{1}{2} \Delta t^2 \mathbf{a}(t) \quad (2.16)$$

Durch Addition dieser beiden Gleichungen ergibt sich der Verlet-Integrations-Algorithmus, der nur noch von der Position \mathbf{x} und der Beschleunigung \mathbf{a} abhängt:

$$\mathbf{x}(t + \Delta t) = 2\mathbf{x}(t) - \mathbf{x}(t - \Delta t) + \Delta t^2 \mathbf{a}(t) \quad (2.17)$$

Die Geschwindigkeit \mathbf{v} eines Teilchens kann hierbei erst im Anschluss mit der folgenden Gleichung berechnet werden:

$$\mathbf{v}(t + \frac{1}{2} \Delta t) = [\mathbf{x}(t + \Delta t) - \mathbf{x}(t)] / \Delta t \quad (2.18)$$

Wie aus dem oben genannten Sachverhalt hervorgeht, bedarf es zur Berechnung der Kräfte einer Positionsänderung der Atome. Zum Starten der Simulation werden daher jedem Atom Anfangsgeschwindigkeiten zugeteilt. Die Geschwindigkeitsverteilung wird auf der Grundlage der Maxwell-Boltzmann-Verteilung für die entsprechende Temperatur berechnet.

Der Zeitschritt Δt ist hierbei ein kritischer Faktor: ist dieser zu groß gewählt, kommt es zu Instabilitäten des Systems und Artefakten; ist er zu klein, steigt die benötigte Rechenzeit massiv an. Durch die numerische Lösung der Bewegungsgleichungen gilt die Energieerhaltung in dem simulierten System nur näherungsweise, sodass bei zu groß gewählten Zeitschritten das Simulationssystem energetisch instabil wird. Daher gilt als Richtwert, dass Δt etwa 1/10 der Zeit, der in einem System vorkommenden kleinsten beobachtbaren Veränderungen, betragen sollte. Die kleinsten vorkommenden Bewegungen in Simulationen von Proteinen, liegen im Größenbereich von 10 fs und stellen die Streckenschwingungen von C-H Bindungslängen dar. Somit dürfte in einem solchen System der maximale Zeitschritt Δt 1 fs betragen, um ein energetisch stabiles System zu erhalten. Da die Schwingungen der C-H Bindungen für die Simulation von Systemen wie Proteinen nur eine untergeordnete Rolle spielen, ist eine vielfach angewendete Lösung das „Einfrieren“ solcher Schwingungen, um die Rechenzeit zu verringern [123]. In AMBER steht hierfür der SHAKE Algorithmus [125] zu Verfügung, der die C-H Bindungslänge auf die Ideal-Werte einfriert. Somit sind Zeitschritte von 2 fs möglich, was zu einer Halbierung der Rechenzeit führt [126].

Ein solches System stellt ein sogenanntes mikrokanonisches System dar (NVE-Ensemble) mit einer konstanten Anzahl an Teilchen N , in einem konstanten Volumen V , und der konstanten Energie E , welche sich durch die klassische Hamiltonfunktion $H(\mathbf{x}, \mathbf{p})$ beschreiben lässt:

$$H(\mathbf{x}, \mathbf{p}) = V(\mathbf{x}) + K(\mathbf{p}) \quad (2.19)$$

Hierbei ist $V(\mathbf{x})$ die potentielle Energie des Systems (abhängig von den Koordinaten der Teilchen) und $K(\mathbf{p})$ die kinetische Energie des Systems (abhängig von den Impulsen der Teilchen) [113]. Ein solches adiabatisches System würde nicht einem realen experimentellen System entsprechen. Die in einem Experiment ablaufenden Reaktionen laufen bei einer entsprechenden Temperatur und einem entsprechenden Druck ab. Um konstante Temperaturen in einer MD-Simulation zu erreichen, bedient man sich eines kanonischen Ensembles (NVT), bei dem sich das System im Austausch mit einem Thermostaten befindet. Dieser „entnimmt“ oder „fügt“ dem System kinetische Energie hinzu, und hält somit die Temperatur konstant. In dieser Arbeit wurde hierzu der Langevin-Thermostat [127] verwendet. Eine vereinfachte Darstellung des Langevin-Thermostats ist, dass große Partikel (in diesem Fall die Atome der Simulation) von sehr viel kleineren Partikeln (in diesem Fall virtuelle Partikel des Thermostats) durch Kollisionen gedämpft werden und umgekehrt von diesen stochastisch kinetische Energie aufnehmen können. Die Simulation ist von diesen virtuellen Teilchen "geflutet", wobei die Anzahl an Kollisionen über einen Kollisionsparameter γ festgelegt werden können. Neben der Temperatur stellt der Druck des Systems eine weitere wichtige Größe dar. Ein System mit konstantem Druck und Temperatur wird als ein isotherm-isobares Ensemble (NPT) bezeichnet. Der Druck wird dort analog durch einen Barostaten konstant gehalten. In den mit AMBER durchgeführten Simulationen wurde hierfür der Andersen-Barostat [128] eingesetzt. Der Druck wird hierbei über das Skalieren der interpartikulären Abstände erreicht. Ein aufgesetztes Simulationssystem ist noch weit von einem stabilen energetischen Gleichgewichtszustand entfernt und muss daher vorsichtig äquilibriert werden. Hierbei wird, nach einer lokalen Energieminimierung, das System zunächst im NVT-Ensemble thermisch, und anschließend im NPT-Ensemble der Druck, äquilibriert².

²Persönliche Korrespondenz mit Carlos Simmerling, Professor für Chemie an der State University of New York at Stony Brook, AMBER Workshop 2018 in Haifa, Israel.

Um die Anzahl an simulierten Wassermolekülen so gering als möglich zu halten und mögliche Artefaktbildung durch ein angrenzendes Vakuum am Rande der Simulationsbox zu vermeiden, werden MD-Simulationen mit sogenannten periodischen Randbedingungen durchgeführt. Die Simulationszelle wird hierbei in jede Richtung durch Kopien vervielfältigt. Verlässt ein Molekül auf der einen Seite die Simulationszelle, wird es durch ein identisches auf der anderen Seite ersetzt. Somit sind Simulationen in einem großen „Bulk“ an Wasser möglich, ohne diese explizit simulieren zu müssen [123].

Die eben aufgeführten theoretischen Grundideen und Teilaspekte von MD-Simulationen stellen nur eine kurze vereinfachte Einführung dar. Für detaillierte Darstellungen dieses komplexen und umfangreichen Themas sei hier auf die Lehrbücher [129, 130] verwiesen.

2.3 Fragmentbasiertes Wirkstoffdesign (Fragment based drug design, FBDD)

2.3.1 Idee des FBDD

Ein Blick in die Milchstraße in einer mondlosen Nacht offenbart eine unüberschaubare Anzahl an Sternen. Überschlägt man nach dem derzeitigen Stand der Wissenschaft die Anzahl der Sterne in dem uns bekannten Universum, kommt man auf eine Zahl zwischen 10^{22} und 10^{24} [131]. Geht man einer analogen Frage im Kontext der medizinischen Chemie nach und fragt, wie viele Moleküle (mit einem Molekulargewicht kleiner 500 Da) es theoretisch im „chemischen Raum“ gibt, so kommt man auf eine Anzahl von 10^{60} Molekülen [132]. Vor diesem Hintergrund wird offensichtlich, dass Hochdurchsatzverfahren (*High throughput screening*, HTS) zur Leitstruktursuche, wie sie in den 1990er und 2000er Jahren der „Goldstandard“ waren, selbst mit Substanzbibliotheken von mehreren 10 Millionen Verbindungen nur einen verschwindend geringen Teil dieses chemischen Raums abdecken können [133].

Mit einer Anzahl von typischerweise weniger als 20 Schweratomen [134] umfassen Fragmente einen sehr viel kleineren chemischen Raum. *Ruddigkeit et al.* zeigten anhand ihrer GDB-17 Datenbank, dass basierend auf der Kombination von bis zu 17 Schweratomen der Elemente C, N, O, S, sowie der Halogene ein chemischer Raum von „nur“ 166 Milliarden Molekülen existiert [135]. Eine Fragmentbibliothek mit mehreren tausend Verbindungen deckt demnach einen solchen Raum viel effizienter ab, als dies bei mehreren Millionen Verbindungen in einem HTS der Fall ist [133].

Ein weiterer Vorteil des FBDD beruht auf dem von *Hann et al.* [136] aufgestellten Konzept der molekularen Komplexität: *Hann et al.* [136] zeigten, dass die Wahrscheinlichkeit, ein Molekül zu finden, das sich an eine spezifische Bindestelle einer Zielstruktur anlagert, wächst, je kleiner dieses Molekül ist. Umgekehrt fällt die Wahrscheinlichkeit drastisch ab, je größer die Verbindung ist. Eine Bibliothek aus kleinen Fragmenten eignet sich demnach besser, mögliche bindende Strukturen zu finden, als eine Bibliothek mit größeren Molekülen. Den „Preis“ für diese höhere Trefferquote zahlt man allerdings mit deutlich schwächeren Affinitäten, die typischerweise im μM - bis mM -Bereich liegen und höhere Anforderungen an die eingesetzten Detektionsmethoden stellen. Ein weiterer Aspekt bezieht sich

auf die Optimierungsmöglichkeiten von Fragmenten. Es ist einfacher, ein kleines schwach bindendes Molekül durch Hinzufügen entsprechender funktioneller Gruppen in seiner Affinität zu optimieren als eine größere, schlecht bindende Verbindung zu verbessern [137]. Vor diesem Hintergrund, ermöglicht das FBDD auch die Adressierung neuer pharmakologischer Zielstrukturen, wie z.B. PPI. Hier bietet der Einsatz von Fragmenten in NMR-Methoden, Röntgenstrukturanalysen oder computerbasierten Methoden die Möglichkeit, die für die Bindung wichtigen *hot-spots* auf der Proteinoberfläche zu identifizieren und einzelne, an unterschiedliche *hot-spots* bindende Moleküle zu einer potenteren Verbindung zu verknüpfen [138, 139]. Auf diese Weise wurden erfolgreich PPI-Inhibitoren, z.B. für IL-2 [140], TNF- α und BCL-X_L [141] entwickelt. Aufgrund dessen etablierte sich das FBDD in den letzten 20 Jahren als ein mächtiges Werkzeug der Leitstruktursuche in der pharmazeutischen Industrie, wie auch im akademischen Bereich. Unterstrichen wird dies dadurch, dass aktuell drei Wirkstoffe aus erfolgreichen FBDD-Kampagnen eine Zulassung erhalten haben (Erdafitinib, Vemurafenib, Venetoclax) und sich über 40 weitere Kandidaten in den klinischen Phasen-I bis -III befinden [133, 138, 142].

2.3.2 Aufbau und Ablauf eines FBDD

Für den Ablauf eines FBDD gibt es kein festes Schema, und jedes Screening enthält individuelle Schritte, die je nach Methode oder Problem gewählt werden müssen. Generell bedarf es aber der folgenden Überlegungen [134]:

1. die Zusammenstellung der Fragmentbibliothek;
2. die Auswahl geeigneter Methoden, um die schwache Bindung der Fragmente zu detektieren und die Bindestellen aufzuklären;
3. Kriterien zur Auswahl der Fragmente zur weiteren Optimierung;
4. Methoden zur Optimierung der Fragmente zu möglichen Leitstrukturen.

Im Folgenden soll auf die oben genannten Punkte im Kontext dieser Arbeit kurz eingegangen werden.

2.3.3 Fragmentbibliotheken

Die Zusammenstellung einer geeigneten Fragmentbibliothek ist ein Kernpunkt im FBDD. Eine ideale Bibliothek sollte (ähnlich wie im HTS) robuste und reproduzierbare Ergebnisse liefern, sowie chemisch und strukturell vielseitig sein. Sie

sollte anhand der Treffer erste Struktur-Wirkungsbeziehungen (SAR) erkennen lassen, die eine hohe Ligandeffizienz (LE) besitzen und sich somit schnell zu Leitstrukturen weiterentwickeln lassen [137].

Die Größen dieser Bibliotheken liegen zwischen einigen hundert bis zu mehreren zehntausend Molekülen, wobei die meisten Bibliotheken im Mittel zwischen 1000 und 2000 Fragmente beinhalten [133, 143]. Neben allgemeinen Kriterien wie Reinheit, Stabilität und vor allem Löslichkeit in wässrigen Medien (typische Konzentrationen bei initialen Screens liegen bei 500 μM und höher), spielen auch die verwendeten Methoden eine Rolle. So gibt es spezielle Bibliotheken, z.B. mit Brom-substituierte Moleküle für kristallographische Screenings, oder mit Fluor ^{19}F -substituierte Moleküle für NMR basierte Techniken. Auch synthetische Aspekte sind für die spätere Optimierung der Fragmente von Bedeutung. So sollten die Moleküle funktionelle Gruppen enthalten, die ein späteres Verknüpfen oder Expandieren der Fragmente ermöglichen [133, 137].

Daneben gibt es verschiedene konzeptionelle Ansätze zur Zusammenstellung von Bibliotheken, zum Beispiel:

- anhand von physikochemischen Kriterien wie z.B. der *Rule of 3* (Ro3) [144];
- basierend auf Naturstoffen oder Metaboliten. Die Idee hierbei ist, den biologisch relevanten Bereich des chemischen Raumes besser abzudecken. Solche Fragmentbibliotheken werden beispielsweise von den Herstellern *AnalytiCon*, *LifeChemicals* oder *Asinex* angeboten und von der bei [145] berichteten *Fragments of Life*-Bibliothek;
- von bereits bekannten Wirkstoffen abgeleitete Fragmente [138];
- für spezielle Zielstrukturen, wie Metalloproteine, PPI oder Kinasen. [146];
- anhand von strukturellen Merkmalen.³

In dieser Arbeit wurde die Fragmentbibliothek anhand von physikochemischen Kriterien zusammengestellt, auf deren Kriterien näher in 5.1.1.1 eingegangen wird. Ein weit verbreitetes Kriterium stellt hier die Ro3 dar [144], die ähnlich der *Rule of 5* (Ro5) [64] physikochemische Grenzen aufstellt, innerhalb derer die Fragmente

³Eine Analyse der Fragmente der ZINC-Datenbank zeigte, dass die meisten Fragmente (75%) eine lineare oder flache Struktur besitzen. Bindungen in der Natur beruhen aber zumeist auf komplexeren räumlichen Strukturen. Daher ist das gegenwärtige Bestreben, Bibliotheken zu synthetisieren, die mit sp^3 Atomen angereichert sind, um solchen räumlich komplexeren Strukturen Rechnung zu tragen. Neben vermutlich höheren Trefferquoten sollen solche Moleküle zudem eine bessere Löslichkeit, sowie bessere ADMET-Eigenschaften besitzen [147–150].

liegen sollten (Tabelle 2.2).

Vor dem Hintergrund von PPI zeigte ein Vergleich aktiver Fragmente gegen PPI mit normalen Fragmentbibliotheken, dass die dort enthaltenen Verbindungen im Mittel ein höheres Molekulargewicht, einen höheren $\log P$, mehr saure oder basische Funktionen und größere hydrophobe Bereiche besitzen [151]. Darauf basierend wurde für Fragmente gegen PPI die *Rule of 4* (Ro4) postuliert [152, 153] mit den in Tabelle 2.2 genannten Definitionen.

Tabelle 2.2: Vergleich der Ro3, mit der Ro4, den *lead-like*-Kriterien [154] und der Ro5

	Ro3	Ro4	Lead-like	Ro5
MW	<300	<400	≤ 460	<500
(c)LogP (o/w)	≤ 3	≤ 4	>-4 und ≤ 4.2	≤ 5
(c)LogS (Wasser)	-	-	≥ -5	-
Rotierbare Bindungen	≤ 3	-	≤ 10	-
Ringsysteme	-	<4	<4	-
H-Brücken-Donatoren	≤ 3	-	≤ 5	≤ 5
H-Brücken-Akzeptoren	≤ 3	<4	≤ 9	≤ 10
TPSA (\AA^2)	≤ 60	-	-	-

Diese Ro4 wird allerdings in [155] kritisch gesehen. In der Vergrößerung der oben genannten Parameter sehen *Jhoti et al.* [155] keinen Vorteil für die Suche nach bindenden Fragmenten im Kontext von PPIs. Aufgrund von 30 Fragmentscreenings⁴ gegen verschiedene Zielprotein-Klassen, zogen *Jhoti et al.* den Schluss, dass Bibliotheken, die der Ro3 folgen, auch für PPI erfolgreich eingesetzt werden können. Die Fragmentbibliotheken, die in dieser Arbeit verwendet wurden, folgten größtenteils der Ro3 und enthielten zusätzlich solche Moleküle, die der Ro4 folgten.

Zur Charakterisierung von Fragmentbibliotheken bedient man sich der bereits vorgestellten PCA oder *Molecular Shape* Analysen [156, 157].

⁴Bei der Firma Astex Therapeutics

2.3.4 Biophysikalische Methoden im FBDD

Generell kann jede Methode zur Detektion der Bindung eines Fragments benutzt werden, wenn sie sensitiv genug ist, um die recht schwachen Affinitäten von Fragmenten im μM - bis mM -Bereich zu erfassen. Weitere praktische Anforderungen stellen der Durchsatz und die benötigten Mengen an Fragmenten bzw. Protein dar [158]. Laut einer Umfrage von Daniel Erlanson [159] werden aktuell folgende biophysikalische Methoden im FBDD eingesetzt⁵:

- Röntgenstrukturanalyse (82 %)
- Ligandbasierte NMR-Methoden (76 %)
- Oberflächenplasmonenresonanz-Spektroskopie (*Surface-Plasmon-Resonance Spectroscopy*, SPR-Spektroskopie) (76 %)
- *Thermal-Shift-Assays* (TSA) (74 %)
- Proteinbasierte NMR-Methoden (52 %)
- Funktionale Assays (44 %)
- Isotherme Titrationskalorimetrie (*Isothermal Titration Calorimetry*, ITC) (40 %)
- Massenspektrometrie (MS)
- *Microscale Thermophoresis* (MST) und andere (weniger als 30 %)
- Biolayer-Interferometrie (BLI) (8 %)

NMR, SPR, TSA und die Röntgenstrukturanalyse stellen hier die am häufigsten angewandten Methoden dar [160]. In dieser Arbeit kamen die ligandbasierte Sättigungstransfer-Differenz (STD)-NMR-Methode (siehe Kapitel 2.4.1), die Röntgenstrukturanalyse (siehe Kapitel 2.4.2) und die Biolayer-Interferometrie (siehe Kapitel 2.4.3) zum Einsatz.

Der Aufbau eines Fragment-Screenings enthält im Wesentlichen ein initiales-Screening, eine Prüfung der Konzentrationsabhängigkeit der initial gefundenen Treffer und ein zweites, sogenanntes orthogonales Screening. Im orthogonalen Screening werden alle im initialen-Screening gefundenen Treffer erneut auf ihre Aktivität hin vermessen. Der orthogonale Assay sollte sich grundlegend vom Messprinzip des primären Assays unterscheiden [161, 162]. Dadurch sollen mögliche falsch-positive Treffer (Artefakte) identifiziert und entfernt werden.

Die Aufklärung des Bindemodus eines Fragments ist ein entscheidender Schritt im FBDD, der es ermöglicht, das schwach bindende Molekül weiter zu optimieren und dessen Affinität zu steigern. Der „Goldstandard“ ist hier sicherlich die

⁵An der Umfrage im Dezember 2019 nahmen 137 Teilnehmer aus Industrie und Forschungseinrichtungen teil. Gefragt wurde nach allen benutzten Methoden im FBDD. Die Prozentzahlen sind der in der Veröffentlichung enthaltenen Grafik entnommen [159].

Röntgenstrukturanalyse. Sie ermöglichte als eine der ersten Methoden die Detektion von solchen schwachen bindenden Molekülen. Daneben liefert sie detaillierte Einblicke in die Protein-Fragment-Interaktionen und ermöglicht hierdurch ein strukturbasiertes Optimieren und Designen neuer Moleküle [163, 164]. Doch oftmals ist es gerade in der Initial-Phase des FBDD nicht möglich, Kristallstrukturen der Fragmente in Komplex mit ihrer Zielstruktur zu gewinnen, weswegen andere Methoden eingesetzt werden, um strukturelle Informationen zu erhalten [165]. Hierzu zählen vor allem NMR-basierte Techniken, wie ^1H - ^{15}N -TROSY-HSQC oder ^1H - ^{13}C -HSQC Experimente mit isotopenmarkierten Proteinen, ILOE-, INPHARMA- oder DEEP-STD-Experimente [166–168].

Aber auch chemische Ansätze, wie die Vermessung von auf der Grundlage von SAR-Überlegungen zusammengestellten Katalogen initialer Treffer oder kovalente Ansätze, bei denen Fragmente mit reaktiven Gruppen modifiziert werden und die Bindung an Proteine anschließend massenspektrometrisch detektiert werden kann, spielen hier eine Rolle [165]. Nicht zuletzt kommen auch computerbasierte Verfahren wie Docking und MD-Simulationen zum Einsatz, um mögliche Bindestellen der Fragmente zu lokalisieren [165].

2.3.5 Auswahl geeigneter Fragmente über Ligandeffizienz-Indizes

Nach einem erfolgreichen Screening folgt die Optimierung der gefundenen Fragmente zu möglichen Leitstrukturen. Die Optimierung beinhaltet nicht nur die Erhöhung der Affinität, sondern auch die Verbesserung des pharmakokinetischen Profils eines Moleküls. Das pharmakokinetische Profil lässt sich über das Molekulargewicht und die Lipophilie eines Moleküls beeinflussen [169]. Die Optimierung bedarf der rechten Balance zwischen den drei Faktoren (Affinität, Molekulargewicht und Lipophilie). Nicht jeder initiale Treffer lässt sich in diesem Sinne weiterentwickeln. Im Wirkstoffdesign wurden verschiedene Effizienz-Indizes entwickelt, um den Prozess der Optimierung zu überwachen.

Die von *Kuntz et al.* [170] vorgeschlagene *ligand efficiency* (LE) ist sicherlich der am weitesten verbreitete Index. Er stellt das Verhältnis der freien Enthalpie ΔG^6 bzw. der Affinität zur Molekülgröße dar (Gleichung 2.20). Die Molekülgröße wird

⁶Hierbei handelt es sich genau genommen nicht um den korrekten thermodynamischen Sachverhalt des Zusammenhangs zwischen freier Enthalpie und der Dissoziationskonstante K_D , da dieser nur bei Standardbedingungen gegeben ist. Vielmehr ist dies eine Näherung, um die gemessene Affinität von Fragmenten zur Molekülgröße in Relation setzen und Verbindungen somit besser relativ zu einander vergleichen zu können. Es geht bei den Ligandeffizienz-Indizes nicht

hierbei über die Anzahl der Schweratome (*heavy atom*, HA) ausgedrückt.

$$LE = \frac{\Delta G}{HA} = \frac{-RT \ln K_D}{HA} \quad (2.20)$$

Anhand dieser Normalisierung der Affinität bezüglich der Molekülgröße bietet die LE einen annähernd größenunabhängigen Vergleich von Molekülen. In ihrer Arbeit zeigten *Kuntz et al.* des Weiteren, dass die maximal erreichbare Bindungsaffinität mit steigender Molekülgröße zunimmt, während gleichzeitig die maximale Bindungsaffinität pro Schweratom abnimmt. Vor diesem Hintergrund zeigt sich, dass die Überwachung der Optimierung von Fragmenten, bei der die Molekülgröße um 10-25 Schweratome ansteigt, mittels der LE problematisch und nicht optimal für Fragmente geeignet ist [170]. Trotz allem ist die LE einer der meistgenutzten Indizes im frühen Wirkstoffdesign. Ein weit verbreiteter Grenzwert für die LE ist, dass eine Verbindung einen Wert von $0.3 \text{ kcal mol}^{-1}$ pro Schweratom (HA^{-1}) erreichen sollte [171, 172].

Der „größenunabhängige Ligandeffizienz-Index“ (*size-independent ligand efficiency*, SILE) stellt eine Weiterentwicklung des LE dar, um dem eben genannten Trend der sinkenden maximalen Bindungsaffinität pro Schweratom bei Vergrößerung der Moleküle entgegenzuwirken [173]. Bei dem SILE-Index wird die Anzahl der Schweratome mit einem auf experimentellen Daten basierenden Exponenten $x = 0.3$ modifiziert:

$$SILE = \frac{pIC_{50}}{HA^{0.3}} \quad (2.21)$$

Hierbei wird die Näherung $\Delta G \sim pIC_{50} \sim pK_D$ angenommen, da hier nur die relativen Werte zum Einordnen der Moleküle untereinander von Interesse sind und nicht absoluten Bindungsenergie berechnet werden. Während bei der Verwendung des LE die Werte während der Optimierung von Fragmenten keinen eindeutigen Trend zeigen, steigt der SILE-Index während einer erfolgreichen Fragmentoptimierung deutlich an, und bietet somit einen effizienten Weg zum Überwachen des Prozesses [173].

Um der Lipophilie der Moleküle Rechnung zu tragen, wurden Indizes entwickelt, die den $\log P$ miteinbeziehen. Einen der ersten Indizes, der die Lipophilie mit der Affinität verknüpft, stellt die von *Leeson & Springthorpe* [174] entwickelte *lipophilic ligand efficiency* (LLE) dar:

$$LLE = pIC_{50} - \log P \quad (2.22)$$

um die genaue Berechnung der freien Enthalpie einer Reaktion. Dies wird nicht zuletzt durch die in unterschiedlichen Indizes eingesetzten Näherungen $\Delta G \sim pK_D \sim pIC_{50}$ deutlich.

Mit einem angenommenen $\log P$ von 2.5 und Affinitäten zwischen 1 bis 10 nM wären typische LLE-Werte von oral verfügbaren Wirkstoffen größer 5. Für typische Fragmente ($\log P$ -Werte von 2 und kleiner, Affinitäten von 100 μM bis 1 mM) ist das Erreichen solcher Werte fast unmöglich, sodass der LLE typischerweise eher für die Optimierung von Leitstrukturen hin zu ersten Wirkstoffkandidaten eingesetzt wird.

Einen weiteren Index stellt der *lipophilicity corrected ligand efficiency* (LELP)-Index dar [171]:

$$LELP = \log P / LE \quad (2.23)$$

Er kann als Maß aufgefasst werden, inwiefern die Bindung eines Moleküls auf dessen Lipophilie beruht. Mit der oben bereits erwähnten LE von 0.3 kcal mol⁻¹ HA⁻¹ und für Leitstrukturen typischen $\log P$ -Werten zwischen -3 und 3 ergibt sich ein akzeptabler Bereich von -10 bis 10. Auch dieser Index ist nicht speziell für Fragmente optimiert und hat hier den Nachteil, dass LELP-Werte fälschlicherweise im angestrebten Bereich liegen können, wenn ein Molekül eine niedrige LE und einen sehr geringen $\log P$ besitzt, was bei Fragmenten durchaus nicht ungewöhnlich ist.

Da weder der LLE noch der LELP optimal für Fragmente sind, entwickelten *Mortenson & Murray*⁷ einen weiteren Index speziell für Fragmente [172]. Dieser Index beruht auf der Überlegung, den ΔG -Term in zwei Terme zu zerlegen: in einen ΔG^* -Term und einen ΔG_{lipo} -Term, der der Überführung eines Liganden aus der wässrigen Umgebung des Solvens in die lipophilere Umgebung des Proteins Rechnung trägt:

$$\Delta G = \Delta G^* - \Delta G_{\text{lipo}} \approx RT \ln(K_D) + RT \ln(P) \approx \ln(10) \cdot RT (\log P - pK_D) \quad (2.24)$$

ΔG^* wird dabei wieder über den $RT \ln(K_D)$ bzw. $RT \ln(IC_{50})$ -Wert abgeschätzt (mit der Gaskonstanten R und der Temperatur T). ΔG_{lipo} wird über $RT \ln(P)$ abgeschätzt. Der so erhaltene Term $\ln(10) \cdot RT (\log P - pK_D)$ ist dabei proportional zu LLE über:

$$\ln(10) \cdot RT (\log P - pK_D) = -1.37(LLE) \quad (2.25)$$

Um in den gleichen Größenbereich des weitverbreiteten LE-Index zu kommen,

⁷bei Astex Therapeutics

wird der erhaltene Term über eine Konstante $x = 0.11$ im Skalenbereich verschoben und mit der Anzahl an Schweratomen HA normiert. Der so erhaltene LLE_{AT} -Index (AT, für ASTEX) lässt sich somit mit der Folgenden finalen Gleichung berechnen:

$$LLE_{AT} = 0.11 - 1.37(LLE/HA) \quad (2.26)$$

Ein akzeptabler Grenzwert liegt somit, wie bei der LE, bei 0.3. Eingesetzt wurde dieser LLE_{AT} -Index unter anderem bei der Optimierung des HSP90-Inhibitors AT13387 (basierend auf einem Fragment) durch das Pharmaunternehmen Astex Therapeutic [175].

Zusammenfassend lässt sich festhalten, dass es nicht den einen „wahren“-Wert oder Index gibt. Vielmehr kommt es auf die richtige Kombination bzw. den Einsatz des passenden Index abhängig vom jeweiligen Problem an. Dennoch helfen die vorgestellten Indizes, diejenigen Fragmente aus einer Liste von Treffern herauszusuchen, deren Optimierung am meisten verspricht. Aktuell werden beim FBDD vor allem der LE-Index (>90 %) und der LLE-Index (>70 %) in der Optimierung verwendet [175].

2.3.6 Methoden zur Optimierung der Fragmente

Nach der Auswahl von geeigneten Fragmenten schließt sich der Prozess der Optimierung der Moleküle zu möglichen Leitstrukturen an. In dieser Arbeit konnte das Stadium der Optimierung noch nicht erreicht werden. Daher soll es nur kurz der Vollständigkeit halber erwähnt werden. Generell gibt es drei Methoden (das Wachsen, das Zusammenführen und das Verknüpfen von Fragmenten), um in ihrer Affinität gesteigerte neue Verbindungen zu erhalten (Abbildung 2.6) [176].

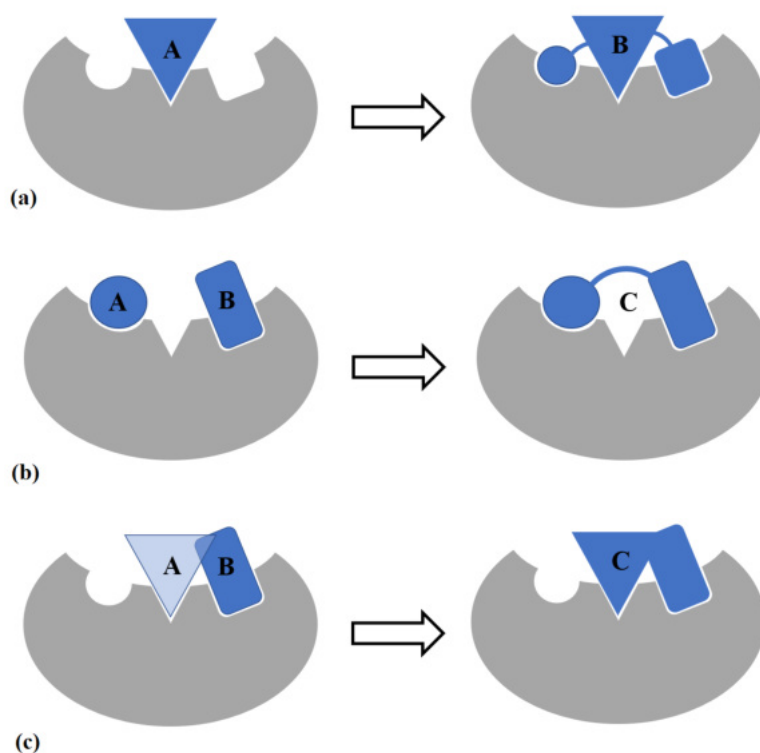


Abbildung 2.6: Fragmentoptimierung Möglichkeiten der Optimierung von Fragmenten mit (a) dem Wachsen, (b) dem Verknüpfen und (c) dem Zusammenführen von Fragmenten.

Beim Wachsen (a) wird (im günstigsten Fall auf der Grundlage von Röntgenstrukturdaten) das Fragment Stück für Stück vergrößert, um eine höhere Affinität zu erhalten. Hierbei kommen vor allem computerbasierte Methoden zum Einsatz, die das Design leiten. Beim Verknüpfen (b) werden benachbart bindende Fragmente über einen Linker miteinander verknüpft. Die dritte Möglichkeit ist das Zusammenführen (c) von Fragmenten. Im Gegensatz zum Verknüpfen überlappen hier die Bereiche der Fragmente, sodass aus den Bindungsepitopen beider Fragmente ein neues Molekül entsteht.

2.3.7 Computerbasierte Methoden im fragmentbasierten Kontext

Computerbasierte Methoden können im FBDD in allen Bereichen der Wirkstofffindung eingesetzt werden, wie sie bereits in Kapitel 2.2 genannt wurden. Gerade zu Beginn eines FBDD, wenn nähere Strukturinformationen durch Röntgenstrukturanalyse oder NMR-Techniken fehlen, bieten computerbasierte Methoden die Möglichkeit der Identifizierung von möglichen Bindestellen der Fragmente oder *hot-spots* für Interaktionen. In diesem Kontext sollen hier zwei in dieser Arbeit

verwendete Methoden näher erläutert werden.

Der FTMap-Algorithmus wurde als computerbasierte Analogie zu Röntgenstruktur- oder NMR-Experimenten zur Identifizierung von potentiellen Bindestellen an Proteinen entwickelt. Bei diesen Experimenten werden kleine organische Moleküle (z.B. Aceton, Acetonitril, DMSO, Isopropanol, Harnstoff oder Methanol) als chemische „Sonden“ benutzt, um mögliche *hot-spots* an Proteinen zu lokalisieren [177]. Der FTMap-Algorithmus benutzt 16 solcher chemischer Sonden (Tabelle 2.3) unterschiedlicher Gestalt, Größe und Polarität [178]. Das Verfahren kann in vier Schritte unterteilt werden, die im Folgenden kurz umrissen werden sollen [179].

1. Die 16 chemischen „Sonden“ werden unter Verwendung eines schnellen Fourier-Transformations-Ansatzes (*fast Fourier transform*, FFT) einem starren Docking (*rigid body docking*) unterzogen. Der FFT-Ansatz kann hierbei Milliarden an möglichen Posen eines Moleküls untersuchen, bietet aber nur die Möglichkeit einer einfachen Scoringfunktion. Die verwendete Scoringfunktion enthält vier Terme: einen van-der-Waals Term; einen elektrostatischen Term; einen sogenannten Kavitätsterm, der die als günstig angenommene Bindung in nicht-polare Kavitäten des Proteins beschreibt; sowie eine wissensbasierte Potentialfunktion zur Beschreibung von Solvatationseffekten. Von jedem Sondenmolekül werden die 2000 best-bewerteten Posen weiterverwendet.
2. Alle 2000 Bindeposen jeder Sonde werden unter Verwendung des CHARMM-Kraftfeldes minimiert. Das Protein wird hierbei als starr definiert.
3. Die minimierten Posen jeder Sonde werden geclustert (jede Sonde wird getrennt behandelt).
4. Alle im vorherigen Schritt erstellten Cluster werden überlagert, um sogenannte Consensus-Cluster aller Sonden zu bilden (Radius 4 Å). Der Cluster mit den meisten Sonden ist hierbei der höchstbewertete Cluster, u.s.w.

Eine detaillierte Beschreibung des FTMap-Algorithmus mit einer genauen Beschreibung der Scoringfunktion ist in [179] zu finden.

Tabelle 2.3: Liste der verwendeten chemischen Sonden des FTMap-Algorithmus

Acetaldehyd	Benzaldehyd	Dimethylether	Harnstoff
Acetamid	Benzen	Cyclohexan	Isopropanol
Aceton	<i>tert</i> -Butanol	Ethan	Methylamin
Acetonitril	N,N-Dimethylformamid	Ethanol	Phenol

Mixed-solvent Simulationen versuchen ebenfalls, die grundlegende Frage nach möglichen günstigen Bindestellen an Proteinen zu klären. Während Ansätze wie der FTMap-Algorithmus mit einem starren Protein ohne explizites Solvensmodell arbeiten, versuchen *mixed-solvent* Simulationen ein realistischeres physikalisches Modell unter expliziten Solvens-Bedingungen und mit flexiblem Protein zu liefern. Alle in der Literatur berichteten Ansätze [180–183] simulieren hierbei eine hohe Konzentration an Fragmenten, die in der Solvensbox „gelöst“ werden und analysieren anschließend den Besetzungsgrad der Fragmente an der Proteinoberfläche. Die Konvergenz solcher Simulationen ist hierbei von der Konzentration der Fragmente, der Flexibilität des Proteins und der Simulationszeit abhängig [184].

Der in dieser Arbeit verwendete *mixed-solvent*-Ansatz basiert auf der in [180, 185] beschriebenen SILCS (*Site Identification by Ligand Competitive Saturation*)-Methode. Um die Agglomeration vor allem lipophiler Fragmente bei hohen Konzentrationen von 1 M zu verhindern, werden hier Dummy-Atome in den Molekülen definiert, die untereinander durch ein Lennard-Jones-Potential repulsiv wechselwirken und somit einer Aggregation entgegenwirken. Dieses repulsive Potential wirkt nur zwischen den definierten Dummy-Atomen und hat keinen Einfluss auf andere Atome in der Simulation. Die Position der Dummy-Atome wird mit Hilfe des Massenschwerpunktes eines Moleküls festgelegt. Die Simulationen werden in mehreren Replikas durchgeführt, um die nötige Konvergenz zu erreichen.

2.4 Theoretische Hintergründe zu den eingesetzten biophysikalischen Methoden

2.4.1 Kernspinresonanz-Spektroskopie

Kernspinresonanz (*Nuclear Magnetic Resonance*, NMR)-Techniken stellen zusammen mit der Röntgenstrukturanalyse die ältesten Methoden im FBDD dar und bereiteten mit den Arbeiten von *Shuker et al.* im Jahr 1996 den Weg für das FBDD [186]. Hier soll auf die in dieser Arbeit verwendete Sättigungstransfer-Differenz

(STD)-Methode näher eingegangen werden. In diesem Kontext kann nicht auf die komplette dahinterstehenden Theorie eingegangen werden; der Leser sei auf die weiterführende Literatur, auf deren Grundlage die folgenden Ausführungen beruhen, verwiesen [187, 188].

2.4.1.1 Grundlagen

Jeder Atomkern besitzt die Kernspinquantenzahl I , die als Drehimpuls (Spin) des Kerns vorgestellt werden kann und positive oder negative Werte (je nach Drehrichtung des Kerns) aufweist. Durch den Impuls und seine Ladung besitzt jeder Kern ein magnetisches Moment μ . Durch ein äußeres, angelegtes Magnetfeld (z.B. des Erdmagnetfelds) besitzen die Kerne (im Folgenden am Beispiel eines ^1H dargestellt) unterschiedliche Energieniveaus E_α und E_β und somit eine Energiedifferenz ΔE (Abbildung 2.7 (a)). Diese Energieniveaus sind laut der Boltzmannverteilung:

$$\frac{N_\alpha}{N_\beta} = e^{\Delta E/k_B T} \quad (2.27)$$

mit k_B der Boltzmannkonstanten, bei einer Temperatur T unterschiedlich stark besetzt. Durch die unterschiedliche Besetzung existiert ein Bulk magnetisierungsvektor M , der durch die Summe der durch das Magnetfelds B_0 induzierten magnetischen Momente μ in Richtung der z -Achse dargestellt werden kann (Abbildung 2.7 (b)). Legt man ein zusätzliches magnetisches Feld B_1 in Form von niederfrequenter elektromagnetischer Strahlung (Radiofrequenz, RF) entlang der x -Achse an, werden einzelne Kerne angeregt und auf das höhere Energieniveau angehoben, wodurch der Vektor M sich in Richtung der xy -Ebene bewegt (Abbildung 2.7 (c)). Bei der Änderung des Vektors M wird eine niederfrequente elektromagnetische Strahlung erzeugt, die das Messsignal des NMR-Experiments darstellt (Gezeigt ist hier ein 90° Puls bei dem der Vektor M in die xy -Ebene fällt). Diesen Effekt beobachtet man nur für Kerne mit nicht-ganzzahligen Kernspinquantenzahlen.

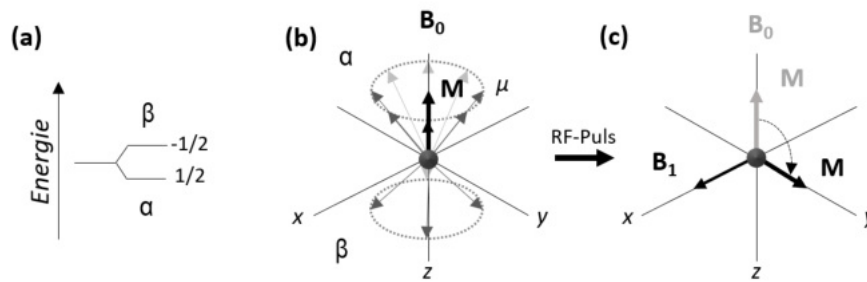


Abbildung 2.7: NMR-Vektormodell

(a) In Gegenwart eines Magnetfelds spaltet sich die Kernspinquanzenzahl in unterschiedliche Energieniveaus auf. (b) Die magnetischen Momente der Kerne bilden hierbei einen Bulkmagnetisierungsvektor M . (c) Strahlt man nun einen RF-Puls in die Probe, werden die Kerne auf das höhere Energieniveau E_β angeregt, wodurch sich der Bulkmagnetisierungsvektor in Richtung der xy -Ebene bewegt. Diese Änderung wird im NMR gemessen.

2.4.1.2 T_1 und T_2 -Relaxationszeit

Die T_1 -Relaxationszeit (longitudinale Relaxationszeit) beschreibt, wie lange es nach einem RF-Puls dauert, bis der Bulkmagnetisierungsvektor M wieder seine ursprüngliche Orientierung einnimmt. Die durch den RF-Puls erhaltene Energie wird hierbei in Form von Wärme an die Umgebung abgegeben. Kleine Moleküle haben hier aufgrund weniger Spin-Spin-Wechselwirkungen und ihrer schnellen Umorientierung in Lösung lange T_1 -Relaxationszeiten. Je nach Frequenz des Spektrometers durchläuft die T_1 -Relaxationszeit bei größeren Molekülen (wenige tausend Da) ein Minimum und steigt für sehr große Moleküle (mehrere tausend Da) wieder an. Die T_2 -Relaxationszeit (transversale Relaxationszeit) beschreibt den Verlust an Magnetisierung in der xy -Ebene. Dieser wird durch kleine lokale Differenzen des Magnetfeldes am Kern verursacht. Die T_2 -Relaxationszeit setzt sich zusammen aus einem vom Magnetfeld abhängigen Teil T_2^* und einem molekularen Anteil T_2 , der auf der Umgebung des Kerns bzw. Protons im Molekül beruht. Die T_2 -Relaxationszeit ist hierbei umgekehrt proportional zur Molekülgröße. Kleine Moleküle besitzen lange, große Moleküle (wie Proteine) aufgrund der vielen intramolekularen Spin-Spin-Wechselwirkungen kurze T_2 -Relaxationszeiten.

2.4.1.3 Spindiffusion und Sättigung

Der Spindiffusionsprozess beruht auf der Kopplung der magnetischen Dipole der Kerne in einem Molekül. Hierbei überträgt ein Kern die Magnetisierung auf benachbarte Kerne. Die Kopplung ist demnach vor allem vom Abstand der Spins

untereinander abhängig. Diese Kopplung von Spins basiert auf dem sogenannten Kern-Overhauser-Effekt (*nuclear Overhauser effect*, NOE) und ist besonders effektiv bei großen Molekülen, wie Proteinen. Als Maß der Spindiffusion kann die T_2 -Relaxationszeit eines Moleküls herangezogen werden: je kürzer diese ist, desto besser ist die Spindiffusion in einem Molekül [189]. Diese Spindiffusionsprozesse sind vor allem bei Proteinen äußerst effektiv und führen dazu, dass eine selektive Anregung, z.B. von Methyl-Gruppen, schnell auf das ganze Protein übertragen wird. Bei der Sättigung wird einer Probe dauerhaft Energie zugeführt, so dass der Bulk magnetisierungsvektor M dauerhaft in der xy -Ebene bleibt. Hierdurch kommt es zu keiner Änderung des Vektors mehr, wodurch kein Signal mehr von der Probe abgestrahlt wird.

2.4.1.4 Das Sättigungstransfer-Differenz (STD)-NMR-Experiment

Das STD-NMR Experiment, zur Detektion der Ligandbindung an Proteine, wurde erstmals bei Mayer & Meyer [190] beschrieben und baut auf der schnellen Spindiffusion in Proteinen auf. Es werden hierzu zwei Spektren aufgenommen: ein sogenanntes *on*-Spektrum und ein *off*-Spektrum [167, 191, 192]. Beim *on*-Spektrum wird das Protein durch einen RF-Puls selektiv gesättigt. Der Einstrahlpunkt muss hierbei so gewählt werden, dass keine Resonanzfrequenzen der Protonen des Liganden getroffen werden. Er liegt daher typischerweise im Bereich von 0.5 bis -1 ppm. Bindet nun ein Ligand an das gesättigte Protein, wird über die Spindiffusion ebenfalls Magnetisierung auf den Liganden übertragen, welche nach Dissoziation von diesem mitgenommen wird. Durch die Bindung werden die Liganden also ebenfalls zu einem gewissen Grad gesättigt, was dazu führt, dass die gemessenen Intensitäten des Liganden abnehmen. Voraussetzung hierfür ist, dass der Ligand lange genug am Protein verweilt, um entsprechend Magnetisierung zu erhalten. Ist seine Residenzzeit zu kurz, oder, wie bei sehr starken bindenden Liganden, zu lange, wird nicht genug Magnetisierung übertragen. Darüber hinaus muss der Ligand im Vergleich zum Protein im Überschuss vorhanden sein, um einen messbaren Effekt zu erhalten. Das STD-NMR Experiment eignet sich demnach optimal für Liganden mit Affinitäten im oberen nM- bis einstelligen mM-Bereich, in dem auch die Affinitäten von Fragmenten typischerweise anzusiedeln sind. Darüber hinaus wirken sich größere Proteine positiv auf die Stärke des STD-Effekts aus, wobei Proteine eine molare Masse von mindestens 15 kDa besitzen sollten [191]. Neben dem *on*-Spektrum wird ein zweites Referenzspektrum, das *off*-Spektrum benötigt, um den STD-Effekt sichtbar zu machen. Hierbei wird weit weg von den Protein- und Ligand-Resonanzen in die

Probe eingestrahlt (40 ppm) und somit ein Spektrum erhalten in dem keine Sättigungseffekte mehr sichtbar sind. Die Intensität I_{on} des *on*-Spektrums werden daraufhin von der Intensität I_{off} des *off*-Spektrums abgezogen. Wird die Intensität I_{on} von Ligandsignalen durch Sättigung erniedrigt bleibt hierbei eine Differenz I_{diff} übrig. Bindet der Ligand nicht an das Protein, erfährt er keine Sättigung, wodurch die Differenz I_{diff} null beträgt. Die STD-Signalstärke wird im Verhältnis zur Intensität des I_{off} Spektrums angegeben:

$$std = \frac{I_{off} - I_{on}}{I_{off}} = \frac{I_{diff}}{I_{off}} \quad (2.28)$$

2.4.2 Röntgenstrukturanalyse

Die Röntgenstrukturanalyse ist die Methode der Wahl, wenn es darum geht, detaillierte dreidimensionale Strukturinformationen eines Proteins bzw. im Kontext des Wirkstoffdesigns eines Protein-Ligand-Komplexes zu erhalten. Kristallstrukturen von Proteinen bilden seit den Anfängen der Proteinkristallstrukturanalyse die Grundlage des rationellen Wirkstoffdesigns. Die Röntgenstrukturanalyse stellt daher eine der Hauptmethoden im Wirkstofffindungsprozess dar [193]. Eine detaillierte Ausführung dieser Methode, deren Ursprünge bis in das neunzehnte Jahrhundert zurückreichen, kann in dieser Arbeit nicht gegeben werden. Es kann hier nur auf einige grundlegende Aspekte der dahinterstehenden Theorie eingegangen werden. Der Leser sei hier auf die weiterführende Literatur verwiesen, auf deren Grundlage die folgenden Ausführungen aufbauen [194–196].

2.4.2.1 Theoretische Grundlagen

Ein Kristall setzt sich aus einer Vielzahl sogenannter Elementarzellen zusammen. Die Elementarzelle kann hierbei durch drei Vektoren \mathbf{a} , \mathbf{b} , \mathbf{c} und die dazugehörigen Winkel α , β , γ beschrieben werden. Hierbei lassen sich auf Grund von unterschiedlichen Symmetriebedingungen 14 sogenannte *Bravais*-Gitter ableiten. Ein räumliches Gitter lässt sich durch beliebig viele Netzebenen teilen. Die Schnittpunkte der Netzebenen mit den Koordinaten-Achsen werden über die sogenannten Millerschen-Indizes h, k, l dargestellt.

Das Prinzip der Röntgenstrukturanalyse fußt auf der Streuung von Röntgenstrahlung an den Elektronen eines Atoms bzw. genauer an den Elektronendichten der Orbitale. Das elektrische-Feld E eines Röntgenphotons (gedacht als Wellenpaket im Sinne des Welle-Teilchen-Dualismus) interagiert hierbei mit dem

elektrischen-Feld der Elektronen, durch Austausch von Photonen. Im Teilchenbild dargestellt induziert ein Photon der Röntgenstrahlung über sein elektrisches Feld ein Oszillieren der Elektronen. Am Punkt der Interaktion hört das Photon der Röntgenstrahlung auf „zu existieren“ und die Elektronen emittieren nun Photonen, die zueinander in Kohärenz stehen und somit miteinander nach dem Superpositionsprinzip interferieren können. Man unterscheidet hierbei zwischen der elastischen Streuung, bei der die emittierten Photonen die gleiche Energie wie die eingestrahlten Röntgenphotonen aufweisen, und der inelastischen Streuung, bei der ein Teil des Impulses auf ein Elektron übertragen wird, wodurch die emittierten Photonen weniger Energie als die eingestrahlten Photonen besitzen.

Sind die Atome bzw. Moleküle in einem Gitter angeordnet und besitzen somit periodisch wiederkehrende Strukturen beobachtet man aufgrund von Phasendifferenzen zwischen den emittierten Photonen die Ausbildung von Interferenzen. Die mathematische Beschreibung dieser Interferenzen liefert die Bragg-Gleichung. Sir William S. Bragg entwickelte aufbauend auf den Arbeiten von Max von Laue eine Gleichung mit der es möglich ist die Streuung an Kristallen als Reflexionen an Netzebenen zu beschreiben:

$$n\lambda = 2d_{hkl}\sin\theta \quad (2.29)$$

Hierbei ist λ die Wellenlänge der elektromagnetischen-Strahlung, n die Beugungsordnung, θ der Einstrahl- bzw. Austrittswinkel der Strahlung in Bezug zu den Netzebenen und d der Abstand zwischen den Netzebenen.

Das so erhaltene Beugungsbild besteht aus vielen Punkten (Reflexen) mit einer definierten Lage und Intensität. Jeder Reflex beinhaltet dabei Informationen über alle in der Einheitszelle vorhandenen Atome und korrespondiert zur Reflexion des einfallenden Röntgenstrahls an einer bestimmten Netzebene des Kristalls. Die Gitterkonstanten und Symmetrie eines Kristalls lassen sich anhand der Lage der Reflexe beschreiben. Die Positionen der Atome innerhalb der Elementarzelle können indirekt über die Intensität der Reflexe bestimmt werden.

Die totale Streuung eines Kristalls ist abhängig von der Anzahl der darin enthaltenen Atome:

$$F_S = \sum_{j=1}^{Atome} f_{S,j}^0 e^{2\pi i h x_j} \quad (2.30)$$

Der sogenannte Strukturfaktor F_S gibt hier die Summe aller Streuungen aller Atome j eines Kristalls an einem Punkt im reziproken Raum h (ausgedrückt über die

Millerschen-Indizes (h, k, l)) wider (i imaginärer Anteil). Atom j ist dabei durch seinen Formfaktor $f_{S,j}^0$ und der Position x_j ausgedrückt in fraktionellen Koordinaten, in der Einheitszelle definiert. Die betrachteten Atome sind hierbei nicht starre Punkte in einem Gitter, sondern oszillieren aufgrund der thermischen Energie um diesen. Dieses „Zittern“ um seine Mittelpunktposition wird durch den Temperaturfaktor (B-Faktor) B_j beschrieben

$$B_j = 8\pi^2 \overline{u_j^2} \quad (2.31)$$

mit der mittleren Abweichung $\overline{u_j^2}$ in \AA^2 . Mit diesem Faktor muss der Strukturfaktor F_S korrigiert werden. Der Einfachheit halber wird er nicht weiter in den Gleichungen mitbetrachtet.

Setzt man anstatt des Formfaktors $f_{S,j}^0$ die lokale Elektronen Dichtefunktion $\rho(x)$ ein, erhält man für den Strukturfaktor F_S :

$$F_{hkl} = V \cdot \int \rho(x) e^{2\pi i h x_j} dx \quad (2.32)$$

Die Elektronendichte eines Atoms in der Elementarzelle kann über ein Fourier-Integral beschrieben werden:

$$\rho(x)_j = \frac{1}{V} \int_h \int_k \int_l (|F_{hkl}|) \cdot e^{-2\pi i h x_j + i \alpha_{hkl}} \quad (2.33)$$

Diese Elektronendichte eines Atoms lässt sich demnach, wenn der Strukturfaktor F_{hkl} mit der dazugehörigen Phase α , das Volumen V der Elementarzelle, und die Millerschen Indizes (h, k, l) der Reflexe bekannt sind, über eine numerische Lösung des Integrals berechnen.

Das Zellvolumen kann über die Gitterkonstanten, die über die geometrische Anordnung der Reflexe bestimmt werden, berechnet werden. Der Strukturfaktor F_{hkl} wird über die Intensität der Reflexe bestimmt:

$$F_{hkl} = \frac{\sqrt{I}}{k} \quad (2.34)$$

mit der der Intensität I und einem Skalierungsfaktor k . Die Phasen α der einzelnen Reflexe sind aus dem Experiment nicht zugänglich und gehen verloren. Dieser unter dem Begriff des „Phasenproblems“ bekannte Sachverhalt wird über zusätzliche experimentelle Messungen (isomorphe Ersatz Methoden, SIR/MIR, oder Anomale Dispersion Methoden, SAD/MAD) oder ähnliche Strukturen im

molekularen Ersatz (*molecular replacement*, MR) gelöst. In dieser Arbeit kam der MR zum Lösen des Phasenproblems zum Einsatz.

Der MR basiert auf der von Patterson entwickelten Methode zur Ableitung eines Strukturmodells. Über die Pattersonfunktion ist es möglich die interatomaren Abstandsvektoren der Atome in einer Einheitszelle zu berechnen, die nur von den in den Intensitäten enthaltenen Strukturinformationen abhängen und keine zusätzliche Phaseninformation benötigen. Der so aufgespannte Pattersonraum enthält alle relativen atomaren Abstände ausgehend von einem „Ur“-Atom (wobei dies für jedes Atom in einem Molekül gilt, und somit ein Molekül mit N Elektronendichten einen Pattersonraum mit N^2 Pattersondichten erzeugt). Ausgehend von einer bekannten Struktur lässt sich deren Pattersonraum berechnen. Im molekularen Ersatz wird der experimentell ermittelte Pattersonraum mit dem aus einer bekannten Struktur errechneten Pattersonraum durch Rotations- und Translationsoperationen versucht maximal zur Deckung zu bringen [197]. Die Qualität der MR kann durch den R-Faktor wie folgt berechnet werden:

$$R = \frac{\sum_{hkl} ||\mathbf{F}_{Exp.}| - k|\mathbf{F}_{Modell}||}{\sum_{hkl} |\mathbf{F}_{Modell}|} \quad (2.35)$$

mit F den theoretischen bzw. experimentellen Strukturfaktoren und dem Skalierungsfaktor k . Es gilt hier: je größer der R-Faktor, desto mehr weicht das Modell vom Experiment ab.

Neben dem MR kam die anomale Dispersion zur Lokalisierung eines anomalen Streuers, in diesem Fall eines Bromatoms, in dieser Arbeit zum Einsatz. Einzelne Elektronen eines Atoms können durch die Röntgenstrahlung auf höhere Energieniveaus angehoben werden und durch Aussendung eines Röntgenphotons wieder relaxieren. Der Prozess wird auch als Röntgenfluoreszenz bezeichnet, wobei hier vor allem Elektronen der inneren Schalen angeregt werden. Hierzu muss die eingesetzte Energie der Röntgenstrahlung höher der entsprechenden Absorptionskante des Elements liegen.

Liegt die Energie unterhalb der Absorptionskante gilt für die elastische Streuung das Friedelsche Gesetz:

$$\begin{aligned} |\mathbf{F}_{hkl}| &= |\overline{\mathbf{F}_{hkl}}| \\ \alpha_{hkl} &= -\alpha_{hkl} \end{aligned} \quad (2.36)$$

welches besagt, dass das Beugungsbild immer zentrosymmetrisch ist. Der Strukturfaktor F_{hkl} ist also seinen zentrosymmetrischen Strukturfaktor $\overline{F_{hkl}}$ genau gleich und unterscheidet sich nur im Vorzeichen der Exponentialfunktion (Friedel-Paar). Das gleiche gilt für α . Liegt die Energie oberhalb der Absorptionskante regen die Röntgenstrahlen einige Atome zur Röntgenfluoreszenz an, wodurch der Formfaktor eines Atoms $f_{S,j}^0$ die zusätzlichen, von der Wellenlängen abhängigen, Faktoren f'_λ und f''_λ erhält. Diese setzen sich zur sogenannten Cromer-Mann Formel zusammen (i imaginärer Anteil):

$$f_{S,\lambda} = f_{S,j}^0 + f'_\lambda + i \cdot f''_\lambda \quad (2.37)$$

Durch den Term $i \cdot f''_\lambda$ wird das Friedel Gesetz verletzt mit dem Resultat, dass für ein Friedel-Paar die Strukturfaktoren F_{hkl} und $\overline{F_{hkl}}$ nicht mehr gleich sind, sondern eine Differenz ΔF aufweisen. Im Beugungsbild zeigen die Reflexe nicht mehr die gleiche Intensität. Anhand der Differenz ΔF können die Phasen α der Reflexe berechnet und somit die anomal streuenden Atome in der Elementarzelle lokalisiert werden.

2.4.2.2 Qualitätsbeurteilung von Datensätzen

Nach der Aufnahme eines Röntgenstrukturdatensatzes bedarf es der Korrektur der erhaltenen Rohdaten, um diese von experimentellen Fehlern wie z.B. Einflüssen des Detektors, Fluktuationen des Röntgenstrahls oder Volumenänderungen des Kristalls zu bereinigen. Der so erhaltene Datensatz wird im Anschluss auf seine Qualität über das mittlere Signal zu Rausch Verhältnis $\langle I/\sigma(I) \rangle$, der Multiplizität der Signale, der Vollständigkeit des Datensatzes und einer Reihe von R-Faktoren (R_{merge} , R_{meas} , $R_{p.i.m.}$) abgeschätzt [198]. Die genannten R-Faktoren berechnen sich hierbei im Einzelnen wie folgt:

$$R_{merge} = \frac{\sum_h \sum_i |I_{hi} - \langle I_h \rangle|}{\sum_h \sum_i \langle I_h \rangle} \quad (2.38)$$

$$R_{meas} = \frac{\sum_h \sum_i \left(\frac{n_h}{n_h - 1} \right)^{1/2} |I_{hi} - \langle I_h \rangle|}{\sum_h \sum_i \langle I_h \rangle} \quad (2.39)$$

$$R_{pim} = \frac{\sum_h \sum_i \left(\frac{1}{n_h - 1} \right)^{1/2} |I_{hi} - \langle I_h \rangle|}{\sum_h \sum_i \langle I_h \rangle} \quad (2.40)$$

mit I_{hi} der Intensität eines einzelnen Reflexes, $\langle I_h \rangle$ der mittleren Intensität aller symmetrieverwandten Reflexe und n der Anzahl aller symmetrieverwandten Reflexe.

2.4.3 Biolayer-Interferometrie (BLI)

Optische Biosensoren stellen ein vielseitiges Instrument in der Wirkstoffentwicklung dar. Seit der Einführung kommerzieller Biosensorplattformen Anfang der 1990er Jahre wächst die Zahl an technischen Systemen und deren Möglichkeiten stetig.

Eingesetzt werden sie vor allem in Primärscreenings sowie für bindungskinetische und thermodynamische Messungen in der darauf folgenden Optimierungsphase [199]. Im Kontext des FBDD sind hier vor allem die SPR- und BLI-Technik von Interesse. Wie oben bereits erwähnt stellt hier die SPR-Technik die dritthäufigste biophysikalische Methode im FBDD dar, die schon erfolgreich in zahlreichen Projekten zur Findung von potenten Fragmenten gegen Zielproteine eingesetzt wurde [200].

Mit nur 8 % nimmt die BLI-Technik eher eine Randstellung im FBDD ein, auch wenn ihr Einsatz in den letzten Jahren stetig leicht zugenommen hat [159]. Die Methode stellt eine noch relative junge Technologie dar, und hat sich vor allem im Bereich der Antikörperdetektion und der Impfstoffentwicklung zu einer Standardmethode entwickelt [201]. Während erste Plattformen noch nicht sensitiv genug waren, sind heutige Geräte mit Sensitivitäten bis zu 160 Da (in Abhängigkeit der von der Proteinmasse) in der Lage, auch die Bindung von Fragmenten zu detektieren. So kam die BLI-Methode schon erfolgreich in Fragmentscreenings mit den Proteinen JNK1, eIF4E oder BCL2 zum Einsatz [202]. Aufgrund des noch laufenden Patentschutzes der Technologie werden BLI-Geräteplattformen aktuell ausschließlich von der Firma Sartorius unter dem Namen Octet vertrieben. In der Patentschrift [203] wird das Messprinzip als *Phase-Shifting-Interferometrie* bezeichnet, die auf dem physikalischen Prinzip der Interferenzbildung von weißem Licht an dünnen Grenzflächen beruht und somit in die Kategorie der reflektorischen Interferenz-Spektroskopie eingeordnet werden kann. Im Folgenden soll detailliert auf die BLI-Technik eingegangen werden.

2.4.3.1 Theoretischer Hintergrund der BLI-Technik

Interferenz beschreibt das Phänomen der Änderung der Amplituden von sich überlagernden Wellen nach dem Superpositionsprinzip, die gemäß den Maxwell'schen-Gleichungen für Licht als elektromagnetische Welle gelten. Für zwei sich im Punkt P überlagernde monochromatische elektromagnetische Wellen gleicher Frequenz gilt somit für die elektrische Feldstärke E :

$$E = E_1 + E_2 \quad (2.41)$$

und somit für die Intensität I mit $I = |E|^2$:

$$I = |E_1 + E_2|^2 = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos \Delta\phi \quad (2.42)$$

wobei I_1 und I_2 die Intensitäten am Punkt P der beiden elektromagnetischen Wellen und $\Delta\phi = \phi_1 - \phi_2$ die Phasendifferenz zwischen beiden Wellen darstellt. Bedingung für die Interferenz ist, dass die Teilwellen zueinander kohärent, also eine feste Phasenbeziehung, besitzen müssen.

Um Interferenz zu erzeugen unterscheidet man die sogenannte Wellenfrontaufteilung und Amplitudenaufteilung. Unter Wellenfrontaufteilung würde z.B. das klassische Spaltexperiment fallen. Die reflektorische Interferenz-Spektroskopie basiert auf dem Prinzip der Amplitudenaufteilung [204].

Ein optischer Sensor, wie er in der BLI eingesetzt wird, kann nach [205–207] vereinfacht über zwei transparente planare Ebenen mit unterschiedlichen Brechungs-Indizes n_1, n_2 dargestellt werden (Abbildung 2.8). Der Lichtstrahl einer Strahlenquelle S mit der Wellenlänge λ wird am Punkt A unter dem Winkel α zu Punkt P partiell reflektiert. Der nicht reflektierte Teil durchläuft unter der Ablenkung β das Medium mit der Schichtdicke d und wird an der zweiten Ebene am Punkt B erneut partiell reflektiert. Er läuft über Punkt C unter erneuter Ablenkung ebenfalls zu Punkt P . Unter der Verwendung der Winkel α und β erhält man für die Strecken:

$$AB = BC = d / \cos \beta \quad (2.43)$$

$$AC = 2d \tan \beta \quad (2.44)$$

$$AD = AC \sin \alpha = 2d \tan \beta \sin \alpha \quad (2.45)$$

Damit ergibt sich für die optische Streckendifferenz Δs der beiden Teilstrahlen I_1 und I_2 :

$$\Delta s = n_2(AB + BC) - n_1 AD = 2n_2 d \cos \beta \quad (2.46)$$

Bei einem gleichbleibenden Einstrahlwinkel α ist β konstant und Δs somit nur von der Schichtdicke d und dem Brechungsindex n_2 abhängig. Gemäß der Stokes-Beziehung kommt die Phasendifferenz von 2π hinzu, wodurch sich die Intensität I in Abhängigkeit der Wellenlänge λ am Punkt P wie folgt berechnen lässt:

$$I(\lambda) = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos\left(\frac{2\pi \Delta s}{\lambda}\right) \quad (2.47)$$

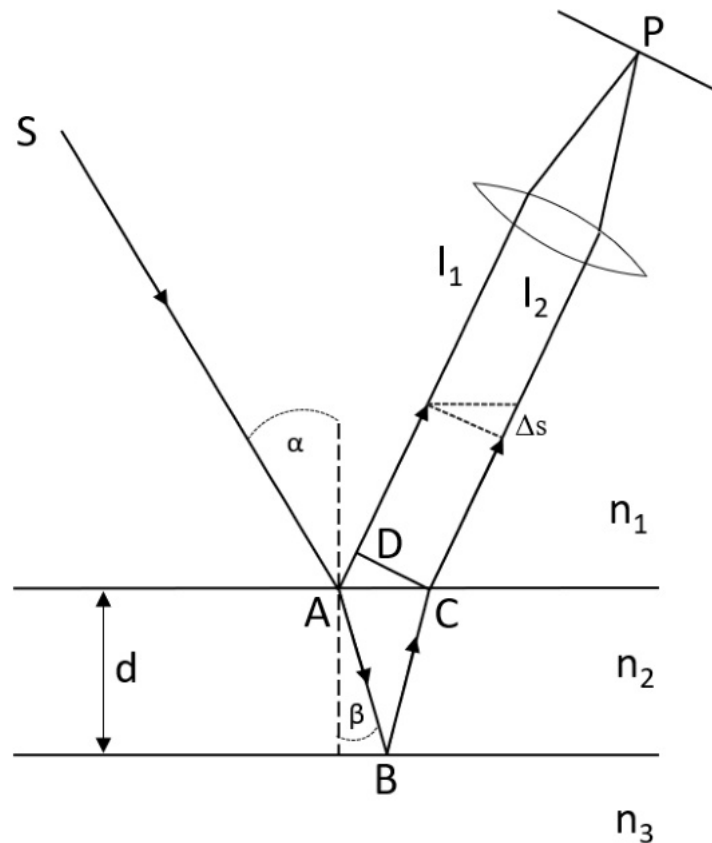


Abbildung 2.8: Interferenz an plan-parallelen Ebenen
 Abbildung modifiziert nach [205]

Abbildung 2.9 zeigt den Aufbau und Strahlengang eines Sensors wie er in Octet Geräten verwendet wird [203]. In einer Lichtquelle **A** wird weißes Licht über ein breites Spektrum erzeugt und mittels eines Optokopplers in den Sensor geleitet. Der Sensor kann nun in folgende Bereiche unterteilt werden:

- **B** beinhaltet eine optische Glasfaser, die die Aufgabe hat, das Licht zur eigentlichen Sensorfläche weiterzuleiten.
- **C** beinhaltet die erste reflektorische Fläche, an der der Referenzstrahl erzeugt wird. Sie besteht aus einem Material mit einem höheren Brechungsindex als das darauf folgende optische Element **D**. Als Beispiel wird hier Ta_2O_5 mit einem Brechungsindex von 2.1 genannt. Die Schichtdicke beträgt in diesem Fall 5 bis 30 nm.
- **D** und **E** stellen das Herzstück des Sensors dar. Bereich **D** besteht aus SiO_2 -Glas mit einem Brechungsindex von 1.4 bis 1.5. Des Weiteren kommen auch transparente Polymere, wie Polystyrene oder Polyethylene zum Einsatz. **E** stellt den Analyt-bindenden Bereich dar. Er besteht aus bifunktionalen Molekülen, die auf der einen Seite z.B. über Siloxangruppen an die SiO_2 -Seite binden und auf der anderen Seite über Hydroxyl-, Amin-, Carboxyl- oder andere reaktive Gruppen eine Verbindung zu einem Biomolekül herstellen. Beide bilden zusammen ein optisches Element mit der Schichtdicke d und dem Brechungsindex n_2 . Bei Bindung eines Liganden an ein an **E** immobilisiertes Biomolekül kann es zu einer Änderung Δd der Schichtdicke sowie des Brechungsindex n_2 kommen. Dies führt zu einer Phasenverschiebung Δs zwischen dem Referenzstrahl $I_{Ref.}$ und dem Signalstrahl $I_{Sig.}$ bei einer Wellenlänge λ , und somit zu einer Änderung in der Intensität in Punkt **F**. Für den gesamten Wellenlängenbereich des eingesetzten weißen Lichts erhält man für unterschiedliche Wellenlängen unterschiedliche Intensitäten, die am Punkt **F** zu dem in **(b)** gezeigten Kurvenverlauf mit unterschiedlichen Minima und Maxima bei unterschiedlichen Wellenlängen λ führen. Die Intensitäten in **F** für verschiedene Wellenlängen λ werden über einen CCD- oder CMOS-Chip nach vorheriger Aufspaltung des Spektrums über ein Spektrometer erhalten. Bei Bindung eines Liganden an den Sensor kommt es durch die wellenlängenabhängige Phasenverschiebung zu einer Verschiebung der beobachteten Minima und Maxima im Spektrum. Diese **Verschiebungen** λ in nm ist das Messsignal der Biolayer-Interferometrie-Methode. Die Höhe der Verschiebung hat hier nichts mit der Signalstärke gemein, weswegen dieser Begriff hier falsch wäre.

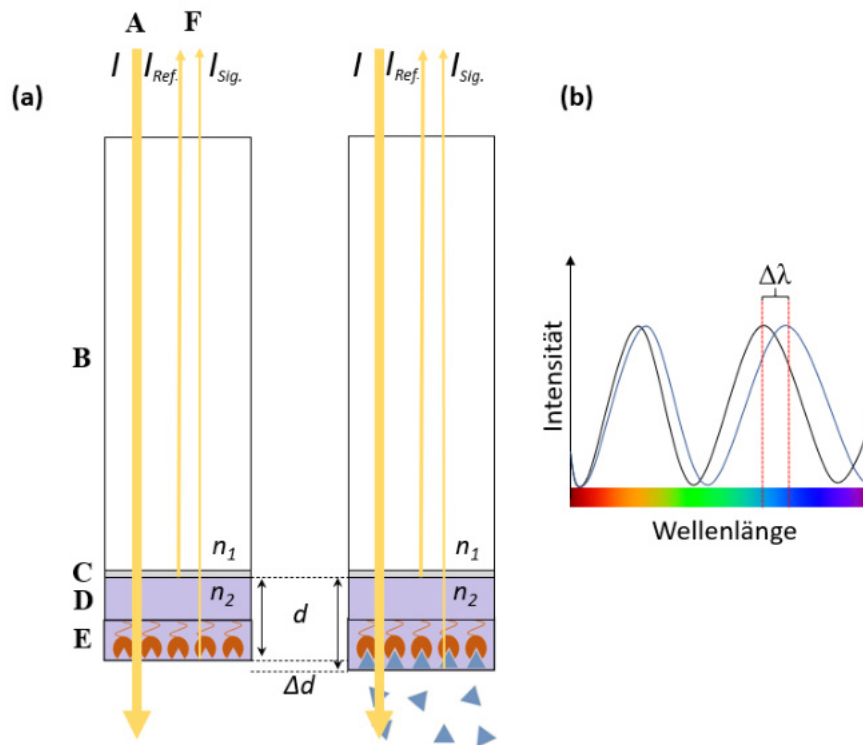


Abbildung 2.9: Aufbau und Funktionsweise der optischen Sensoren der Biolayer-Interferometrie

Gezeigt ist der Strahlengang durch einen optischen Sensor wie er in den Octet Geräten Verwendung findet. Weißes Licht (A) wird durch eine Glasfaser zum eigentlichen optischen Element (C,D,E) geleitet, und durch partielle Reflexionen an den enthaltenen Grenzflächen in den Referenz und den Signalstrahl aufgespalten. Hierdurch kommt es zur Interferenzbildung zwischen den beiden Strahlen. (b) Bindet ein Ligand an den Sensor, erhöht sich der Abstand d bzw. der Brechungsindex n_2 , wodurch es zu einer quantifizierbaren Verschiebung des Interferenzmusters kommt. Für Details siehe Text.

Ein BLI-Experimente lässt sich in zwei Schritte einteilen: in eine Assoziationsphase, an der ein mit Protein beladener Sensor in einen Napf mit einer Ligandlösung eingebracht wird und der Ligand an das Protein bindet, sowie in eine Dissoziationsphase, in der der Sensor mit dem gebildeten Protein-Ligand-Komplex in einen Napf ohne Ligandlösung eingebracht wird und das Molekül wieder von dem Protein dissoziiert. Beide Prozesse lassen sich, wie bei der SPR-Spektroskopie, in Echtzeit verfolgen und bieten dadurch direkten experimentellen Zugang zu den kinetischen Parametern k_d (Geschwindigkeitskonstante der Dissoziation), k_a (Geschwindigkeitskonstante der Assoziation) und k_{obs} (empirische Geschwindigkeitskonstante) und somit zu der Affinität K_D , der Bindungsreaktion.

Vorteile der BLI-Technik im Vergleich zur SPR-Technik ist das Fluid-freie Messsystem. Dadurch, dass der Sensor direkt in einen Reaktionsnapf eingetaucht wird

und keine Durchflusszelle wie in der SPR-Technik existiert, ist das System unempfindlich gegenüber den eingesetzten Puffern. Es können sogar Messungen in Blutseren und Zell-Lysaten durchgeführt werden. Auch bedarf es bei BLI-Messungen keiner Korrektur der BrechungsIndizes der eingesetzten Pufferlösungen. Im Falle der Sensitivität für kleine Moleküle liegt allerdings die SPR-Technik im Vorteil mit einem Detektionslimit bis zu 60 Da [208].

2.4.3.2 Gleichgewichts- und kinetische Analyse von Protein-Ligand-Interaktionen im Kontext der BLI

Auch wenn sich das physikalische Messprinzip der BLI-Methode von der SPR-Methode unterscheidet, so sind die Vorgänge, die an der Sensoroberfläche ablaufen, die gleichen und können mit Adsorptionsvorgängen an Oberflächen beschrieben werden. Das hier zu Grunde liegende Phänomen der Adsorption bedarf einer festen Phase (Adsorbant) und einem in Lösung befindlicher Liganden, der an die feste Phase bindet. Die feste Phase wird hierbei durch das auf dem Sensor immobilisierte Molekül gebildet, an den spezifisch oder unspezifisch die in Lösung befindlichen Ligandenmoleküle binden. Die an der festen Phase stattfindenden Vorgänge lassen sich hier in die folgenden Vorgänge einteilen. Bei Kontakt der festen Phase mit dem in Lösung befindlichen Liganden wird dieser an die feste Phase adsorbiert. Dies geschieht solange, bis sich ein Gleichgewicht zwischen der Assoziation (Adsorption) und der gleichzeitig stattfindenden Dissoziation einstellt. Das Gleichgewicht ist dabei nur von der Menge an fester Phase und von der Konzentration des Liganden abhängig. Alleinige Dissoziationsprozesse des adsorbierten Liganden können durch den Austausch des Mediums mit Ligand-freiem Medium beobachtet werden. Zur Beschreibung dieser an der Sensoroberfläche stattfindenden Prozesse existieren mathematische Modelle (von einfachen 1:1-Bindungsmodellen bis hin zu komplexeren Bindungsmodellen), mit deren Hilfe diese Vorgänge beschrieben werden können [209].

2.4.3.3 1:1-Bindungsmodell

Das 1:1-Bindungsmodell zwischen einem Ligand und einem Protein stellt das einfachste Modell dar und lässt sich mittels folgender Reaktionsgleichung beschreiben:



mit dem immobilisiertem Protein P und dem frei in Lösung vorliegendem Liganden L , sowie dem Protein-Ligand-Komplex PL . Die ablaufende Reaktion lässt

sich mittels der folgenden Differentialgleichungen beschreiben [210]:

$$\begin{aligned} \text{Assoziation : } \frac{d[PL]}{dt} &= k_a[P][L] - k_d[PL] \\ \text{Dissoziation : } \frac{-d[PL]}{dt} &= k_d[PL] \end{aligned} \quad (2.48)$$

Die in der Dissoziation mögliche erneute Bindung von frei gewordenen Liganden an das immobilisierte Protein wird hierbei vernachlässigt. Die Lösung der Differentialgleichung 2.48 für die Assoziationsphase stellt eine Exponentialfunktion folgender Form dar [211]:

$$R_t = \frac{k_a[L]B_{max}}{k_a[L] + k_d} (1 - e^{-(k_a[L]+k_d t)}) \quad (2.49)$$

Experimentell ist die Geschwindigkeitskonstante k_a in der Assoziation nicht direkt zugänglich, sondern es wird hier die empirische Geschwindigkeitskonstante k_{obs} beobachtet. Die empirische Geschwindigkeitskonstante k_{obs} stellt eine Kombination aus Assoziationsprozessen und Dissoziationsprozessen dar und beschreibt die zur gleichen Zeit ablaufende Bildung des Protein-Ligand-Komplex sowie dessen gleichzeitig stattfindende Zerfall. Hierdurch vereinfacht sich Gleichung 2.49 zu:

$$R_t = A(1 - e^{-(k_{obs}t)}) \quad (2.50)$$

Die Asymptote A beinhaltet dabei den schon bekannten Term:

$$A = \frac{k_a[L]B_{max}}{k_a[L] + k_d} \quad (2.51)$$

Zwischen der empirischen Geschwindigkeitskonstante k_{obs} und der Ligandenkonzentration $[L]$ besteht in einem 1:1-Modell hierbei folgender linearer Zusammenhang:

$$k_{obs} = k_a[L] + k_d \quad (2.52)$$

Über die Konzentration $[L]$ des Liganden und die dazugehörige gemessene empirische Geschwindigkeitskonstante k_{obs} sind somit die Geschwindigkeitskonstanten k_a und k_d zugänglich [211].

In gleicher Weise lässt sich für ein 1:1-Modell der Dissoziationsprozess über eine Zerfallsexponentialfunktion erster Ordnung beschreiben:

$$R_t = R_{eq}(e^{-(k_d t)}) \quad (2.53)$$

Sie ist nur von der Geschwindigkeitskonstanten k_d abhängig und somit direkt aus dem Experiment zugänglich. Anhand der Geschwindigkeitskonstanten k_a und k_d kann die Affinität K_D der Bindung berechnet werden:

$$K_D = \frac{k_d}{k_a} \text{ bzw. } K_D = \frac{k_d([L] + k_d)}{k_{obs}} \quad (2.54)$$

Für ein 1:1-Modell lässt sich die Affinität auch anhand der Gleichung 2.55 berechnen, wenn sich das System im chemischen Gleichgewicht befindet. Das Erreichen des chemischen Gleichgewichts zeigt sich hierbei im Sensorgramm in der Ausbildung einer Plateauphase.

$$K_D = \frac{[P][L]}{[PL]} \quad (2.55)$$

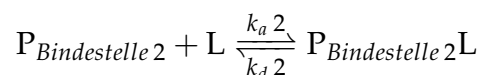
Die Konzentration $[P]$ stellt dabei die Menge an immobilisiertem Protein auf dem Sensor dar und besitzt eine maximale Bindungskapazität, die die maximale Verschiebung R_{max} hervorruft (Besetzung aller Ligandenbindestellen). Misst man bei unterschiedlichen Ligand-Konzentrationen $[L]$, so lässt sich die Affinität über eine Langmuir-Bindungsisotherme berechnen:

$$R_{eq} = \left(\frac{n[L]}{[L] + K_D} \right) R_{max} \quad (2.56)$$

Hierbei stellt n die Stöchiometrie der Protein-Ligand-Bindung dar.

2.4.3.4 Komplexere Bindungsmodelle

Neben dem einfachen 1:1-Modell kann als nächstes ein Szenario betrachtet werden, bei dem neben einer Bindestelle eine oder mehrere weitere Bindestellen in einem Molekül existieren. Zunächst sei hier der Fall betrachtet, bei dem zwei unterschiedliche Bindestellen vorhanden sind (2:1-Modell). Vorausgesetzt die Bindung an diese beide Bindestellen ist unabhängig voneinander, lässt sich ein solcher Prozess in je eine Reaktionsgleichung pro Bindestelle zerlegen:



Die Gesamtreaktion ist hierbei die Summe beider Teilreaktionen, womit sich für die Assoziationsphase ergibt:

$$R_t = A_1(1 - e^{-(k_{obs1}t)}) + A_2(1 - e^{-(k_{obs2}t)}) \quad (2.57)$$

und für die Dissoziationsphase:

$$R_t = R_{eq}(e^{-(k_{d1}t)}) + R_{eq}(e^{-(k_{d2}t)}) \quad (2.58)$$

Kurven solcher Bindungen zeigen im Sensorgramm einen biphasigen Verlauf.

Im chemischen Gleichgewicht lässt sich die Reaktion ebenfalls als Summe zweier 1:1-Langmuir-Modelle beschreiben:

$$R_{eq} = \left(\frac{n_1[L]}{[L] + K_{D1}}\right)B_{max_1} + \left(\frac{n_2[L]}{[L] + K_{D2}}\right)B_{max_2} \quad (2.59)$$

Verallgemeinert heißt dies für die Anzahl i an Bindestellen für die Assoziations-, bzw. Dissoziationsphase:

$$R_t = \sum A_i(1 - e^{-(k_{obs_i}t)}) \quad (2.60)$$

$$R_t = \sum R_{eq_i}(e^{-(k_{d_i}t)})$$

Und für das chemische Gleichgewicht:

$$R_{eq} = \sum \left(\frac{n_i[L]}{[L] + K_{D_i}}\right)R_{max_i} \quad (2.61)$$

Ein bekanntes Beispiel für solche unabhängigen Bindestellen ist das Humane Serumalbumin (HSA). HSA besitzt vier unabhängige Bindestellen, an die unterschiedliche Klassen von Wirkstoffen binden und mithilfe solcher Modelle beschrieben werden können [212].

Daneben existieren Modelle zur Einbeziehung von Konformationsänderungen, die bei der Bindung eines Liganden hervorgerufen werden. Hierbei wird davon ausgegangen, dass ein Ligand zunächst schwach an seine Zielstruktur bindet. Durch die Bindung wird eine strukturelle Änderung hervorgerufen, in der eine neue Konformation eingenommen wird, die eine höhere Affinität für den Liganden besitzt [213].



Mit $K_D = k_d/k_a$ und $K_{conf} = k_{conf}/k_{-conf}$ erhält man hier die beobachtete Affinität:

$$K_{obs} = K_D(1 + K_{conf}) \quad (2.62)$$

Der Vollständigkeit halber sei hier noch das bivalente Modell, das vor allem bei Antikörperbindungen eingesetzt wird, erwähnt [214]. Neben diesen auf der Bindung selbst beruhenden Modellen gibt es weitere Effekte, die die Kinetik und das Erscheinungsbild der Sensorgramme beeinflussen können. Bei Massentransporteffekten beeinflusst die Diffusion eines Liganden aus der Bulk-Lösung die Kinetik der Bindung. Vereinfacht ausgedrückt laufen hier zwei Schritte ab.



Zunächst muss der Ligand durch Diffusion zu dem immobilisierten Molekül auf der Sensoroberfläche gelangen, um dort in einer zweiten Reaktion eine Bindung mit diesem einzugehen. Ist nun die Geschwindigkeit, mit der der Ligand zum Sensor diffundieren kann kleiner als die Geschwindigkeit, mit der der Ligand an das immobilisierte Molekül binden kann, entsteht in der Nähe der Sensoroberfläche ein Konzentrationsdefizit des Liganden. Unter diesen Bedingungen ist die Bindungskinetik nur von der Massentransferrate k_m des Liganden abhängig und nicht mehr von der wahren Bindungskinetik k_a bzw. k_d . Für ein 1:1-Bindungsmodell spielt in der Assoziationsphase neben Gleichung 2.48 auch Gleichung 2.63 für den Massentransport eine Rolle [213, 215].

$$\frac{d[L_{\text{Bulk}}]}{dt} = -k_m[L_{\text{Bulk}}] + k_m[L] \quad (2.63)$$

Womit sich die Assoziation durch die folgende Differentialgleichung beschreiben lässt:

$$\frac{d[L]}{dt} = k_m[L_{\text{Bulk}}] - k_m[L] - k_a[P][L] + k_d[PL] \quad (2.64)$$

In gleicher Weise beeinflusst der Massentransport-Effekt auch die Dissoziationsphase. Hierbei spricht man von transportlimitierter Dissoziation. Der Ligand kann hier nicht schnell genug von seiner Zielstruktur weg diffundieren und geht eine erneute Bindung mit diesem ein. Gleichung 2.65 zeigt ein Modell zu Beschreibung solcher Effekte [215].

$$\frac{d[R(t)]}{dt} = \frac{-k_d R(t)}{1 + \frac{k_a}{k_m}(R_{\text{max}} - R(t))} \quad (2.65)$$

Analyse-Programme, wie für das Octet-System, verfügen über die Möglichkeit durch numerische Integration solche Differentialgleichungen zu lösen und Sensorgramme mit Massentransport-Effekten auszuwerten. Die hier vorgestellten Gleichungen bilden die Grundlage aller später durchgeführten Auswertungen der BLI-Messungen.

Kapitel 3

Zielsetzung dieser Arbeit

Das Ziel dieser Arbeit ist die Identifizierung von Verbindungen, die innerhalb der SHP-Motiv Bindestelle der p97-N Domäne binden und Startpunkte für die Entwicklung eines Inhibitors von Protein-Protein-Interaktionen mit p97 Kofaktoren bilden. Als Strategie wird hierfür ein fragmentbasierter Ansatz gewählt, der sowohl in Form eines virtuellen- als auch mittels eines biophysikalischen Screenings verfolgt wird. Ein weiteres Ziel ist eine nähere Charakterisierung der Bindung des Naturstoffes Xanthohumol an die N-Domäne von p97. Xanthohumol ist derzeit der einzige bekannte p97-Inhibitor mit einem vermutlich über die N-Domäne vermittelten Wirkmechanismus.

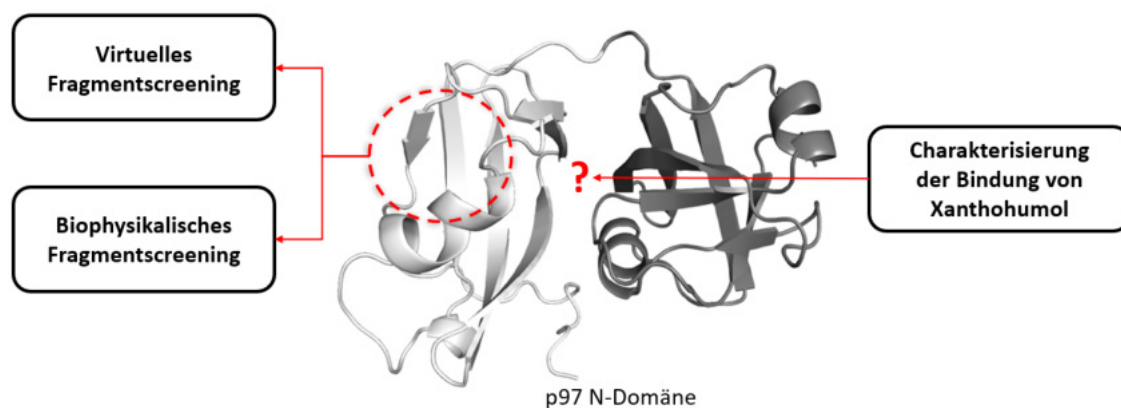


Abbildung 3.1: Zielsetzung der Arbeit

Zur Erreichung der Ziele wird mit Hilfe computergestützter Verfahren eine strukturbasierte Charakterisierung der SHP-Bindestelle vorgenommen, eine virtuelle Fragmentbibliothek zusammengestellt und ein virtuelles Screening unter Verwendung von Pharmakophormodellen, Docking-Verfahren und Molekulardynamik-Simulationen durchgeführt. Für das biophysikalische Screening wird unter Einsatz der Biolayer-Interferometrie ein Assay etabliert, der bindende Verbindungen einer Fragmentbibliothek identifizieren, sowie die Affinitäten dieser Verbindungen bestimmen soll. Zur Bestätigung der Biolayer-

Interferometrie Ergebnisse wird eine orthogonale Messmethode unter Verwendung der STD-NMR-Spektroskopie eingesetzt.

Die Bindestellen von Fragmenten sollen unter Einsatz der Röntgenstrukturanalyse, computerbasierter *mixed-solvent* Simulationen und STD-NMR-Messungen eingegrenzt bzw. aufgeklärt werden.

Kapitel 4

Materialen, Geräte und Software

4.1 Software

Tabelle 4.1: In dieser Arbeit verwendete Software

Bezeichnung	Version	Referenzen
AMBER (Amber18 & AmberTools18)	2018	[216]
ANODE	2013/1	[217]
BUSTER	2.10.3	[218]
CCP4i Suite	7.1	[219]
Coot	0.9.1	[220]
DSX	0.89	[109]
ForteBio Data Analysis	Gerätespezifische Versionen	[221]
Gaussian09	09 Revision C.01	[222]
GOLD	5.2.2	[100]
GROMACS	2019.01	[223]
MarvinSketch	17.28.0	[224]
Microsoft Excel		[225]
MOE (Molecular Operating Environment)	2019.01	[92]
MoKa (Molecular Discovery)	3.1.3	[226–228]
Origin Pro	2019b	[229]
Phenix Suite	1.18.2	[230]
PyMOL	2.3.2	[231]
R	3.6.3	[232, 233]
RDKit	Release 2018.09.1	[234]
SHELX C/D	2016/1	[235]
TopSpin	4.0.8	[236]
VMD	1.9.3	[237]
XDS	31.1.2020	[238]

4.2 Verwendete Puffer

Tabelle 4.2: Verwendete Puffer

Bezeichnung	Zusammensetzung	pH
Elutions-Puffer	50 mM Tris/HCl 500 mM KCl 5 mM MgCl ₂ 5% (v/v) Glycerol 5 mM β-Mercaptoethanol 250 mM Imidazol in H ₂ O	8
Lyse- und Wasch-Puffer	50 mM Tris/HCl 500 mM KCl 5 mM MgCl ₂ 5% (v/v) Glycerol 5 mM β-Mercaptoethanol in H ₂ O	8
NMR-Puffer	NaCl 25 mM K ₃ PO ₄ 35 mM in D ₂ O (99.9%)	7.4 (pD)
PBS-Puffer	NaCl 137.0 mM Na ₂ HPO ₄ (Dihydrat) 8.1 mM KCl 2.7mM KH ₂ PO ₄ 1.5 mM in H ₂ O	7.4
Puffer ^{BLI-1}	0.05% (v/v) Polysorbat 20 1 mM DTT in PBS-Puffer	7.4
Puffer ^{BLI-2}	5% (v/v) DMSO in Puffer ^{BLI-1}	7.4
Puffer ^{BLI-3}	0.05% (v/v) Polysorbat 20 5 mM DTT in PBS-Puffer	7.4
Puffer ^{BLI-4}	5% (v/v) DMSO in Puffer ^{BLI-3}	7.4
SEC- und Dialyse-Puffer	50 mM Tris/HCl 150 mM KCl 5 mM MgCl ₂ 5% (v/v) Glycerol 5 mM β-Mercaptoethanol in H ₂ O	8

4.3 Geräte

Tabelle 4.3: Verwendete Geräte

Geräteart	Bezeichnung	Hersteller
BLI-Plattformen	Octet K2 (Leihgerät Firma Sartorius)	ForteBio (jetzt Sartorius)
	Octet 96RED (Universität Würzburg)	ForteBio (jetzt Sartorius)
	Octet 96REDe (Universität Bergen)	ForteBio (jetzt Sartorius)
Cryo-Loop	Cryo-Loop	Hampton Research
Kristallisationsinkubator	ROCK-IMAGER-1000	Formulatrix
Mikroplatten Reader	CLARIOstar	BMG LABTECH
Mikroskop	SteREO Discovery.V12	ZEISS
Mikroskopkamera	AxioCam MRc	ZEISS
NMR-Spektrometer	Avance 400 MHz	Bruker
	mit BBO BB-H 5 mm Probenkopf und BCU-05 Temperatur Kontrolleinheit	
Pipettierroboter	NT8-Kristallisationsroboter	Formulatrix
	Honeybee 963	Digilab
	Mosquito HV (Universität Bergen)	TTP Labtech
Proteinaufreinigung	Äkta Chromatographiesysteme	GE Healthcare (jetzt Cytiva)
Puck-System	UniPuck	MiTeGen
Rotor (8x 50 mL)	JA-25.50	Beckman Coulter
Rotor (4 x 2250 mL)	JS-5.0	Beckman Coulter
Spektrophotometer	NanoDrop ND 1000	Peqlab
Zellaufschluss System	M-110P	Microfluids

4.4 Chemikalien

Tabelle 4.4: Chemikalien

Substanz (p.a. grade)	Lieferant
Acrylamide/Bis-acrylamide (37.3:1)	Carl Roth
Adenosin 5'-diphosphat-Natriumsalz	Sigma-Aldrich
Adenosin 5'-triphosphat-Dinatriumsalz (Hydrat)	Sigma-Aldrich
Adenylyl-imidodiphosphat-Tetralithiumsalsz	Sigma-Aldrich
Ammoniumpersulfat (APS)	Carl Roth
Ampicillin Natriumsalz	Carl Roth
Biocytin	Sigma-Aldrich
Biotin	Sigma-Aldrich
Chloramphenicol	Carl Roth
Coomassie Brilliant Blue G-250	Carl Roth
Dimethylsulfoxid (DMSO)	Carl Roth
Dimethylsulfoxid deuteriert	Deutero
Dinatriumhydrogenphosphat	Carl Roth
Dithiothreitol (DTT)	Carl Roth
Ethanol	Carl Roth
Glyzerin	Carl Roth
Imidazol	Carl Roth
Isopropyl- β -D-thiogalactopyranosid (IPTG)	Carl Roth
Kalilauge	Carl Roth
Kaliumchlorid	Carl Roth
Kaliumdihydrogenphosphat	Carl Roth
Kaliumphosphat (Wasserfrei)	Sigma Aldrich
Kanamycinsulfat	Carl Roth
Magnesiumchlorid (Hexahydrat)	Carl Roth
Magnesiumchlorid (Wasserfrei)	Sigma Aldrich
β -Mercaptoethanol	Carl Roth
Natriumchlorid	Carl Roth
Natriumdodecylsulfat (SDS)	Carl Roth
Natriumformiat	Sigma Aldrich
Natronlauge	Carl Roth
Polyethylenglykol (PEG 600)	Sigma Aldrich
Polysorbat 20 (Tween 20)	Sigma Aldrich
Salzsäure (HCl bzw. DCl)	Carl Roth bzw. Sigma Aldrich
Tris-(hydroxymethyl)-aminoethan (Tris)	Carl Roth
Wasser deuteriert (99.9%)	Deutero
Xanthohumol	MedChemExpress Europe

Tabelle 4.5: Screening Verbindungen

Katalog Nr.	Interne ID	Lieferant
1023215	TROLL16	OTAVACHemicals
1044950	TROLL19	OTAVACHemicals
1047688	TROLL15	OTAVACHemicals
1055871	TROLL7	OTAVACHemicals
1120063	TROLL11	OTAVACHemicals
1149690	TROLL6	OTAVACHemicals
1196351	TROLL8	OTAVACHemicals
1202258	TROLL12	OTAVACHemicals
1207432	TROLL10	OTAVACHemicals
1523268	TROLL9	OTAVACHemicals
1524668	TROLL14	OTAVACHemicals
1582045	TROLL18	OTAVACHemicals
1770375	VIK50	OTAVACHemicals
1799671	VIK40	OTAVACHemicals
1909733	TROLL2	OTAVACHemicals
2101165	TROLL13	OTAVACHemicals
2204975	TROLL4	OTAVACHemicals
2659358	TROLL1	OTAVACHemicals
4347292	VIK1	OTAVACHemicals
125160275	TROLL17	OTAVACHemicals
7212060151	VIK20	OTAVACHemicals
Z104639526	SHP11	Enamine
Z1145207891	SHP6	Enamine
Z1350606470	SHP12	Enamine
Z1565851368	SHP5	Enamine
Z1694842016	SHP3	Enamine
Z1856567394	SHP4	Enamine
Z31231940	SHP9	Enamine
Z32368274	SHP10	Enamine
Z415728082	SHP2	Enamine
Z446358664	SHP1	Enamine

4.5 Spezielle Verbrauchsmaterialien

Tabelle 4.6: Labeling-Kits, Enzyme und Sonstiges

Bezeichnung	Lieferant
3C Protease	in Haus Produktion
BirA Enzym	in Haus Produktion
Complete EDTA-free Protease Inhibitor-Tabletten	Roche
DNase I	Invitrogen
LB Medium	Carl Roth
Monolith His-Tag Labeling Kit RED-trisNTA 2 nd Generation	Nanotemper

Tabelle 4.7: Sonstiges Verbrauchsmaterial

Bezeichnung	Lieferant
Amicon Ultra Centrifugal Filter (10 und 30 kDa MWCO)	Merck
Crystalgen SuperClear 24 Napf Platen	Jena Bioscience
Dialyse Membran Spectra/Por	Spectrum Laboratories
Glas Abdeckplättchen (22 mm) silikonisiert	Jena Bioscience
HiLoad® 75/200 16/600 Superdex® <i>prep grade</i> Säulen	GE Healthcare (jetzt Cytiva)
INTELLI-96-3 Kristallisationsplatten	Art Robbins Instruments
Microplatten, 96 Näpfchen, F-Boden (Kaminform), schwarz	Greiner Bio-One
Protino™ Ni-TED Harz	Macherey-Nagel

Kapitel 5

Methoden

Im Folgenden werden die in dieser Arbeit eingesetzten Methoden beschrieben. Wo erforderlich werden theoretische Hintergründe, die noch nicht im allgemeinen Theorie-Teil erläutert wurden, den jeweiligen Methoden vorangestellt erläutert.

5.0.1 Abbildungen

Alle molekularen Abbildungen wurden in Pymol [2.3.2] [231] erstellt, mit Ausnahme der mit MOE [2019.01] [92] erstellten Darstellungen der Interaktionspotential-Analysen und der Pharmakophormodelle. Zweidimensionale Moleküldarstellungen wurden in MarvinSketch [17.28.0] [224] erstellt. Alle Graphen wurden mit R [3.6.3] [232, 233] generiert.

5.1 Computerbasierte Methoden

5.1.1 Virtuelle Fragmentbibliothek

5.1.1.1 Auswahl der Fragmente für die Fragmentbibliothek

In dieser Arbeit wurde eine virtuelle Fragmentbibliothek zusammengestellt, die die Grundlage des später durchgeführten virtuellen Screenings bildete. An die Moleküle dieser Bibliothek wurde die Anforderung der schnellen und einfachen Verfügbarkeit gestellt, um vielversprechende Fragmente direkt bestellen und vermessen zu können. Der gewählte fragmentbasierte Ansatz ließ die Wahl weitestgehend auf von Anbietern bereits zusammengestellte Fragmentbibliotheken fallen. Tabelle 5.1 zeigt eine Übersicht der für diese Arbeit relevanten und kommerziell verfügbaren Fragmentbibliotheken. Bei der Auswahl dieser Fragmentbibliotheken wurden mit in Betracht gezogen, dass für spätere experimentelle Untersuchungen eine möglichst hohe Löslichkeit gegeben sein sollte. Daher wurden

das *Essential-Fragment-Set* (Enamine) mit experimentell bestimmten Löslichkeiten von mindestens 1 mM (Wasser), sowie das *Carboxylic-Acids-Fragment-Set* (Enamine) aufgrund der löslichkeitsverbessernden Carbonsäurefunktionen ausgewählt. Die *Golden-Fragment-Library* und *Ro3-Fragment-Library* von Enamine, sowie die Fragmentbibliothek von Life Chemicals wurden als *Ro3*-basierte Fragmentbibliotheken ausgewählt. Die *PPI-Fragment-Library* von Enamine und Life Chemicals wurden auf der Grundlage der *Ro4* zusammengestellt und daher mit in Betracht gezogen [152][153]. Die in Tabelle 5.1 mit aufgeführte Fragmentbibliothek der Universität Bergen (Norwegen) (UiB) wurde nicht im virtuellen Screening eingesetzt, sondern nur für computerbasierte retrospektive Analysen des biophysikalischen Fragment screenings herangezogen.

Tabelle 5.1: Fragmentbibliotheken

Bibliothek	Fragmente	Anbieter	Verwendung	Stand
Carboxylic Acids Fragment Set	4350	Enamine	Virtuelles Screening SHP-Bindestelle (VS-SHP)	23.5.2018
Essential Fragment Set	191	Enamine	VS-SHP	23.5.2018
Golden Fragment Library	1794	Enamine	VS-SHP	23.5.2018
PPI Fragment Library	3600	Enamine	VS-SHP	23.5.2018
Ro3 Fragment Set	69583	Enamine	VS-SHP	23.5.2018
Fragment Library	4938	Life chemicals	VS-SHP	25.5.2018
PPI Fragment Library	2447	Life chemicals	VS-SHP	16.11.2017
Fragmentbibliothek	678	UiB (Norwegen) ^a	<i>mixed-solvent</i> MD-Simulationen & biophysik. Screening	26.9.2019 ^b

^a bezogen von Otava Chemicals

^b Stand am Tag des durchgeführten Screenings mit der ND1-Domäne

5.1.2 Aufbereitung der zusammengestellten Fragmentbibliothek

Die zusammengestellte virtuelle Fragmentbibliothek bestehend aus den sieben in Tabelle 5.1 genannten kommerziellen Fragmentbibliotheken wurde mittels MOE wie folgt aufbereitet:

1. Entfernung aller Salzkomponenten und Gegenionen;
2. Entfernung von nicht gewünschten reaktiven Gruppen¹;
3. Generierung aller Protomere und Tautomere für den pH 7;
4. Entfernung von unerwünschter Protonierungs-, und Tautomerisierungszuständen²;
5. Energieminimierung mittels des MMFF94s Kraftfeldes [119] mit einem Gradienten von $0.0001 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$ als Konvergenzkriterium.

Für die Fragmentbibliothek der UiB Bergen wurden mit RDKit [2018.09.1] [234] alle Stereoisomere erzeugt. In MOE wurden die dominant vorherrschenden Protomer- bzw. Tautomerspezies bei pH 7 bestimmt und diese anschließend mittels des MMFF94s Kraftfeldes [240] bis zu einem Gradienten von $0.0001 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$ energieminiert.

5.1.2.1 Charakterisierung der Fragmentbibliotheken

Für eine Charakterisierung der Fragmentbibliotheken wurden unterschiedliche Moleküldeskriptoren berechnet. Die Auswahl dieser Deskriptoren erfolgte in Anlehnung an die in der Literatur berichteten Analysen chemischer Räume mittels PCA [241–243]. Tabelle 5.2 zeigt eine Übersicht aller berechneten Deskriptoren.

5.1.2.2 Hauptkomponentenanalysen

Alle PCAs in dieser Arbeit wurden mit Hilfe der *prcomp* Funktion in R durchgeführt. Folgende Deskriptoren flossen in die Auswertung mit ein: *HA*, *HDon*, *HAcc*, *TPSA*, *Nrot*, *NChir*, *NRings*, *clogP*, *Narom*, *Fsp3*. Für einen direkten Vergleich der UiB mit der virtuellen Fragmentbibliothek wurden die Kovarianzen

¹Der in MOE eingesetzte Filter beinhaltet unter anderem die folgenden reaktiven Gruppen und ist vollständig in [239] bzw. Anlage I zu finden: Acrylamide, azyklische Diketone, Aldehyde, Aliphatische Imine, Aliphatische Ketone, Aliphatische Nitroverbindungen, Aliphatische Thioester, Halogenalkane, Anhydride, Azide, Aziridine, Epoxide, Halogenpyrimidine, Iso(thio)cyanate, Maleimide, Michaelakzeptoren.

²Eingesetzt wurde eine in MOE verfügbare Standardliste an SMARTS-Definitionen um unerwünschte Protonierungs-, und Tautomerisierungszustände zu entfernen. Eine Liste der eingesetzten SMARTS-Definitionen finden sich in Anlage I.

Tabelle 5.2: Berechnete Deskriptoren der Fragmentbibliotheken

Deskriptor	Beschreibung	Software
MW	Molekulargewicht in Da	MOE
HA	Anzahl Nichtwasserstoffatome	MOE
HDon	Anzahl Wasserstoffbrückendonatoren	MOE
HAcc	Anzahl Wasserstoffbrückenakzeptoren	MOE
Narom	Anzahl aromatischer Atome	MOE
clogP	Octanol/Wasser Verteilungskoeffizient ^a	MOE
TPSA	Topologische polare Oberfläche in Å ² ^b	MOE
NRings	Anzahl an Ringsystemen	MOE
NRot	Anzahl rotierbarer Bindungen	MOE
NChir	Anzahl an chiralen Zentren	MOE
Fsp3	Anteil an sp ³ hybridisierten Kohlenstoffatomen	RDkit
Tanimoto Distance	Ähnlichkeitsmaß	RDkit

^a basierend auf einem unpublizierten linearen Modell (100 Parameter, 1827 Moleküle) von Labute P. (MOE)

^b Parametrisierung von Ertl *et al.* [244]

der Hauptkomponenten über die virtuelle Fragmentbibliothek bestimmt und für die Lage der UiB Fragmentbibliothek benutzt.

5.1.3 Strukturanalysen der SHP-Bindestelle

5.1.3.1 Grundlegende Einstellungen in MOE

Minimierungen von Proteinstrukturen wurden unter Verwendung des Amber14SB [117] Kraftfeldes durchgeführt. Organische Moleküle wurden unter Verwendung des MMFF94s [119] Kraftfeldes minimiert.

5.1.3.2 Strukturanalysen durch visuelle Inspektion in MOE

Die Strukturanalysen der SHP-Bindestelle wurde anhand der PDB-Einträge 5B6C [245] (SHP-Box von UFD1 in Komplex mit p97-N) und 5GLF [246] (SHP-Box von Derlin-1 in Komplex mit p97-N) durchgeführt. Die PDB Strukturen wurden von der PDB [14] bezogen, in MOE mittels der *Structure Preparation*-Funktion aufbereitet und mit der *Protonate3D*-Funktion, unter Verwendung der Standardeinstellungen, für den pH 7.4 protoniert. Die sinnvolle Protonierung von Histidin-Seitenketten innerhalb der SHP-Bindestelle wurde visuell kontrolliert. Alle Wasser- und Puffermoleküle, bis auf die an Wechselwirkungen innerhalb der SHP-Binderegion beteiligten, wurden aus den Strukturen entfernt. Die so aufbereiteten Komplexe wurden visuell auf Interaktionen zwischen den Kofaktoren und p97-N analysiert.

5.1.3.3 Interaktionpotential-Analysen in MOE

Für die Analyse wurden die aufbereiteten Strukturen von p97-N benutzt, bei denen zuvor die Kofaktoren entfernt wurden. Die folgenden chemischen Sonden wurden für die gitterbasierte Analyse möglicher Interaktionspunkte verwendet: Carboxyl-Sauerstoffatom, Kohlenstoff in einer Methylgruppe, aromatischer Kohlenstoff, Hydroxylgruppe (aliphatisch), protonierter sp^3 hybridisierter Stickstoff (neutrales bzw. protoniertes aliphatisches Amin), Chloratom und Stickstoff in Amidbindung (NH). Die Auswertung der Analyse erfolgte grafisch unter Verwendung von Isokarten für einen jeweils eingestellten Energiewert pro verwendeter Sonde.

5.1.3.4 Hot-spot-Analyse unter Verwendung des FTMap-Servers

Beide p97-N Strukturen (PDB: 5B6C und 5GLF) mit entfernten Kofaktoren wurden auf den FTMap-Server³ hochgeladen und dort mit Standardeinstellungen prozessiert. Die erhaltenen Ergebnisse wurden im PDB-Format heruntergeladen und in MOE visuell auf Interaktionen der Fragmente mit p97-N analysiert.

5.1.4 Pharmakophorbasiertes Screening

5.1.4.1 Erstellung der Pharmakophormodelle

Es wurden insgesamt vier unterschiedliche Pharmakophormodelle unter Zuhilfenahme des Pharmakophoreditors in MOE erstellt. Die Pharmakophormodelle basieren auf den Resultaten der vorher durchgeführten Strukturanalysen und sind im Ergebnisteil in Kapitel 6.1.3.1 genauer beschrieben.

5.1.4.2 Konformerbibliotheken und pharmakophorbasiertes Screening

Methodischer Hintergrund

MOE bietet die Möglichkeit der Konformerengenerierung mittels der *Conformation Import*-Funktion [247]. Die *Conformation Import*-Funktion zerlegt ein Molekül in kleinere Fragmente und unterzieht diese einer Konformationsanalyse, um sie anschließend wieder zusammensetzen. Durch die einfache Parallelisierbarkeit dieser Operation zeichnet sich die Methode durch einen hohen Durchsatz an Molekülen aus. MOE bietet drei Algorithmen zur Generierung von Konformeren an: eine systematische Suche, eine stochastische Suche und eine Generierung auf MD-Simulationsbasis. Bei der stochastischen Suche werden zufällige Rotamere generiert und diese anschließend energieminimiert. Die Methode eignet sich vor

³<http://ftmap.bu.edu/serverhelp.php>

allem für kleinere Moleküle mit wenigen rotierbaren Bindungen.

Durchführung

Die benötigte Konformerbibliothek der Fragmentbibliothek wurde in MOE unter Benutzung der *Conformation Import*-Funktion erstellt. Die Generierung der Konformere erfolgte mittels der stochastischen Suche, mit einem Konvergenzkriterium von $0.0001 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$, sowie einer maximalen Anzahl von 300 Konformationen pro Molekül (Die Vollständigen Einstellungen finden sich in Anlage I). Die erhaltene Konformerbibliothek wurde einem pharmakophor-basierten Screening unter Verwendung der vier erstellten Pharmakophormodelle unterzogen. Es mussten hierbei jeweils alle Bedingungen des jeweiligen Pharmakophormodells erfüllt sein, damit ein Fragment als Treffer gezählt wurde. Die Ergebnisse des Screenings wurde jeweils in Subfragmentbibliotheken abgespeichert und getrennt voneinander weiterbehandelt.

5.1.5 Virtuelles Screening

Eine schematische Übersicht über das durchgeführte virtuelle Screening ist in Abbildung 5.1 gezeigt. Es gliedert sich in einen Docking-Teil (weiß hinterlegt) und einen daran anschließenden MD-Simulation-Teil (grau hinterlegt). Im folgenden werden die einzelnen Abschnitte näher erläutert.

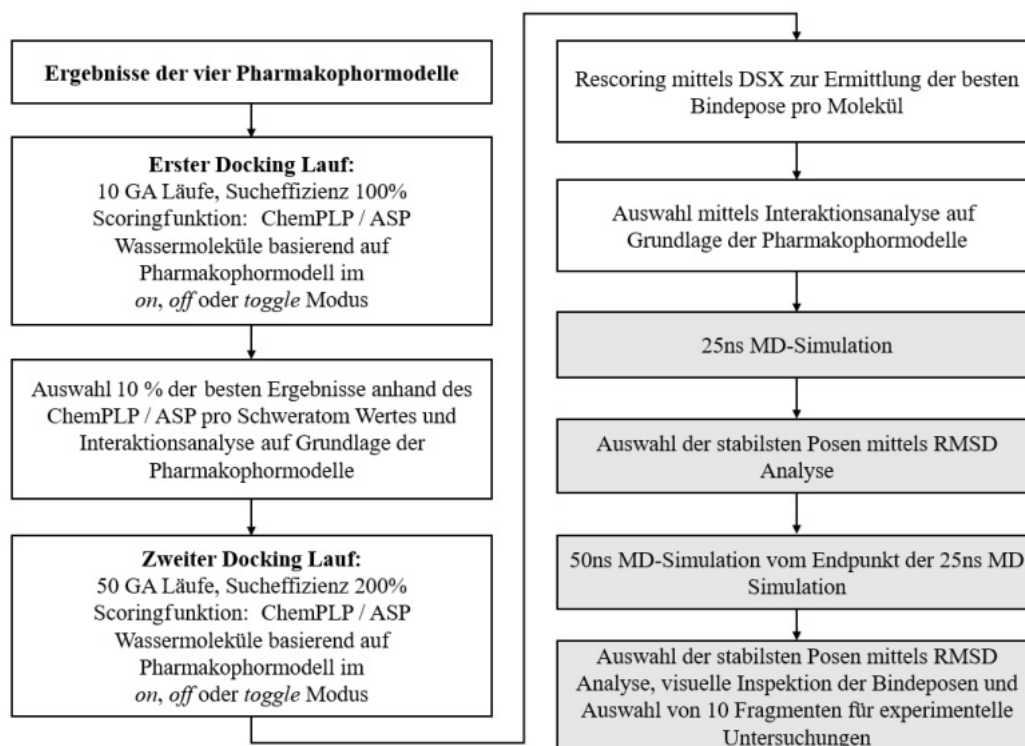


Abbildung 5.1: Schematischer Aufbau des virtuellen Screenings

5.1.5.1 Vorbereitung

Die vier erhaltenen Subfragmentbibliotheken des pharmakophorbasierten Screenings wurden in MOE bis zu einem Gradienten von $0.0001 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$ energie-minimiert und jeweils als multi .sdf Dateien abgespeichert. Als Proteinstruktur diente die präparierte p97-N Struktur mit entfernten Kofaktor UFD1 der Komplexstruktur mit dem PDB-Code 5B6C. Alle noch enthaltenen Wassermoleküle wurden bis auf vier, als strukturell wichtig erachtete (siehe Kapitel 6.1.2.4 für nähere Beschreibung), entfernt.

5.1.5.2 Erster Dockinglauf

Das Docking wurde mit dem Programm GOLD [5.2.2] [100] unter Verwendung der ChemPLP- und ASP-Scoringfunktionen durchgeführt. Jede erhaltene Subfragmentbibliothek des Pharmakophor-Screenings wurde separat mit den folgenden Einstellungen einem Docking unterzogen.

Der erste Dockinglauf beinhaltete die Generierung von jeweils zehn Bindeposen pro Fragment. Die Bindetasche wurde anhand eines Radius von 10 \AA um das Atom C^5 -I182 von p97-N gewählt. Die Einstellungen des GA, wie Populationsgröße, Selektionsdruck, Anzahl an Inseln, maximale Anzahl an Operationen und Nischengröße wurden auf *auto* gestellt. Es wurde eine Sucheeffizienz von 100% gewählt. Für den Liganden wurden die Optionen *Detektierung intramolekularer Wasserstoffbrückenbindungen*, sowie das *Durchschwingen pyramidalen Stickstoffatome* erlaubt. Die vorhandenen Wassermoleküle wurden anhand der vier Pharmakophormodelle (wie in Tabelle 5.3 zu sehen) behandelt. Je Pharmakophormodell wurden diese in den *on*-, *off*- oder *toggle*-Modus gesetzt. Für eine Definition der einzelnen Wassermoleküle **S1-S4** sei auf die Ergebnisse der Strukturanalyse in 6.1.2.4 verwiesen.

Tabelle 5.3: Behandlung der Wassermoleküle in GOLD basierend auf den vier Pharmakophormodellen

	Modell 1	Modell 2	Modell 3	Modell 4
Wassermolekül S1	on	on	on	on
Wassermolekül S2	on	on	on	on
Wassermolekül S3	toggle	toggle	off	off
Wassermolekül S4	on	off	on	off

5.1.5.3 Auswahl der Fragmente für den zweiten Dockinglauf

Methodischer Hintergrund

Zur Selektion von Fragmenten wurden zwischen dem *ersten* und *zweiten* Dockinglauf, sowie nach dem *zweiten* Dockinglauf, die erhaltenen Bindeposen auf ihre Interaktionen mittels der *Protein Ligand Interaction Fingerprint* (PLIF) Funktion in MOE analysiert [248]. Hierbei wird für jede Aminosäure eines Proteins, die in Kontakt mit dem Liganden steht, ein sogenannter „Fingerabdruck“ erstellt, in dem die gefundenen Interaktionen als *bit pattern* abgespeichert werden. Der schematische Aufbau eines solchen Fingerabdrucks ist in Abbildung 5.2 zu sehen. Die gefundenen Interaktionen werden in folgende drei Kategorien unterteilt: Wasserstoffbrückeninteraktionen der Seitenketten, Wasserstoffbrückeninteraktionen des Rückgrats und sonstige Interaktionen. Die Kategorien eins und zwei werden weiter in die Kategorie Wasserstoffbrückendonator bzw. Akzeptor und die dritte Kategorie in ionische und Oberflächenkontakte unterteilt. Die jeweilige Kategorie wird weiter in schwache bzw. starke Wechselwirkungen (basierend auf dem eingestellten Kraftfeld) unterteilt. Eine 0 ist als „keine Interaktion vorhanden“ und eine 1 als „Interaktion vorhanden“ definiert. Durch die so erzeugten *bit pattern* lassen sich Interaktionen zwischen Liganden und Proteinen schnell analysieren und filtern.

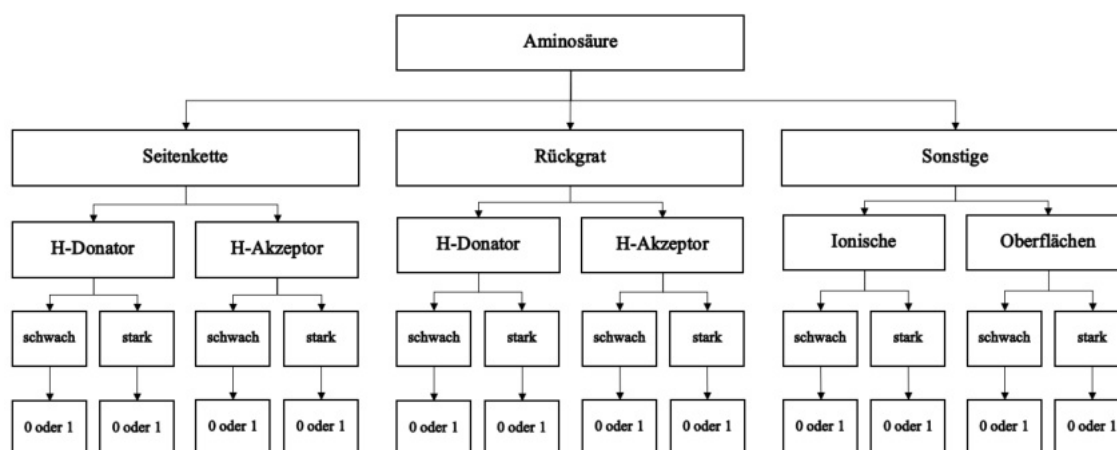


Abbildung 5.2: Schematischer Aufbau der *Protein Ligand Interaction Fingerprints* in MOE

Durchführung

Von den zehn Bindeposen eines Fragmentes wurde jeweils die bestbewertete Bindepose anhand ihres Score-Wertes bestimmt. Von dieser wurde der $Score_{HA}$

durch $Score_{HA} = Score_{Scoringfunktion} / N_{Anzahl\ Schweratome}$ berechnet. Die 10 % bestbewerteten Fragmente wurden anhand ihres $Score_{HA}$ für jede Subfragmentbibliothek des jeweiligen Pharmakophormodells ausgewählt. Die so erhaltenen Moleküle (pro Subfragmentbibliothek) wurden gemäß folgender Kriterien mittels des PLIF-Werkzeugs in MOE weiter selektiert (es wurden nur starke Wechselwirkungen betrachtet mit einem gesetzten Grenzwert von 1.5 kcal/mol):

- alle Fragmente, die keine Wasserstoffbrückenbindung mit p97-N zeigten, wurden entfernt;
- für Moleküle aus den Subfragmentbibliotheken der Pharmakophormodelle 1 und 2 sollte mindestens eine Wasserstoffbrückenbindung mit einer der folgenden Aminosäuren vorhanden sein: I182, T180, P178, A177 oder Wassermolekül (S4);
- für Moleküle aus den Subfragmentbibliotheken der Pharmakophormodelle 3 und 4 sollte mindestens eine Wasserstoffbrückenbindung mit einer der folgenden Aminosäuren vorhanden sein: T180, P178, A177, E141 oder Wassermolekül (S4).

Die so erhaltenen Fragmente wurden wieder energieminiert und gesondert als multil .sdf Dateien abgespeichert.

5.1.5.4 Zweiter Dockinglauf

Der zweite Dockinglauf wurde mit den gleichen Parametern wie der erste Dockinglauf durchgeführt, bis auf die zwei Änderungen, dass diesmal 50 Posen pro Molekül erstellt wurden und die Suche effizienz auf 200 % gesetzt wurde. Im Anschluss wurde wieder von den 50 Bindeposen eines Fragmentes die bestbewertete Bindepose anhand ihres Score-Wertes bestimmt und der $Score_{HA}$ berechnet.

5.1.5.5 Rescoring der Bindeposen des zweiten Dockinglaufs

Für die darauf folgende Selektion wurden die generierten Posen des zweiten Dockinglaufs einem Rescoring in DSX [0.89] [109] mittels der Scoringfunktion DSX, unter Verwendung der CSD-Potentiale, unterzogen. Die jeweiligen vorhandenen Wassermoleküle wurden in diesen Prozess mit einbezogen. Für die erhaltenen DSX-Score wurde analog der $Score_{HA}^{DSX}$ über $Score_{DSX} / N_{Anzahl\ Schweratome}$ berechnet. Für jedes Fragment wurde die Bindepose mit dem negativsten $Score_{HA}^{DSX}$ für die weitere Analyse ausgewählt.

5.1.5.6 Auswahl der Fragmente für MD-Simulationen

Die Auswahl der Fragmente für die MD-Simulationen wurde durch visuelle Analyse und unter Zuhilfenahme des PLIF-Werkzeugs in MOE durchgeführt. Die Herangehensweise war hierbei die gleiche wie in der ersten Selektion mittels des PLIF-Werkzeugs, mit dem Unterschied, dass die Selektionskriterien hier nun enger gewählt wurden. Da die genaue Auswahl auf den einzelnen Eigenschaften der Pharmakophormodelle und den Ergebnissen der vorausgegangenen Docking-Läufe basiert, sei für eine detaillierte Beschreibung der Kriterien auf Tabelle 6.4 im Ergebnisteil verwiesen.

5.1.6 MD-Simulationen

Alle MD-Simulationen des virtuellen Screenings, sowie die MD-Simulationen zur Analyse der Wassermoleküle in der SHP-Bindestelle wurden mit den Programmen der AMBER-Suite (Amber18 und AmberTools18) [2018] [216] durchgeführt. Die *mixed-solvent* Simulationen ausgewählter Fragmente aus dem biophysikalischen Fragmentscreening wurden in GROMACS [2019.01] [223] durchgeführt. Die AMBER Protokolle richten sich nach den beim AMBER-Workshop 2018 in Haifa (Israel) vorgestellten MD-Simulationsprotokollen und sind in der Anlage J zu finden⁴.

5.1.6.1 Präparation der verwendeten Liganden und Proteinstrukturen, sowie Erstellung der Startkonformation der Simulationen

Die Startkonformation jedes Liganden war die im Docking identifizierte beste $Score_{HA}^{DSX}$ -Bindenpose. Die elektrostatischen Potentiale der Verbindungen wurden mit dem Programm Gaussian09 [09 Revision C.01] [222] unter Verwendung der Hartree-Fock-Methode mit dem 6-31G* Basissatz berechnet. Die daraus ableitbaren Atomladungen wurden zusammen mit den Atomtypen des GAFF Kraftfeldes [118] mit Hilfe von *Antechamber/RESP* (AmberTools18) zugewiesen. Fehlende Kraftfeldparameter der Liganden wurden mittels des Programms *parmchk* (AmberTools18) abgeschätzt. Das für das Docking präparierte Protein (PDB: 5B6C) wurde mittels *pdb4Amber* (AmberTools18) in das AMBER konforme Format überführt. Die Protein-Ligand-Komplexe wurden mit Hilfe des Programms *tleap* (AmberTools18) unter Verwendung des AMBER ff14SB-Kraftfeldes [117] für das Protein erstellt und anschließend mit dem Programm *sander* (Amber18) für 500 Schritte einer Energieminimierung (100 Schritte mit der *steepest descent*-Methode und weitere 400 mit der *conjugate gradient*-Methode),

⁴Die Protokolle stammen aus der Arbeitsgruppe von Carlos Simmerling (Stony Brook University, New York, USA)

mit einem *generalized Born* (GB) impliziten Solvationsmodell [249], unterworfen. Der so erhaltene minimierte Komplex wurde im Anschluss in *tleap* mittels des TIP3P [120] Wassermodells mit einer oktaedrischen Wasserbox (10 Å) solvatisiert und mit Gegenionen neutralisiert.

Das Solvens wurde mit dem Programm *sander* für weitere 1,000 Schritte energie-minimiert (200 Schritte mit der *steepest descent*-Methode und weitere 800 mit der *conjugate gradient*-Methode), wobei der Komplex mit einem harmonischen Potential mit einer Kraftkonstanten von $500.0 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ fixiert wurde. Die Minimierung wurde über die Verfolgung des Verlaufs der potentiellen Energiewerte des Systems geprüft.

5.1.6.2 Generelle Einstellungen aller AMBER MD-Simulationen

Zur Temperaturkontrolle kam bei allen Simulationen die Langevin Dynamik mit einer Kollisionsfrequenz γ von 1 ps^{-1} zum Einsatz. Alle Simulationen im NPT-Ensemble wurden bei 300 Kelvin und 1.0 bar durchgeführt. Als Barostat wurde der Berendsen-Barostat, unter Verwendung der isotropischen Positionskalierung und einer Druckrelaxationszeit von 2 ps (Produktionsläufe virtuelles Screening) bzw. 5 ps (sonstige Simulationen) verwendet. Der SHAKE-Algorithmus wurde benutzt und die Integrationszeit somit auf 2 fs gesetzt. Der Van-der-Waals- und Coulomb-Wechselwirkungen wurden paarweise bis 8 Å summiert und für Abstände größer 8 Å über die PME-Methode berechnet.

5.1.6.3 Äquilibrierung der Systeme

Alle Systeme wurden mit dem folgenden Protokoll in *pmemd* (Amber18) äquilibriert:

- im NVT-Ensemble in 250,000 Schritten (500 ps) von 100 auf 300 Kelvin erhitzt, wobei der Protein-Ligand-Komplex durch ein harmonisches Potential mit einer Kraftkonstanten von $100.0 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ fixiert wurde;
- im NVT-Ensemble in 125,000 Schritten (250 ps) wieder von 300 auf 100 Kelvin abgekühlt, wobei die Kraftkonstante auf $10.0 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ verringert wurde;
- im NVT-Ensemble in 250,000 Schritten (500 ps) wieder von 100 auf 300 Kelvin erhitzt, wobei der Protein-Ligand-Komplex nur noch leicht mit einer Kraftkonstanten von $1.0 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ festgehalten wurde;
- im NPT-Ensemble in 250,000 Schritten (500 ps) zur Äquilibrierung des Drucks.

Die Kontrolle der einzelnen Schritte erfolgte über die Verfolgung des Verlaufs der Temperatur, der Energien (E_{tot} , E_{kin} , E_{pot}), des Drucks und des RMSD (*root mean square deviation*)-Werts für das Proteinrückgrat und den Liganden im Vergleich zur ersten Konfiguration der Äquilibration.

5.1.6.4 Auswertung mittels der *root mean square deviation*

Methodischer Hintergrund

Der eben genannte RMSD-Wert berechnet sich über:

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N r_i^2} \quad (5.1)$$

und gibt die mittlere quadratischen Abweichung der N -Atome eines Moleküls zu einem Referenzpunkt wieder.

Durchführung

Da alle Simulationen mit *periodic boundaries* durchgeführt wurden, wurde alle Trajektorien zunächst mittels der *autoimage* Funktion in *cpptraj* (AmberTools18) prozessiert. Alle in unterschiedlichen periodischen Boxen befindliche Moleküle werden hier auf die Ursprungs-Box zurück projiziert. Die erhaltene Trajektorie wurde in *cpptraj* auf das Peptidrückgrat (definiert über die Atome $C\alpha$, $C\beta$, N und O) aligniert und die erhaltene Trajektorie für die Bestimmung der RMSD-Werte der Liganden benutzt, wobei jeweils die Startkonfiguration des Simulation als Referenz diente.

5.1.6.5 MD-Simulationen des virtuellen Screenings

Alle ausgewählten Fragmente wurden nach der Äquilibration zunächst für 12,500,000 Schritte (25 ns) in *pmemd* im NPT-Ensemble simuliert, wobei die Konfigurationen alle 5,000 Schritte in die Trajektorie geschrieben wurden. Die RMSD-Werte der Liganden wurden anschließend in *cpptraj* für jedes Fragment berechnet. Die Stabilität der Bindeposen wurde über diese RMSD-Werte abgeschätzt und nur Liganden für weitere Simulationen ausgewählt, die eine stabile Pose zeigten (für detailliertere Beschreibung der Kriterien sei auf den Ergebnisteil 6.1.3.3 verwiesen).

Von ausgewählten Fragmenten wurden die letzte Konfiguration der 25 ns Simulation als Startkonfiguration für eine weitere Simulation herangezogen. Es wurden 25,000,000 Schritte (50 ns) in *pmemd* im NPT-Ensemble simuliert, wobei die Geschwindigkeiten von der letzten Konfiguration der 25 ns Simulation benutzt

wurden. Konfigurationen wurde erneut alle 5,000 Schritte in die Trajektorie geschrieben und anschließend die RMSD-Werte der Liganden berechnet.

5.1.6.6 MD-Simulation zur Abschätzung der Wassereigenschaften der SHP-Bindestelle

Für eine Abschätzung der thermodynamischen Eigenschaften des Wassers innerhalb der SHP-Bindestelle, wurde eine 100 ns MD-Simulation in AMBER durchgeführt und anschließend mit der GIST (*Grid Inhomogeneous Solvation Theory*)-Methodik [250] (implementiert in *cpptraj*) und dem Programm GISTPP [251] ausgewertet.

Methodischer Hintergrund

Die GIST-Methodik beruht auf der Inhomogenen-Solvatisierungs-Theorie (IST) und stellt eine Methode der statistischen Thermodynamik dar. Sie versucht die Frage zu beantworten, welche freie Enthalpie ΔG_{solv} bei der Solvatisierung eines Moleküls in Wasser eine Rolle spielt. Die freie Enthalpie ΔG_{solv} der Solvatisierung besteht gemäß der Gibbs-Helmholtz-Gleichung aus dem enthalpischen Term ΔH_{solv} und dem entropischen Term ΔS_{solv} :

$$\Delta G_{solv} = \Delta H_{solv} - T\Delta S_{solv} \quad (5.2)$$

Der Enthalpieterm besteht hierbei aus

$$\Delta H_{solv} = \Delta H_{sw} + \Delta H_{ww} \quad (5.3)$$

mit den Enthalpietermen ΔH_{sw} für die Molekül/Wasser Interaktionen und ΔH_{ww} für die Wasser/Wasser Interaktionen. Analog gilt für die Entropie

$$\Delta S_{solv} = \Delta S_{sw} + \Delta S_{ww} \quad (5.4)$$

mit dem Entropietermen ΔS_{sw} für den Molekül/Wasser Entropie und ΔS_{ww} für den Wasser/Wasser Entropie, wobei in der IST die Vereinfachung $\Delta S_{solv} \approx \Delta S_{sw}$ gemacht wird. Während die Enthalpie über das Kraftfeld einer Simulation zugänglich ist, wird der entropische Term ΔS_{sw} über eine Dichteverteilungsfunktion basierend auf der Boltzmann-Statistik abgeschätzt:

$$\Delta S_{sw} = -k_B \frac{\rho^0}{8\pi^2} \int g_{sw}(\mathbf{x}, \omega) d\mathbf{x}d\omega \quad (5.5)$$

Hierbei ist k_B die Boltzmannkonstante, ρ^0 die Referenzdichteverteilung im „Bulk“

Wasser und $g_{sw}(\mathbf{x}, \omega)$ die Molekül/Wasser Korrelationsfunktion, mit der Position \mathbf{x} der Teilchen und der Orientierung ω . ΔS_{sw} setzt sich also aus einem Translationsbeitrag ΔS_{sw}^{trans} und einem Orientierungsbeitrag ΔS_{sw}^{orient} zusammen. Die Lösung dieser Integrale erfolgt bei der GIST-Methode anhand eines gitterbasierten Ansatzes. Detaillierte Beschreibungen dieser Methode, die an dieser Stelle zu weit führen würden, sind in [250, 251] zu finden.

Durchführung der GIST-Analyse

Als Proteinstruktur wurde die für das Docking benutzte Struktur von p97-N (ohne Kofaktor, PDB: 5B6C) verwendet. Das Aufsetzen und die Äquilibration der MD-Simulation erfolgte analog der oben beschriebenen Methode, nur ohne Ligandstruktur. Für den Produktionslauf wurden 50,000,000 Schritte (100 ns) im NPT-Ensemble simuliert, wobei ein harmonisches Potential mit einer Kraftkonstanten von $10 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ auf alle Proteinschweratome wirkte. Dies war nötig, um im Anschluss die Trajektorie mittels GIST auswerten zu können. Rotationen oder Translationen des Proteins verfälschen das Ergebnis ⁵.

Die erhaltene Trajektorie wurde im Anschluss mit der in *cyptraj* implementierten GIST-Methode ausgewertet. Ein Gitter mit den Maßen $40 \times 40 \times 40$ und einem Gitterabstand von 0.5 \AA wurde um die SHP-Bindestelle definiert und jede zehnte Konfiguration (alle 100 ps) für die Analyse herangezogen. Es wurden hierbei Dichtekarten für die Sauerstoffatome und Wasserstoffatome der Wassermoleküle generiert. Des Weiteren wurden die folgenden thermodynamischen Größen berechnet:

- die Enthalpie ΔH_{solv} , berechnet aus der Summe der Molekül/Wasser Enthalpie ΔH_{sw} und der Wasser/Wasser Enthalpie ΔH_{ww} ;
- die Solvent Entropie ΔS_{sw} , berechnet aus der Summe der ΔS_{sw}^{orient} Entropie und der ΔS_{Trans} Entropie;
- und die freie Enthalpie ΔG_{solv} , berechnet über $\Delta G_{solv} = \Delta H_{solv} - T\Delta S_{sw}$ mit $T = 300 \text{ Kelvin}$;

Die Berechnungen wurden mit GISTPP durchgeführt. Anhand dieser Werte wurden diejenigen Volumenelemente ausgewählt, die eine 15-fach höhere Dichte an Sauerstoffatomen der Wassermoleküle gegenüber dem „Bulk“ Wasser (Dichte = $0.0334 \text{ Wassermoleküle pro \AA}^3$) und gleichzeitig eine freie Enthalpie

⁵Persönliche Korrespondenz mit Tom Kutzman (Lehman College, Department of Chemistry, New York, USA) auf dem Amber-Workshop März 2018 am Technion in Haifa, Israel

ΔG_{soln} kleiner -0.5 kcal/mol besaßen. Die Volumenelemente wurden anschließend in GISTPP geclustert und die Cluster in Pymol mit den Wassermolekülen der Kristallstrukturen von p97-N verglichen.

5.1.6.7 *mixed-solvent* Simulationen

mixed-solvent Simulationen wurden in dieser Arbeit benutzt, um mögliche Bindestellen der im biophysikalischen Screening identifizierten Fragmente zu identifizieren. Auch wenn die gefundenen Fragmente mit einer mittleren Anzahl von 15 Schweratomen größer als die in der Literatur eingesetzten Verbindungen waren [180, 185], wurde die Methode auf Grund des in [252] berichteten erfolgreichen Screenings mit Fragmenten bis zu 16 Schweratomen gewählt. Da AMBER nicht über die Möglichkeit verfügt geometrisch definierte Dummy-Atome zu simulieren, wurden alle Simulationen mit GROMACS unter Verwendung des OPLS-AA Kraftfeldes [121, 122] und des TIP3P Wassermodel durchgeführt.

Die Fragmente wurden mit dem LigParGen-Servers⁶ unter Verwendung des 1.14*CM1A Ladungsmodells parametrisiert [253–255]. Der Masseschwerpunkt eines zu simulierenden Fragmentes wurde mit Pymol berechnet, in VMD das Dummy-Atom erzeugt und die entsprechenden Einträge für dieses in den *.gro* und *.itp* Dateien hinzugefügt. Mit den GROMACS-*tools* wurde eine Lösungsmittelbox mit einer 1 M Konzentration des jeweiligen Fragments erstellt. Die Anzahl an Molekülen wurde mit Hilfe der Avogadro-Konstanten und dem Boxvolumen von 125 nm^3 berechnet. Die Solvatisierungsbox wurde mit Gegenionen neutralisiert, anschließend für 50,000 Schritte minimiert, für jeweils 100 ps im NVT und NPT Ensemble äquilibriert und für 1 ns im NPT Ensemble simuliert. Die Temperatur wurde unter Verwendung des Geschwindigkeits-Reskalierungs-Thermostaten und der Druck mithilfe des Parrinello-Rahman Barostaten [256] kontrolliert.

Mit der so erhaltenen homogenen Lösungsbox wurde die in MOE präparierte p97-N Struktur mit entfernten Kofaktor UFD1 (PDB: 5B6C) solvatisiert, mit einem Mindestabstand von 1 nm zwischen Proteinoberfläche und den Grenzen der kubischen Solvatisierungsbox. Das so erhaltene System wurde minimiert, äquilibriert und für 40 ns simuliert (Der Integrationszeitschritt lag bei 2 fs und es wurde alle 10 ps eine Konfiguration gespeichert). Für jedes Fragment wurden diese

⁶<http://zarbi.chem.yale.edu/ligpargen/>

Schritte viermal unabhängig voneinander wiederholt, und somit vier unabhängige Replikate mit jeweils 40 ns Simulationszeit erhalten⁷.

Die erhaltenen Trajektorien der vier Replikate wurden anschließend mit VMD [1.9.3] [237] auf das Proteinrückgrat ($C\alpha$ -Atome) aligniert und die prozentuale Besetzungsdichte der Dummy-Atome der Fragmente mithilfe der Funktion *volmap* (VMD) bestimmt. Die eingesetzte Gitterauflösung betrug hierbei 1.0 Å. Die so erhaltenen vier Dichtekarten wurden zu einer Gesamtdichtekarte aufsummiert und für eine erste Abschätzung möglicher Bindestellen herangezogen. Das gleiche Vorgehen wurde auch zur Bestimmung weiterer Dichten von anderen ausgewählten Schweratomen der Fragmente genutzt. Für die Bestimmung möglicher Bindestellen wurden visuell, die zwei höchst besetzten Gitterpunkte der Dummy-Atome in VMD ausgewählt. Die beteiligten Aminosäuren wurden durch einen Radius von 6.0 Å um die ausgewählten Gitterpunkte bestimmt, und als Binderegion definiert.

⁷Die Skripte zur Automation der Simulationen und Umsetzung in GROMACS wurden von Josef Kehrein geschrieben und mir dankenswerter Weise zu Verfügung gestellt.

5.2 Biochemische Methoden

5.2.1 p97-Konstrukte

Eine Übersicht über die verwendeten p97 Konstrukte findet sich in Tabelle 5.4. Alle aufgeführten Konstrukte wurden von Dr. Petra Hänzelmann⁸ mit teilweiser Unterstützung von Monika Kuhn⁹ hergestellt. Die Vervielfältigung der humanen cDNA-Plasmide wurde mittels der PCR-Methode [257] durchgeführt. Folgende Konstrukte wurden hergestellt:

- p97-N (aa 2-187) und p97-NL (21-199) im pETM14-Vektor (EMBL *plasmid collection*) mit einem N-terminalen, über eine 3C-Protease spaltbaren His6-Tag. Beide Konstrukte wurden mittels restriktionsfreier Klonierung (*Restriction free cloning*, RF) [258] hergestellt.
- p97-NL (aa 21-199, pETM14 mit N-terminalem 3C-Protease spaltbarem His6-Tag) und p97-ND1 (aa 2-480 kloniert KasI/HindIII im pProEx HTA-Vektor mit N-terminalem TEV-Protease spaltbarem His6-Tag), beide versehen mit einem C-terminalen AviTag (GLNDIFEAQKIEWHE). Der AviTag wurde mittels inverser Mutagenese, wie von *Fairhead & Howarth* [259] beschrieben, hergestellt;
- p97-ND1 (aa 1–480) (NdeI/Sall im pET21b-Vektor (Novagen) mit einem C-terminalen His6-Tag verbunden über einen 9 Aminosäuren (SVDKLAAL-LE) langen Linker [260]. Das gleiche Konstrukt wurde für die Alanin-Mutanten benutzt. Die ortsspezifische Mutagenesen erfolgten mittels des *Quick Change site-directed mutagenesis kit* von Agilent®;
- p97-ND1 (aa 1–460) mit einem C-terminalen His6-Tag verbunden über einen 2 Aminosäuren (RS) langen Linker gemäß dem von *Tang & Xia* [261] berichteten Konstrukt. Das Konstrukt wurde mittels RF Klonierung [258] hergestellt;
- FAF1-UBX (aa 568-650) (NcoI(PciI)/XhoI) im pETM11-Vektor (EMBL *plasmid collection*) mit N-terminalem TEV-Protease spaltbarem His6-Tag [26]

Die korrekten Nukleotidsequenzen aller Konstrukte wurde durch DNA-Sequenzierung bestätigt (Microsynth Seqlab, Göttingen).

⁸Wissenschaftliche Mitarbeiterin am Rudolf-Virchow-Zentrum - Center for Integrative and Translational Bioimaging der Universität Würzburg

⁹Technische Angestellte am Rudolf-Virchow-Zentrum - Center for Integrative and Translational Bioimaging der Universität Würzburg

Tabelle 5.4: Konstrukte von p97 und deren Verwendung

Konstrukt	Aminosäuren	Tag	Verwendung	Komentar	Resistenzen
p97-NL ^a	21-199	N-His-3C/C-Avi	BLI	enzymatisch mittels BirA biotinyliert	Ampicillin
p97-ND1 ^a	2-480	N-His-TEV/C-Avi	BLI	enzymatisch mittels BirA biotinyliert	Ampicillin
p97-ND1	1-480	C-His	Kristallisation, NMR		Ampicillin
p97-ND1-F131A ^a	1-480	C-His	NMR		Ampicillin
p97-ND1-I182A ^a	1-480	C-His	NMR		Ampicillin
p97-ND1-L198W ^a	1-460	C-His	Kristallisation, NMR		Ampicillin
p97-N	2-187	N-His-3C	Kristallisation, NMR		Kanamycin
p97-NL	21-199	N-His-3C	Kristallisation		Kanamycin
FAF1-UBX ^b	568-650	N-His-TEV	Fluoreszenzpolarisations- Assay		Kanamycin

^a Expression und Aufreinigung durchgeführt von Monika Kuhn

^b zur Verfügung gestellt von Petra Hänzelmann

5.2.2 Protein Charakterisierung

5.2.2.1 Ermittlung von Proteinkonzentrationen

Methodischer Hintergrund

Die Bestimmung der Proteinkonzentration beruht auf der Messung der Absorption von Licht bei einer Wellenlänge von 280 nm gemäß dem Lambert-Beer Gesetz:

$$c = \frac{A}{\epsilon d} \quad (5.6)$$

mit der Konzentration c , der Absorption A , der Schichtdicke d , sowie dem Extinktionskoeffizienten ϵ . Für die Absorption bei einer Wellenlänge von 280 nm lässt sich ϵ als Summe der einzelnen Extinktionskoeffizienten der Aminosäuren Tryptophan, Tyrosin und Cystein (vorliegend als Disulfidbrücken) wie folgt berechnen (n : Anzahl der entsprechenden Aminosäuren):

$$\epsilon(280 \text{ nm}) [M^{-1} \text{ cm}^{-1}] = 5500n_{\text{Trp}} + 1490n_{\text{Tyr}} + 125n_{\text{Disulfid}} \quad (5.7)$$

Durchführung

Alle Proteinkonzentrationen wurden an einem Nanodrop ND-1000 Spektrophotometer der Firma PEQLAB bei der Absorption von 280 nm vermessen.

5.2.2.2 SDS-PAGE Gele

SDS-PAGE (*Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis*)-Gele (15%) [262] wurden zur Kontrolle der Reinheit und der Proteinmasse während der Aufreinigung der Proteine benutzt. Alle Gele wurden bei 200 Volt zwischen 40 und 60 min laufen gelassen und mittels Coomassie G250 gefärbt.

5.2.3 Expression und Aufreinigung von rekombinanten p97-Spezies

5.2.3.1 Expression von p97-Spezies

Expression und Aufreinigung der verschiedenen Konstrukte von p97 wurden, wie bei Hänzelmann *et al.* [12] beschrieben, durchgeführt. Alle Proteine wurden durch T7-Promotoren [263] und Lac-Repressor [264] gesteuerte Expression in kompetenten *Escherichia coli* BL21(DE3) RIL-Zellen (Stratagene), wie folgt hergestellt: Das Plasmid wurde in kompetente *E. coli* BL21(DE3) RIL-Zellen transformiert. Die Bakterienzellen wurden hierzu für zwanzig Minuten auf Eis ruhen gelassen. Der Hitzeschock wurde bei 42 °C für eine Minute durchgeführt und danach die

Bakterien erneut für 2 Minuten auf Eis inkubiert. Anschließend wurden 300 μL LB-Medium hinzu gegeben und die Zellen bei 37°C und 200 rpm für 1 1/2 Stunden kultiviert. Der Ansatz wurde anschließend in 125 mL LB-Medium (versetzt mit den jeweiligen Antibiotika, siehe Tabelle 5.4) überführt und die Vorkultur über Nacht, unter gleichen Bedingungen wachsen gelassen.

30 mL eines solchen Ansatzes wurden in 2 L LB-Medium mit den jeweiligen Antibiotika überführt und bei 37°C bis zu einer optischen Dichte OD_{600} von 0.4 bis 0.6 wachsen gelassen. Anschließend wurde mit einer frisch hergestellten 0.5 mM IPTG-Lösung induziert und bei 16°C über Nacht exprimiert. Die Ansätze wurden durch Zentrifugation bei $4,000 \times g$ für 20 min geerntet und bis zur weiteren Aufreinigung bei -80°C gelagert.

5.2.3.2 Aufreinigung von p97 Spezies

Methodischer Hintergrund

Zur Aufreinigung der rekombinanten p97 Proteine kamen zwei Methoden zum Einsatz: die immobilisierte Metallchelate-Affinitätschromatographie (IMAC); sowie die Größenausschluss-Chromatographie (*Size Exclusion Chromatography*, SEC). Die IMAC beruht auf dem Prinzip der Komplexbildung von nicht protonierten Histidinen an divalenten Kationen, wie Ni^{2+} und der Säulenmatrix. Hierbei wird das rekombinante Protein mit einem Polyhistidin-Tag versehen und kann somit durch Immobilisierung von den restlichen, nicht modifizierten Proteinen getrennt werden. Das Protein kann danach durch Elution mit Imidazol, Protonierung der Histidine durch pH-Erniedrigung, oder Einsatz von Chelatoren wieder mobilisiert werden [265]. Eine SEC beruht auf dem unterschiedlichen Permeationsverhalten von Molekülen, die aufgrund ihrer Molekülgröße und Form aufgetrennt werden. Kleine Proteine können in die Poren der Säulenmatrix eindringen und wandern dadurch langsamer durch die Matrix im Vergleich zu großen Proteinen [266].

Durchführung

Für die IMAC kam das ProtinoTM Ni-TED-Harz der Firma Macherey-NagelTM, bei dem mittels Tris-(carboxymethyl)ethylendiamin (TED)-Gruppen chelatierte Ni^{2+} Kationen durch ein Kunstharz an Silikat immobilisiert sind, zum Einsatz [267]. Für die SEC wurden HiLoad®16/600 Superdex® *prep grade* Säulen der Firma GE Healthcare (jetzt Cytiva) mit unterschiedlichen Porengrößen verwendet. Die Trennmatrix besteht hier aus quer vernetzter Agarose mit kovalent gebundenem Dextran.

Die Zell-Pellets wurden langsam bei 4 °C aufgetaut und in Lyse-Puffer bei 4 °C schonend suspendiert (etwa 30 mL Zellen auf 100 mL Lyse-Puffer versetzt mit DNase- und Proteaseinhibitoren). Die Zellen wurden unter Eiskühlung durch 2-malige Passage durch einen Hochdruckhomogenisator mit einem Druck von 1.5 kbar aufgeschlossen. Nicht lösliche Lysat-Bestandteile wurden durch einstündiges Zentrifugieren bei 60,000 x g und 4 °C entfernt. Der lösliche Überstand wurde mit ca. 4 g Ni-TED-Harz (in Wasch-Puffer für eine Stunde äquilibriert) versetzt und für eine Stunde bei leichtem Rühren bei 4 °C inkubiert. Das inkubierte Ni-TED-Resin wurde in eine Gravitationssäule überführt, mit 300 mL Wasch-Puffer gewaschen und anschließend mit Fraktionen von 8 mL, 2 x 30 mL und 3 x 10 mL mit Elutions-Puffer eluiert, wobei die Fraktionen von 30 mL zusätzlich zehn Minuten inkubiert wurden. Die initiale Proteinausbeute wurde mittels Messung der Absorption bei 280 nm ermittelt, sowie zur Prüfung der molekularen Masse und der Reinheit wurde jede Fraktion außerdem mit einem SDS-Page Gel analysiert. Ausgewählte Fraktionen wurden auf ein Volumen von 5 mL zur weiteren Aufreinigung konzentriert.

Bei p97-N Konstrukten für die Kristallisation wurde zusätzlich der N-terminale 6xHis-Tag, mittels proteolytischer Spaltung durch rekombinante HRV-3C-Protase und Dialyse in Dialyse-Puffer bei 4 °C über Nacht, entfernt. Zur Entfernung der abgespaltenen HisTags, sowie nicht gespaltenem p97-N, wurde das Dialysat wieder einer IMAC mit Ni-TED unterzogen, wobei der Durchfluss ungetaggtetes p97-N enthielt.

Eine SEC wurde je nach Konstrukt für p97-N Proteine mit einer HiLoad® 75 16/600 Superdex® *prep grade*, bzw. für p97-ND1 Proteine mit einer HiLoad® 200 16/600 Superdex® *prep grade* durchgeführt. Die Chromatographie erfolgte an einem Äkta Chromatographiesystem der Firma GE Healthcare (jetzt Cytiva). Die Proben wurden vor Injektion durch einen Agarose Filter mit 0.2 µm Porengröße filtriert, um eventuelle Präzipitate zu entfernen. Die Proben wurde unter konstanter Flussrate mit SEC-Puffer eluiert. Bei Fraktionen die Protein enthielten wurde die Konzentration durch Messung der Absorption_{280nm} ermittelt, sowie die Reinheit mittels SDS-PAGE Gele überprüft. Ausgewählte Fraktionen wurden vereinigt, konzentriert, aliquotiert, in flüssigen Stickstoff Schockgefroren und bei -80 °C gelagert.

5.2.4 Herstellung und Biotinylierung von AviTag-p97

5.2.4.1 Biotinylierung von AviTag-p97-Konstrukten unter Verwendung von BirA

Die Biotinylierung wurde wie von *Fairhead & Howarth* [259] beschrieben durchgeführt. Ein Reaktionsansatz (1 mL) enthielt 100 μM p97-N/ND1-AviTag, 5 mM MgCl_2 , 2 mM ATP (eingestellt auf pH 7), 150 μM Biotin (50 mM Biotin Stocklösung in DMSO) und 1 μM GST-BirA (zur Verfügung gestellt von P. Hänzelmann). Alle Bestandteile wurden in PBS-Puffer gelöst und auf 1 mL ergänzt. Der Reaktionsansatz wurde für 1 Stunde bei 30 °C und 300 rpm inkubiert. Anschließend wurden dem Reaktionsansatz erneut 150 μM Biotin, 1 mM ATP, sowie 1 μM BirA hinzugegeben und erneut für 1 Stunde bei 30 °C und 300 rpm inkubiert. Dem Reaktionsansatz wurden anschließend zur Abtrennung von GST-BirA 100 μL Glutathion-Agarose (Pierce) hinzugegeben und für 30 min bei Raumtemperatur unter leichtem Schütteln inkubiert. Nach Überführung in eine Spin-Säule wurde für eine Minute bei 1,000 \times g zentrifugiert und 3-mal mit 200 μL PBS-Puffer gewaschen. Der Durchfluss mit enthaltenem biotinylierten p97 wurde abschließend mittels SEC, je nach Konstrukt mit einer HiLoad® 75 16/600 Superdex® *prep grade*, bzw. einer HiLoad® 200 16/600 Superdex® *prep grade* unter Verwendung von PBS-Puffer eluiert. Die Konzentrationen von Protein enthaltenen Fraktionen wurden mittels der Absorption_{280nm} ermittelt, gepoolt, konzentriert, aliquotiert, in flüssigen Stickstoff Schockgefroren und bei -80 °C gelagert. Der Erfolg der Biotinylierung wurde durch Beladen von Super-Streptavidin (SSA)-Biosensoren geprüft.

5.3 Herstellung der Stammlösungen ausgewählter Fragmente

Die bei den Herstellern bestellten Fragmente wurden allesamt in deuterierten DMSO- d_6 zu Konzentrationen von 100 mM oder 200 mM gelöst und bei -20 °C gelagert. Die Stocklösungen wurden in allen Experimenten verwendet.

5.4 Kristallisation von p97-ND1

Für eine spätere Aufklärung der Bindestelle von identifizierten Fragmenten, sowie zukünftigen Fragmentscreenings mittels Röntgenstrukturanalyse, wurde eine robuste Kristallisationsbedingung von p97 benötigt, die reproduzierbare Kristalle mit hoher Beugungsqualität liefern sollte. Zunächst wurde mit den beiden

Konstrukten der p97 N-Domäne (p97-N und p97-NL) ein Kristallisationsscreening durchgeführt. Dieses ergab zwar für das p97-NL-Konstrukt reproduzierbare Bedingungen mit großen Einzelkristallen, die jedoch alle eine starke Verzwilligung aufwiesen und somit nur bedingt für die später geplanten Experimente geeignet waren. Daher wurde entschieden sich auf Kristallisationsbedingungen für das p97-ND1-Konstrukt zu konzentrieren, nicht zuletzt wegen einer höheren biologischen Relevanz und größeren Nähe zum Vollängen-p97. Zum Zeitpunkt der Arbeit war die von *Tang & Xia* [261] berichtete Struktur mit 2.2 Å die am höchst-aufgelöste von p97-ND1 (ADP-Konformation). Daher wurde das dort berichtete Proteinkonstrukt mit der die IBMPFD-Krankheit verursachenden Mutation L198W für die Kristallisation ausgewählt. Die im Linkerbereich zwischen der D1-, und der N-Domäne vorhandene Mutation hatte hierbei keinen direkten Einfluss auf die favorisierte SHP-Bindestelle der N-Domäne.

5.4.1 Initiales-Kristallisationsscreening und Optimierung der gefundenen Bedingung

Methodischer Hintergrund

Ziel eines Kristallisationsscreenings ist es, diejenigen Bedingungen und Zusammensetzungen zu finden, bei denen das gelöste Protein in eine geordnete, dreidimensionale Kristallstruktur übergeht. Bei diesem Prozess wird das gelöste Protein (Einquphasen-System) durch eine gezielte Übersättigung in einen metastabilen Zustand überführt. Durch die noch vorhandene kinetische Energie des Systems wird ein Übertritt in einen zweiphasigen Gleichgewichtszustand verhindert. Wird diese Energiebarriere übersprungen, trennt sich die metastabile Lösung in ein zweiphasiges System. Dies führt entweder zur Präzipitation des Proteins oder im günstigsten Fall zur Bildung eines Kristalls. Um diese Übersättigung der Proteinlösung zu erreichen wurde in dieser Arbeit die Dampfdiffusionstechnik mittels der hängenden- (*hanging drop*) oder sitzenden-Tropfen (*sitting drop*) Methode angewendet. Hierbei wird das Protein mit der Präzipitatlösung (auch Mutterlauge genannt) gemischt und dampfdicht in einem Gefäß, welches im Überschuss Mutterlauge enthält, verschlossen. Der, das Protein enthaltene Tropfen, weist zu Beginn einen höheren Wasseranteil auf als die Mutterlauge. Wasser verdunstet nun solange aus dem Tropfen, bis mit der Mutterlauge ein Gleichgewicht eingestellt wird, wobei die Proteinkonzentration im Tropfen idealerweise bis zur Übersättigung ansteigt [268].

Durchführung

Da die von *Tang & Xia* berichtete Kristallisationsbedingung nur Sphärolite und Phasenseparation zeigte, wurde ein initiales-Kristallisationsscreening durchgeführt. Gescreent wurden die in Tabelle 5.5 gezeigten Kristallisationsscreens bei drei unterschiedlichen Proteinkonzentrationen (12, 14, 16 mg/ml), sowie bei 4 °C und 20 °C. Es wurde hierzu das gleiche Volumen an Mutterlauge mit den genannten Proteinkonzentrationen gemischt. Die Platten wurden mittels eines NT8-Kristallisationsroboters pipettiert (*sitting-drop*-Verfahren) und in einem ROCK-IMAGER-1000-Kristallisationsinkubator der Firma Formulatrix inkubiert. Das Kristallwachstum wurde täglich für 2 Wochen lang verfolgt. Gefundene Kristalle wurden mittels UV-Aufnahmen geprüft, um Salzkristallbildung auszuschließen.

Anschließend wurde ein Optimierungsscreening mit einer vielversprechenden Bedingung (Na-Formiat pH 6, Glycerol, PEG600) durchgeführt. Variiert wurden hierbei die Formiat-Konzentration, die Glycerol-Konzentration, sowie der pH-Wert in engen Grenzen (± 1.5 pH in 0.5 pH-Schritten).

Tabelle 5.5: Screens des initialen-Kristallisationsscreenings

Screens
Wizard I, II, III, IV
Protein Complex
pH Clear Suite I
PEG Suite I,II
OptiSalts
Optimix 4 PEG
JCSG+
Index-Screen HT
Ammonium Sulfat
Individueller Natriumformiat-Screen

5.4.2 Ko-Kristallisation und Soaking ausgewählter Fragmente

Alle (Ko-)Kristallisationen wurden im *hanging-drop*-Verfahren in 24 Napfplatten wie folgt durchgeführt:

- SEC-Puffer wurde auf eine ADP-Konzentration von 4 mM und 40 mM MgCl₂ mittels Stocklösungen eingestellt;
- das in SEC-Puffer^{ND1L198W} gelagerte Protein p97-ND1(L198W) wurde mit dem ADP/MgCl₂ enthaltenen SEC-Puffer auf eine Konzentration von 15 mg/mL eingestellt;
- es wurde DTT in einer finalen Konzentration von 1 mM hinzugegeben;
- die Proteinlösung wurde bei 25,000 x g und 4 °C für 30 min zentrifugiert;
- die Proteinlösung wurde mit der in 6.4.3.1 genannten Mutterlauge im Verhältnis 1:1 gemischt. Bei Ko-Kristallisationen enthielt die Mutterlauge das jeweilige Fragment in Konzentrationen zwischen 1 bis 5 mM (je nach Löslichkeit des Fragments) und somit in einem stöchiometrischen Überschuss von etwa 3:1 bis 16:1;
- das Kristallwachstum wurden für mindestens 4 Tage bei 20 °C durchgeführt.

Für Soaking-Experimente wurden apo-Kristalle oder Kristalle mit dem gleichen ko-kristallisierten Fragment benutzt. Die jeweiligen Fragmente wurden, je nach Löslichkeit bis zu 20 mM in Mutterlauge gelöst. Ein mögliches Ausfallen der Fragmente wurde toleriert, mit dem Hintergedanken eine gesättigte Fragmentlösung zu erhalten. Die Soaking-Lösungen enthielten, abhängig von der Fragmentkonzentration, DMSO-Konzentrationen zwischen 1 bis 10% (v/v), die von den Kristallen toleriert wurden. Die Kristalle wurden in die Soaking-Lösungen überführt und für unterschiedliche Zeiten (0.5, bis >24h) in den Soaking-Lösungen inkubiert.

Die Kryoprotektion der Kristalle erfolgte durch kurzes nacheinander folgendes Überführen der Kristalle in zwei Mutterlauge-Lösungen mit 20% bzw. 30% Glycerol (v/v). Die Kryoprotektionslösung enthielt hierbei das entsprechende Fragment in der im Soaking eingesetzten Konzentration. Die so erhaltenen Kristalle wurden mittels Nylon-Schleifen „gefischt“ und bei –80 °C in flüssigem Stickstoff bis zu ihrer Vermessung gelagert.

5.4.3 Datensammlung und Auswertung

Die Sammlung der hier vorgestellten Daten erfolgte an der European Synchrotron Radiation Facility (ESRF, Grenoble, Frankreich). Die Kristalle wurden an der Beamline ID30B vermessen und die Beugungen mit einem PILATUS3 6M Detektor der Firma Dectris aufgenommen. Alle Datensätze wurden bei 100 Kelvin, unter Rotation bei einem Winkel-Inkrement von 0.1 Grad (3600 Bilder pro Kristall) aufgenommen. Um das anomale Signal des Broms für die Lokalisation des Fragmentes TROLL2 nutzen zu können, wurde die Wellenlänge der Beamline auf 0.9184 Å nahe der Bromkante eingestellt.

Die Integration, Laue- und Raumgruppenbestimmung, sowie Skalierung der Daten wurde mit XDS [31.1.2020] [238] durchgeführt. Die initialen Phasen zur Strukturaufklärung wurden mittels molekularem Ersatz (*Molecular Replacement*, MR) unter Verwendung von PHASER [269] in der CCP4-Suite [7.1] [219], basierend auf der von *Tang & Xia* [261] berichteten Struktur bestimmt. Die Strukturverbesserung erfolgte anschließend in BUSTER [2.10.3] [218] und COOT [0.9.1] [220]. TLS-Gruppen zur Verfeinerung wurden mittels des TLSMD-Webservers [270] bestimmt.

Geometrische Beschränkungen (*restraint*)-Informationen für die Fragmente wurden mit Hilfe des Programms *Electronic Ligand Builder and Optimization Workbench* (eLBOW) in PHENIX [1.18.2] [230] unter der Verwendung einer kraftfeldbasierten Optimierung erstellt.

Für die Auswertung des anomalen Signals des Broms wurde die in XDS prozessierte Reflektionsdatei mittels SHELXC [2016/1] [235] weiterbearbeitet und die Strukturfaktoren (FA) und Phasen (α) bestimmt. Die Substrukturbestimmung wurde mit SHELXD [2016/1] [235] und die Analyse der anomalen Elektronendichtedichte des Broms mit ANODE [2013/1] [217] durchgeführt.

5.5 Biolayer-Interferometrie (BLI)-Messungen

5.5.1 Allgemeine Vorgehensweise

Alle BLI-Messungen wurden an einem Octet RED96e Instrument der Universität in Bergen (Norwegen), an einem Octet K2 Leihgerät der Firma Sartorius und an einem Octet RED96 (Messungen der Funktionalität der Biotinylierung) am Rudolf-Virchow-Zentrum der Universität Würzburg durchgeführt. Als Sensoren kamen High-Precision-Streptavidin (SAX)-Biosensoren, sowie Super-Streptavidin (SSA)-Biosensoren der Firma Sartorius zu Einsatz. Alle Messungen wurden bei 25 °C, einer Schüttelgeschwindigkeit von 1000 rpm in 96 Well-Micro-Plates der Firma Greiner Bio-One durchgeführt. Die Prozessierung der BLI-Messdaten erfolgte an der zum jeweiligen Messgerät zugehörigen Auswertungssoftware der Firma Sartorius. Die anschließende kinetische Auswertung der prozessierten Daten erfolgte durch hierfür geschriebene Python- und R-Skripte, Excel [225] und Origin Pro [2019b] [229]. Sämtliche Experimente und Prozessierungen wurden gemäß der Doppelreferenz-Methode durchgeführt.

Methodischer Hintergrund

Die Doppelreferenz-Methode ist eine Standardmethode um schwache Bindungssignale, wie sie von kleinen organischen Molekülen hervorgerufen werden, in SPR- und BLI-Messungen detektieren zu können [271].

Sie beinhaltet für die BLI-Methode das Messen der folgenden vier Signale Sig :

- die Messung der Probe mit immobilisierten Protein $Sig_{Protein/Probe}$
- die Messung des Assay-Puffers mit immobilisierten Protein $Sig_{Protein/Puffer}$
- die Messung der Probe mit einem Referenzsensor $Sig_{Ref./Probe}$
- die Messung des Assay-Puffers mit einem Referenzsensor $Sig_{Ref./Puffer}$

Das korrigierte Signal errechnet sich anschließend über:

$$Sig_{Doppel Ref.} = (Sig_{Protein/Probe} - Sig_{Protein/Puffer}) - (Sig_{Ref./Probe} - Sig_{Ref./Puffer}) \quad (5.8)$$

Für die Wahl eines Referenzsensors werden drei Verfahren berichtet [272, 273], bei dem der gleiche Sensortyp mit entweder Biocytin¹⁰, oder mit Streptavidin geblocktem Biocytin, oder einer inaktiven Mutante des Proteins beladen wird.

Da die gesamte N-Domäne im Fragmentscreening im Fokus stand, fiel die Benutzung einer inaktiven Mutante weg und es wurde ein mit Biocytin beladener

¹⁰Bei Biocytin handelt es sich um ein Carbonsäureamid aus Biotin und der Aminosäure Lysin.

Sensor als Referenzsensor benutzt. Die Variante eines Referenzsensors mit Streptavidin geblocktem Biocytin, die ein leicht besseres Signal-Rausch-Verhältnis gegenüber der Variante mit Biocytin besitzen soll [272], wurde aufgrund der kurzen Aufenthaltszeit in Bergen nicht in Betracht gezogen.

5.5.2 Beladen der Sensoren

5.5.2.1 Beladen der Sensoren mit p97-ND1

Das Beladen der Sensoren wurde nach dem Protokoll der Firma Sartorius wie folgt durchgeführt [274]:

- 30 min Äquilibrieren der SSA-Sensoren in PBS-Puffer auf der Laborbank und anschließendes Überführen in das Octet RED96e Gerät
- 10 min Äquilibrieren in Puffer^{BLI-1}
- 30 min Beladen der Sensoren mit biotinylierten p97-ND1 mit einer eingesetzten Proteinkonzentration von 50 µg/mL in Puffer^{BLI-1}
- 5 min Quenchen der Sensoren mit Biocytin¹¹ mit einer eingesetzten Konzentration von 10 µg/mL in Puffer^{BLI-1}
- 20 min Waschen der Sensoren in Puffer^{BLI-1}

Das Beladen der Referenzsensoren mit Biocytin wurde nach dem gleichen Protokoll durchgeführt, wobei der Schritt der Proteinbeladung weggelassen und die Sensoren direkt mit Biocytin gequencht wurden. Sensoren, die nicht eine Verschiebung von > 5 nm erreichten, im Waschschrift einen starken Abfall > 0.5 nm zeigten oder anderweitig von den zu erwartenden Kurven abwichen wurden nicht weiter benutzt und verworfen. Alle Sensoren wurden vor ihrem ersten Einsatz mit einer frisch hergestellten 500 µM ADP-Lösung auf ihre Funktionalität hin überprüft. Beladene Sensoren wurden maximal einen Tag eingesetzt und bei Nichtverwendung bei 4 °C für maximal 24 h in Puffer^{BLI-1} gelagert.

Anmerkung zur Beladung von p97-ND1 auf SSA-Sensoren

Bei den Messungen wurde festgestellt, dass die Beladung von großen Proteinen (Proteinmassen größer 70 kDa) bei SSA-Sensoren neueren Herstelungsdatums zu Problemen führte. Unterschiedliche Chargen zeigten hier ein unterschiedliches Beladungsverhalten bzw. ließen sich überhaupt nicht mehr mit Protein beladen. Aufgrund der Problematik wurde teilweise versucht mit SAX-Sensoren, die eigentlich nicht für die Detektion kleinerer Moleküle gedacht sind, durch eine ausreichende Beladung ein signifikantes Signal zu erhalten. Die Proteinkonzentrationen zur Beladung wurden hierbei auf bis zu 250 µg/mL erhöht. Es bestand

¹¹Hierdurch werden noch nicht besetzte Streptavidin Bindestellen abgesättigt

hier ein enger Austausch mit dem Hersteller Sartorius. Bis zum Ende der durchgeführten Messungen lag noch keine Problemlösung seitens Sartorius vor.

5.5.2.2 Beladen der Sensoren mit p97-N

Das Beladen der SSA-Sensoren für das Screening mit p97-N wurde, basierend auf den Ergebnissen des Screenings mit p97-ND1, sowie nach *Martin et al.* [273], hinsichtlich der eingesetzten Konzentration an biotinylierten Protein modifiziert¹². Das Beladen wurde wie folgt durchgeführt:

- 60 min Äquilibrieren der SSA-Sensoren in PBS-Puffer auf der Laborbank und anschließendes Überführen in das Octet RED96e Gerät
- 10 min Äquilibrieren in Puffer^{BLI-4}
- 20 min Beladen der Sensoren mit biotinyliertem p97-N mit einer eingesetzten Proteinkonzentration von 100 µg/mL in Puffer^{BLI-4}
- 5 min Quenchen der Sensoren mit Biocytin mit einer eingesetzten Konzentration von 10 µg/mL in Puffer^{BLI-4};
- 20 min Waschen der Sensoren in Puffer^{BLI-4}

Das Beladen der Referenzsensoren mit Biocytin erfolgte analog wie bereits beschrieben. Sensoren, die nicht eine Verschiebung von > 10 nm erreichten, im Waschschrift einen starken Abfall > 1.0 nm zeigten oder anderweitig von den zu erwartenden Kurven abwichen wurden nicht weiter benutzt und verworfen. Aufgrund einer fehlenden Positivkontrolle für die isolierte N-Domäne wurden mit Protein beladene Sensoren maximal für zwei Messungen (ca. 8 Stunden) verwendet. Referenzsensoren mit Biocytin wurden für einen Tag eingesetzt und bei Nichtverwendung bei 4 °C in Puffer^{BLI-4} für maximal 24 h gelagert.

5.5.3 Validierung des BLI-Assays

Die Bindung von ADP an die Nukleotidbindestelle der p97-ND1-Domäne wurde zur Validierung und Überprüfung des BLI-Assays verwendet. Hierzu wurden ADP-Lösungen in Puffer^{BLI-2} in folgenden Konzentrationen frisch hergestellt: 8; 21; 125; 500; 1,000; 6,250; 25,000; 100,000 nM. Folgende Assay-Einstellungen wurden gewählt¹³:

¹²*Martin et al.* berichten über die Evaluation von BLI-Assays für vier unterschiedliche Proteine bei dem Pharmaunternehmen Novartis. Alle Proteine wiesen Molekulargewichte zwischen 31 und 37 kDa auf und liegen somit in der Größenordnung der N-Domäne (23 kDa). Die dort berichteten, ansonsten gleichen Protokolle, wiesen eine deutlich höhere Beladung der Sensoren von über 10 nm auf. Die eingesetzten Konzentrationen an biotinylierten Protein lagen hier bei 100 µg/ml und mehr.

¹³Die einzelnen Zeiten wurden aufgrund der Erfahrungen der Arbeitsgruppe von Ruth Brenk (Universität Bergen, Norwegen), in der die Messungen durchgeführt wurden gewählt. Sowie auf Basis der analogen Vorgehensweise bei SPR-Messungen [275].

- Basislinienmessung und Assoziationszeit: 120 s
- Dissoziationszeit: 300 s

Basislinie und Dissoziation wurden jeweils im gleichen Napf gemessen. Das Signal einer 500 nM ADP-Lösung wurde zu Beginn des Assays gemessen und mit diesem die Messungen unterschiedlicher Sensoren gegeneinander normiert.

Zum Validieren des Assays für die isolierte N-Domäne stand zum Zeitpunkt der Messungen keine in der Literatur berichtete, geeignete Positivkontrolle zur Verfügung. Das von *Chimenti et al.* [61] berichtete Fragment (ID5), welches bei SPR-Messungen eine Bindung mit p97-N zeigte, lieferte im BLI-Assay keine verwertbaren Signale. Es wurde daher entschieden mit den p97-ND1 Assaybedingungen das Screening durchzuführen. Aufgrund der höheren Neigung der isolierten p97-N-Domäne Aggregate zu bilden, wurde die DTT-Konzentration auf 5 mM erhöht. Auf die Zugabe von $MgCl_2$ wurde verzichtet, da die N-Domäne keine Nukleotidbindestelle besitzt.

5.5.4 Fragmentscreening

Für das Fragmentscreening an p97 wurde der in Abbildung 5.3 gezeigte Ablauf gewählt.

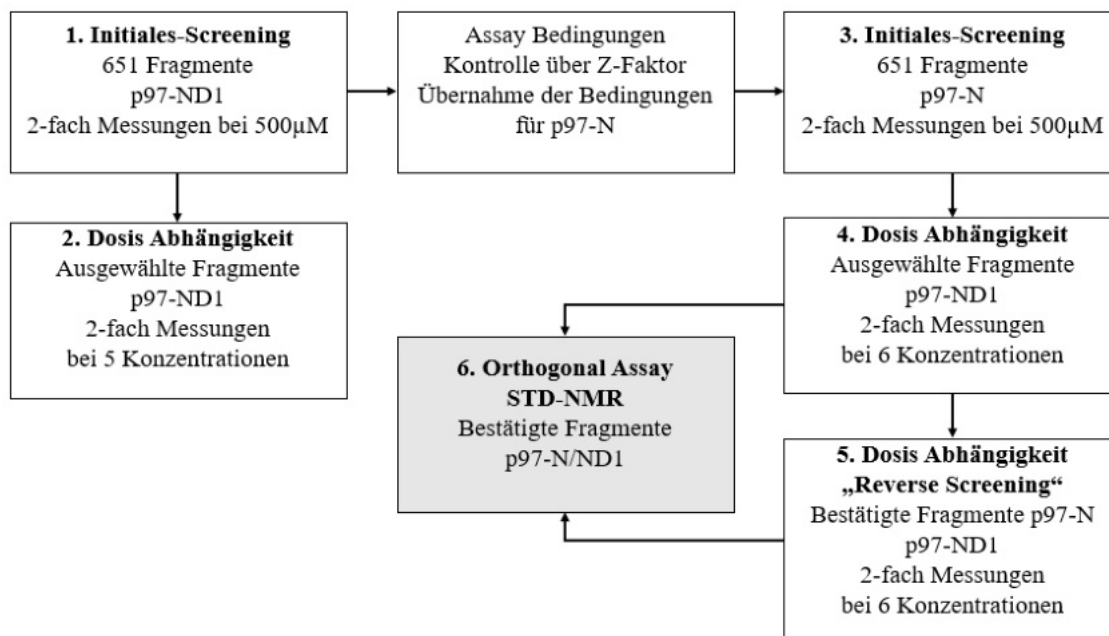


Abbildung 5.3: Schematischer Ablauf des Fragmentscreenings

Zunächst wurde ein initiales-Screening (1.) der Fragmente mit anschließender Messung der Dosisabhängigkeit ausgewählter Fragmente (2.) mit p97-ND1 durchgeführt. Die Qualität des Screenings wurde über den Z-Faktor ausgewertet und die gefundenen Assaybedingungen für p97-N übernommen. Es folgte ein weiteres initiales-Screening (3.) mit anschließender Messung der Dosisabhängigkeit (4.) mit p97-N. Bestätigte Fragmente aus (4.) wurden darauf mit p97-ND1 noch einmal auf ihre Dosisabhängigkeit geprüft (5.). Bestätigte Fragmente aus (4.) und (5.) wurden mittels eines STD-NMR orthogonal Assays (6.) noch einmal vermessen.

5.5.4.1 Initiale-Screenings

Mit p97-ND1 und p97-N wurden jeweils ein initiales-Screening der Fragmente der UiB-Bibliothek durchgeführt. Die Bibliothek umfasste zum Zeitpunkt des Screenings 678¹⁴ Fragmente der Firma Otava Chemicals. Die Bibliothek bestand aus 96 Napfplatten mit 200, 20 und 10 mM Stocklösungen in DMSO. Alle Platten wurden bei 4 °C unter Stickstoff gelagert.

¹⁴Aufgrund der beschränkten Messzeit bei den Aufenthalten an der UiB, wurden bei dem initialen-Screening mit p97-N nur 651 Fragmente vermessen. Eine Platte mit 27 Fragmenten wurde hier nicht vermessen.

Für das initiale-Screening wurden 10 μL der 10 mM DMSO Stocklösungen mittels eines Pipettierroboters vom Typ Mosquito HV *liquid handling robot* (TTP Labtech) (p97-ND1-Screening) bzw. mittels einer Multipipette (p97-N-Screening) überführt und mit 190 μL Puffer^{BLI-1} (p97-ND1) bzw. Puffer^{BLI-3} (p97-N) versetzt. Die finale Konzentration lag somit bei 500 μM , mit einer DMSO Anteil von 5 % (v/v).

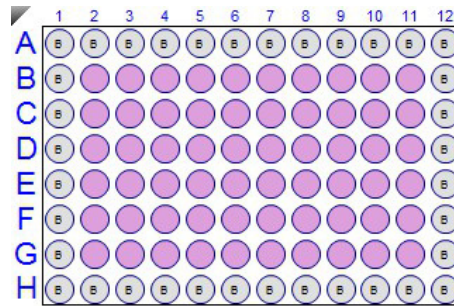


Abbildung 5.4: Aufbau einer Screeningplatte für ein initiales-Screening

Gezeigt ist die Belegung einer Screeningplatte eines initialen-Screenings. Mit B bezeichnete Näpfchen sind mit Assaypuffer, magentafarbige Näpfchen mit jeweils einem Fragment (500 μM) befüllt. Pro Platte konnten somit 60 Fragmente vermessen werden.

Der Aufbau einer Screeningplatte ist exemplarisch in Abbildung 5.4 gezeigt. Jede Platte enthielt 60 unterschiedliche Fragmente. Die Spalten 1 und 12 waren jeweils nur mit Puffer^{BLI-2} bzw. Puffer^{BLI-3} beschickt und dienten der Messung der Basislinien und der Dissoziation. Wie bei *Martin et al.* [273] berichtet stellte eine Kreuzkontamination und Verschleppung der unterschiedlichen Fragmente kein Problem dar. Ein Übertrag von einem Napf in den nächsten war dort mit 0.04% angegeben und somit vernachlässigbar. Die Reihen A und H wurden ebenfalls nur mit Puffer^{BLI-2} bzw. Puffer^{BLI-3} beschickt und dienten den Korrekturmessungen für den Puffer. Jede Platte wurde zuerst mit den p97 beladenen Sensoren und im Anschluss mit den Biocytin-Referenzsensoren vermessen. Die Assay Einstellungen waren wie folgt:

- Basislinienmessung: 15 s
- Assoziationszeit: 60 s
- Dissoziationszeit: 300 s

Die Platte wurde einmal vorwärts (Reihe 2 bis 11) und einmal rückwärts (Reihe 11 bis 2) gemessen, womit jedes Fragment zweimal vermessen wurde. Zur Berechnung der Platten Z-Faktoren des p97-ND1 initialen-Screenings wurden vor und nach jeder Screeningplatte eine Positivkontrolle (500 nM ADP-Lösung), sowie eine Negativkontrolle (Puffer) gemessen.

Methodischer Hintergrund

Zur Validierung des initialen-Screenings mit p97-ND1 wurde für jede Screeningplatte der Z-Faktor bestimmt. Hierbei handelt es sich um eine statistische Größe, welche vor allem in Hochdurchsatzverfahren zur Anwendung kommt und Aussagen über die Qualität eines Assays liefert. Er errechnet sich über

$$Z - \text{Faktor} = 1 - \frac{3(\sigma_p + \sigma_n)}{|\mu_p - \mu_n|} \quad (5.9)$$

wobei σ die Standardabweichung, μ der Mittelwert, n die Negativkontrolle und p die Positivkontrolle darstellen. Mit dem erhaltenen Wert lässt sich einschätzen, inwiefern ein Assay zwischen Positivkontrolle und Negativkontrolle unterscheiden kann. Die Interpretation der Z-Faktoren zeigt die folgende Tabelle 5.6.

Tabelle 5.6: Qualität eines Assays anhand des Z-Faktors nach Zhang et al. [276]

Z-Faktor	Assay Ergebnisse	Interpretation für Screening
$Z = 1$	$\sigma = 0$, Positiv und Negative Kontrolle liefern immer gleichen Wert	Theoretischer Idealfall
$1 > Z \geq 0.5$	Separation zwischen Negativ- und Positivkontrolle ist groß (bei 0.5 immer noch 12σ)	Exzellenter Assay
$0.5 > Z > 0$	Separation zwischen Negativ- und Positivkontrolle klein	Assay mit schwacher Aussagekraft
$Z = 0$	Keine Separation zwischen σ von Negativ- und Positivkontrolle	Assay nicht aussagekräftig
$Z < 0$	σ von Negativ- und Positivkontrolle überlappen	Nicht funktionierender Assay

5.5.4.2 Auswertung der initialen-Screenings

Die erhaltenen Ergebnisse wurden mit der Octet Analysesoftware [gerätespezifische Versionen] [221], wie folgt bearbeitet:

- alle Daten wurden mittels Doppelreferenzmethode prozessiert
- alle Sensorgramme wurden auf die Basislinien ausgerichtet
- der Beginn der Dissoziation wurde auf das Ende der Assoziation ausgerichtet (*Interstep-Correction*).

Das Signal eines Fragments wurde über den Mittelwert der Daten, der Assoziationsphase zwischen 55 und 59 s bestimmt. Das Signal der ADP-Positivkontrolle

wurde als Mittelwert der Daten, der Assoziationsphase zwischen 175 bis 179 s und die Negativkontrolle als Mittelwert der Daten, der Dissoziationsphase zwischen 270 und 300 s der ADP-Kontrolle bestimmt. Alle Daten wurden im .csv bzw. .xlsx Format für die weitere Bearbeitung exportiert. Die Beladungsstärke mit p97-N wurde, nach Ausrichtung der Basislinien mit dem Signal am Ende des Waschschriffs bestimmt.

Die prozessierten Daten wurde, wie bei *Giannetti et al.* [277] für SPR-Messungen beschrieben, ausgewertet. Das dort beschriebene Protokoll wurde für die hier durchgeführten BLI-Messungen adaptiert. Über die Positivproben war es bei dem initialen-Screening mit p97-ND1 möglich die Daten hinsichtlich Beladungsunterschiede der Sensoren, Beladungsverluste bzw. Funktionsverlust (z.B. durch Denaturieren des Proteins an der Sensoroberfläche) durch die gemessene ADP-Verschiebung (siehe 6.2.1.3) zu korrigieren. Anhand der Signale der ADP-Positivproben, die vor und nach jeder Screeningplatte vermessen wurden, wurde für jeden Sensor eine Regressionsgerade erstellt, die den Funktionsverlust von p97-ND1 über die Zeit darstellt. Neben der Korrektur für die Inaktivierung von p97 wurden darüber hinaus alle über 40 benutzten Sensoren hinsichtlich ihrer leicht unterschiedlichen Beiladungsstärken gegeneinander normiert. Die doppelreferenzierten Verschiebungen der Fragmente wurden daraufhin über die folgende Gleichung korrigiert:

$$R(t)_{norm} = R(t)/F_{reg}(t) \quad (5.10)$$

Wobei R die Verschiebung eines Fragments zum Zeitpunkt t der Messung, F_{reg} die errechnete Verschiebung der Regressionsgerade zum Zeitpunkt t und R_{norm} die korrigierte und normierte Verschiebungen zum Zeitpunkt t sind. Da die Experimente nach dem gleichen Schema durchgeführt wurden und somit die relativen Zeiten zwischen den Positivproben immer gleich waren, wurde zur einfacheren Handhabung nicht die wahre Zeit t , sondern ein fortlaufender Index t_i verwendet. Durch die Korrektur konnten die Signale unterschiedlicher Sensoren zu unterschiedlichen Zeitpunkten des Screenings miteinander verglichen werden. Die Korrektur wurde mit Python- und R-Skripten durchgeführt.

Durch das Fehlen einer geeigneten Positivprobe bei dem durchgeführten initialen-Screening mit p97-N, konnten hier weder Z-Faktoren noch die eben dargestellte

Korrekturen angewendet werden. Die Signale wurden hier nur für unterschiedliche Beladungszustände der Sensoren korrigiert:

$$R_{norm} = (R/R_{Beladung\ Sensor}) \cdot 10 \quad (5.11)$$

wobei R_{norm} die normierte Verschiebung eines Fragments, R die Verschiebung des Fragments und $R_{Beladung\ Sensor}$ die Verschiebung der Beladung des dazugehörigen Sensors sind. Um die erhaltenen Daten wieder in den gleichen Skalenbereich zu überführen, wurden sie mit dem Faktor 10 multipliziert.

5.5.4.3 Auswahl der Fragmente aus dem initialen-Screening für Konzentrationsabhängigkeitsmessungen

Die Sensorgramme von auffällig hohen oder negativen Signalen wurden visuell auf mögliche Artefakte inspiziert. Zur Auswahl der „Treffer“ in den initialen Screenings, wurden all diejenigen Fragmente herangezogen, die ein Signal R_{norm} größer der Summe des Mittelwertes μ_{Frag} und der Standardabweichung σ_{Frag} der Signale aufwiesen.

Die Sensorgramme der ausgewählten Fragmente wurden anhand der folgenden zwei Kriterien visuell selektiert: Alle Fragmente deren Sensorgramme nicht dem zu erwartenden, exponentiellen Kurvenverlauf folgten und die keine ersichtliche Plateauphase zeigten, wurden verworfen.

5.5.4.4 Konzentrationsabhängigkeitsmessungen ausgewählter Fragmente

Ausgewählte Fragmente wurden auf ihre Konzentrationsabhängig geprüft. Hierzu wurde für Fragmente des p97-ND1-Screenings Messungen bei fünf unterschiedlichen Konzentrationen und bei Fragmenten des p97-N-Screenings Messungen bei sechs unterschiedlichen Konzentrationen durchgeführt. Alle Konzentrationsreihen wurden als 1:1-Verdünnungsreihen ausgehend von 1,000 μM hergestellt. Verdünnungsreihen wurden mittels eines Pipettierroboters vom Typ Mosquito HV *liquid handling robot* (p97-ND1), bzw. mittels einer Multipipette (p97-N) hergestellt. Die Assay Einstellungen beliefen sich wie folgt:

- Basislinienmessung: 15 s (p97-N)¹⁵ bzw. 60 s (p97-ND1)
- Assoziationszeit: 120 s bzw. 90 s (Octet K2 Messungen)¹⁶
- Dissoziationszeit: 180 s bzw. 150 s (Octet K2 Messungen).

¹⁵Die kürzere Basislinienmessung bei p97-N wurde zur Reduktion der Messzeit gewählt, um 2 Messplatten pro Sensorbeladung vermessen zu können.

¹⁶Zur Reduktion der Messzeit wurden die Assay-Zeiten angepasst, da es sich bei dem Octet K2 nur um ein Zweikanal Gerät handelte und dieses nur eine begrenzte Zeit zur Verfügung stand.

Jede Konzentration wurde zwei mal gemessen ($n=2$). Die erhaltenen Daten wurden, wie bereits oben beschrieben mittels der Doppelreferenzmethode prozessiert. Auf eine anschließende Normierung der Daten konnte verzichtet werden, da keine unterschiedlichen Sensoren in die spätere Datenauswertung mit eingingen. Die Auswertung der erhaltenen Bindungskurven wurde anschließend in Origin Pro [229] durchgeführt. Die Fragmente wurden, anhand ihrer über das Langmuir-Modell ermittelten Affinitäten, in folgende Kategorien eingeteilt:

- **Binder** ($K_D < 1000 \mu M$)
- **schwache Binder** ($1000 \leq K_D < 2000 \mu M$)
- **unspezifische Binder** (sichtbare lineare Abhängigkeit bzw. Affinitäten $K_D > 2000 \mu M$)
- **keine Binder** (keine Konzentrationsabhängigkeit)

Ausgewählte Fragmente aus der Kategorie der Binder wurden anschließend auf Grundlage des linearen Zusammenhangs zwischen der empirischen Geschwindigkeitskonstanten k_{obs} und der Konzentration $[L]$ des Liganden untersucht.

5.5.4.5 Analyse der Plateauphasen

Das Erreichen eines chemischen Gleichgewichts, stellt sich bei optischen Sensoren durch das Ausbilden einer Plateauphase im Sensorgramm dar. Zur Auswertung der Affinität über das Langmuir-Modell ist das Erreichen dieses Plateaus essenziell. Um eine bessere qualitative Aussage treffen zu können, ob eine Plateauphase in einem Sensorgramm ausgebildet wurde, wurde die im Folgenden beschriebene Methode angewandt. Beim Erreichen des chemischen Gleichgewichts kommt es zu keiner Änderung des Signals mehr in der Zeit. Das Sensorgramm bildet also eine Gerade, deren Steigung idealerweise null betragen sollte. Um diese Steigung zu messen wurden von jedem Sensorgramm für jede Konzentration die Datenpunkte der letzten zehn Sekunden der Assoziationsphase genommen (Abbildung 5.5). Die Datenpunkte wurden einer linearen Regression (Abbildung 5.5, **Lineare Regression 1.**) unterzogen und somit die Steigungen $m_{Konz.}$ der Sensorgramme erhalten. Die so erhaltenen Steigungen $m_{Konz.}$ wurden gegen den Logarithmus der Konzentration aufgetragen. Ein Ligand der das chemische Gleichgewicht bei allen Konzentrationen erreicht hat (Abbildung 5.5, **(a)**), sollte hier einen Datensatz zeigen, der sich mit einer Geraden mit der Steigung $m_{Plateau}$ von null darstellen lässt (Abbildung 5.5, **Lineare Regression 2.**). Bei unspezifische Bindung kommt es nicht zum Erreichen der Plateauphase, vorallem bei höheren Konzentrationen (Abbildung 5.5, **(b)**). Legt man durch die so erhaltenen Datenpunkte eine Gerade, so wäre die Steigung $m_{Plateau}$ größer null. Erreichen die kleinen Konzentrationen auf Grund einer zu kurzen Assoziationsphase nicht das

Plateau, wäre der Fall umgekehrt und die Gerade hätte eine Steigung m_{plateau} von kleiner null. Anhand dieser Steigungen m_{plateau} eines jeden Fragments wurde das Erreichen des Plateaus qualitativ abgeschätzt. Als Positivkontrolle der Methode wurde hier der Ligand ADP benutzt.

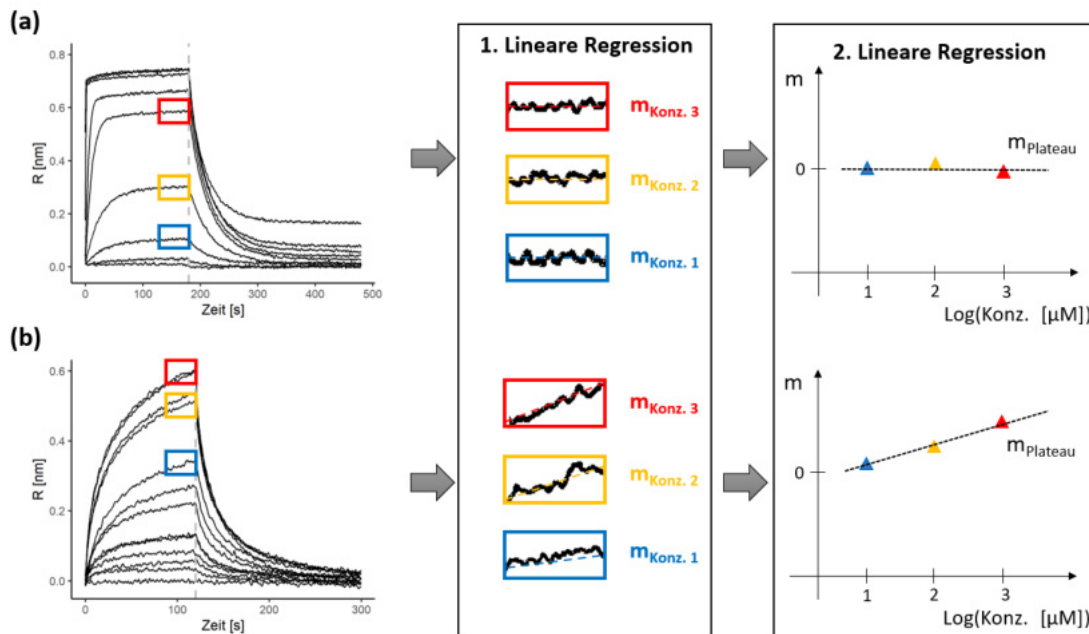


Abbildung 5.5: Methode zum Abschätzen des Erreichens der Plateauphasen in Sensorgrammen

Ein Ausbilden eines chemischen Gleichgewichts wird in den Sensorgrammen durch eine Plateauphase sichtbar (a). Sehen Sensorgramme wie bei (b) aus liegt kein chemisches Gleichgewicht vor und die über ein Langmuir-Modell abgeschätzte Affinität ist extrem fehlerbehaftet. Zur Ermittlung, ob ein chemisches Gleichgewicht erreicht wurde, wurden die Steigungen der Sensorgramme aller Konzentrationen in den letzten 10 Sekunden ermittelt und gegen diese aufgetragen. Die so erhaltenen Steigungen m_{plateau} dienen der qualitativen Analyse der Sensorgramme.

5.5.5 Messungen der Affinitäten der Bindungen von Xanthohumol (XN) an die N- und ND1-Domäne von p97

Die Messungen erfolgten mit p97-N an einem Octet RED96e der Universität Bergen (Norwegen) und mit p97-ND1 an einem Octet K2 (Leihgerät der Firma Sartorius) an der Universität Würzburg.

Da, wie bei *Sasazawa et al.* [69] berichtet, XN einen *in vitro* Effekt bei Konzentrationen zwischen 3 bis 10 μM zeigte, wurde eine 1:1-Verdünnungsreihe von 20 bis 0.625 μM für die Messungen gewählt. Die XN-Stammlösung wurde mit DMSO verdünnt und mit Puffer^{BLI-3} auf die entsprechenden Konzentrationen eingestellt. Der finale DMSO-Anteil im Assay betrug 5% (v/v).

Für p97-N wurden drei Konzentrationsreihen mit drei SSA-Sensoren vermessen. Für die ND1-Domäne wurden vier Konzentrationsreihen mit vier SAX-Sensoren bzw. zwei Konzentrationsreihen mit zwei SSA-Sensoren vermessen. Jede Konzentrationsreihe wurde hierbei zweimal mit dem gleichen Sensor vermessen. Die Assay-Einstellungen waren:

- Basislinienmessung 15 s
- Assoziationszeit: 120 s
- Dissoziationszeit: 180 s

Alle Daten wurden mittels der Doppelreferenzmethode prozessiert und mittels des Beladungssignals der Sensoren normiert. Die weitere Auswertung erfolgte in Origin Pro.

Um Einflüsse der Nukleotid-abhängigen Konformation auf die Bindung von XN zu klären, wurden Messungen mit p97-ND1 unter Verwendung von SSA-Sensoren durchgeführt. Die Verschiebung einer 5 μM XN-Lösung wurde hierbei in Abwesenheit eines Nukleotids, sowie in Gegenwart von 1 mM ADP bzw. 1 mM des ATP-Analogen AMP-PNP vermessen. Für jeden Nukleotid-Zustand wurden vier Messungen durchgeführt. Die Auswertung erfolgte in Origin Pro.

5.6 STD-NMR-Messungen

5.6.1 Durchführung der STD-NMR-Messungen

5.6.1.1 Einstellungen des Spektrometers

Alle STD-NMR Messungen wurden an einem Bruker (Karlsruhe, Deutschland) Avance 400 MHz Spektrometer am Institut für Pharmazie und LMC der Universität Würzburg durchgeführt. Das Spektrometer arbeitete bei 400.13 MHz mit einem BBO BB-H 5 mm Probenkopf bei 300 Kelvin. Die Temperatur wurde durch eine BCU-05 Temperaturkontrolleinheit konstant gehalten. Für die Datenerfassung wurde das Pulsprogramm *stdiffesgp.3* aus der Bruker Pulsprogramm-Bibliothek verwendet. Alle STD-NMR Messungen wurden mit den folgenden Einstellungen durchgeführt [278]: Die Sättigungsfrequenzen des *on*- bzw. des *off*-Spektrums wurden auf -400 Hz bzw. -16.000 Hz gesetzt, mit einer Sättigungszeit von 3 s (Relaxationsverzögerung D1: 4 s); das HDO-Signal wurde bei 4.703 ppm mittels der *excitation-sculping*-Methode unterdrückt¹⁷.

5.6.1.2 Umpufferung des Proteins

Die Proteine wurden mittels 0.5 mL Centricons (Merck Millipore) mit einem Molekulargewichtsgrenzwert (*Molecular weight cut-off*, MWCO) von 10kDa für p97-N bzw. MWCO von 30kDa für p97-ND1, durch mehrmaliges Waschen und Zentrifugieren bei 10.000 × g und 4 °C in den NMR-Puffer überführt. Die finale Proteinkonzentration wurde durch Messung der Absorption bei $\lambda = 280$ nm mit $n = 3$ Messungen bestimmt.

5.6.1.3 Messungen ausgewählter Fragmente des virtuellen Screenings

Für die STD-NMR Messungen ausgewählter Fragmente des virtuellen Screenings wurde die überführte N-Domäne auf eine Konzentration von 40 μ M mit NMR-Puffer eingestellt. Die jeweilige Fragment-Konzentration belief sich auf 500 μ M, mit einem finalen DMSO- d_6 Anteil von 5% (v/v). Die Bedingungen wurden auch für weitere Messungen an der N-Domäne beibehalten. Negativproben, d.h. in Abwesenheit eines Fragments, wurden unter ansonsten identischen Bedingungen durchgeführt.

¹⁷Die Einstellung, Programmierung und Übertragung der STD-Messmethode auf das oben genannte Spektrometer wurde von AD Dr. Curd Schollmayer und Markus Zehe durchgeführt.

5.6.1.4 Messungen an der ND1-Domäne

Für STD-NMR Messungen an der ND1-Domäne (einschließlich der Alanin-Punkt-Mutanten) wurde das überführte Protein auf eine Konzentration von 30 μM mit NMR-Puffer eingestellt. Die jeweilige Fragment Konzentration belief sich auf 428 μM ¹⁸, mit einem finalen DMSO- d_6 Anteil von 5% (v/v). Alle Messungen wurden in Gegenwart von 500 μM ADP und 5 mM MgCl_2 durchgeführt. Die 100 mM ADP-Stammlösung und 1 M MgCl_2 (wasserfrei)-Stammlösung wurden in D_2O hergestellt.

5.6.2 Auswertung der STD-NMR-Messungen

Die Auswertung der Messungen erfolgte mit der Software TopSpin [4.0.8] [236]. Die pseudo-2D-Spektren wurden hierzu mittels der in TopSpin implementierten Funktion *proc_std* in ein 1D Referenz-Spektrum (*off*-Spektrum) und ein 1D Differenz-Spektrum (*off*-Spektrum „minus“ *on*-Spektrum) prozessiert. Zur Auswertung der Spektren wurden nur Signale, die ein Signal-zu-Rauschverhältnis (S/N) > 3 % im Differenzspektrum zeigten, benutzt. Zur Auswertung wurden nur Signale im aromatischen Bereich (> 5.0 ppm) der Fragmente berücksichtigt. Dieser Bereich wurde wegen folgender Gesichtspunkte gewählt: Aliphatische Systeme relaxieren schneller als aromatische Systeme, weswegen ein möglicher STD-Effekt im aromatischen Bereich am größten ist; Signale bei 4.0 ppm liegen wurden durch die Wasserunterdrückung entweder ausgelöscht oder stark beeinflusst und waren somit für eine Auswertung nicht verwendbar; da das Einstrahlprofil bei -1 ppm dem einer Gaußkurve entspricht, könnten nahe liegende aliphatische Systeme eines Liganden angeregt werden und ein falsch positives Signal erzeugen.

Für jedes Fragment wurde der prozentuale STD-Effekt eines Protons durch die Intensität des Referenz-Spektrums I_{off} und die Intensität des Differenz-Spektrums I_{diff} , über $std = I_{diff}/I_{off} \cdot 100$ berechnet. Die Auswertung erfolgte auf Basis des Mittelwerts std_{mean}^{arom} aller STD-Effekte und des größten gefundenen STD-Effekts std_{max}^{arom} .

¹⁸Der Wert entstand durch die zur Verfügung stehende Feinpipette und dem eingesetzten Volumen von 1.5 μL der 200 μM Stocklösung.

5.7 Fluoreszenzanisotropie Wettbewerbsassay

Theoretischer Hintergrund

Regt man ein Fluorophor mit linear polarisiertem Licht an, emittiert dieses ein Photon, welches ebenfalls eine Polarisation aufweist. Das Fluorophor besitzt hierbei die Fluoreszenzlebensdauer τ , die die mittlere Zeit angibt, wie lange ein Molekül im angeregten Zustand verweilt, bis es durch die Emission eines Photons wieder in den Grundzustand zurückfällt. Diese Zeit τ liegt im Größenbereich der Rotationen von Molekülen in Lösungen. Das bedeutet vereinfacht gesagt, wird ein Molekül angeregt, rotiert es, und strahlt daher ein Photon in einer anderen Polarisationsrichtung aus, als die in der es angeregt wurde. Die Rotation des Moleküls in Lösung ist hierbei von dessen Volumen, der Temperatur und der Viskosität der Lösung abhängig. Betrachtet man einen ein Fluorophor enthaltenen Liganden der an ein weiteres Molekül bindet, ändert sich das Volumen des Liganden zu dem des entstehenden Komplexes. Somit ändert sich seine Rotationsgeschwindigkeit, die zur Änderung in der beobachteten Polarisation der Lösung führt. Gemessen wird hier die Differenz der bei zwei unterschiedlichen Polarisationsrichtungen gemessenen Intensitäten des emittierten Lichts [279].

Der messtechnische Aufbau eines Fluoreszenzanisotropie Experiments ist wie folgt (Abbildung 5.6) [279]: Über einen **Polarisator** wird polarisiertes Licht erzeugt, mit dem das in der Probe enthaltene **Fluorophor** angeregt wird. Die Intensität des emittierten Lichts des Fluorophors wird einmal bei **paralleler Anordnung** beider Polarisationsfilter (Intensität I_{para}), sowie bei **senkrechter Anordnung** beider Polarisationsfilter gemessen (Intensität I_{senk}). Die Fluoreszenzanisotropie FA lässt sich hierdurch über die folgende Gleichung berechnen:

$$FA = \frac{I_{para} - I_{senk}}{I_{para} + 2I_{senk}} \quad (5.12)$$

Das gemessene Signal wird hierbei um so größer, je größer die Volumenänderung bei der Bindung ist. Somit zeigt ein kleines Molekül einen viel größeren Effekt bei Bindung an ein großes Molekül, als umgekehrt.

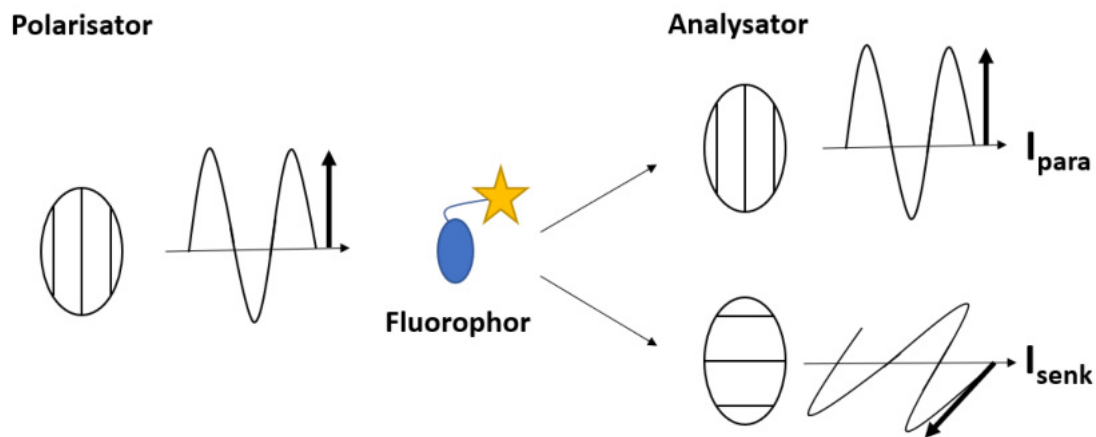


Abbildung 5.6: Aufbau eines Fluoreszenzanisotropie-Experiments

Die Fluoreszenz-Intensität eines mit polarisiertem Licht angeregten Fluorophors wird einmal über eine parallele und senkrechte Anordnung des zweiten Polarisationsfilters (Analysator) vermessen. Bei einem Bindungsereignis ändert sich das Volumen des Fluorophors und somit die Differenz der beiden gemessenen Intensitäten.

Durchführung

Es wurden vier Konzentrationsreihen mit acht Konzentrationen von p97-ND1 in Form einer 1:1-Verdünnungsreihe ausgehend von 20 μM in PBS-Puffer¹⁹ hergestellt. Die einen His6-Tag enthaltene FAF1-UBX-Domäne wurde mit dem Monolith His-Tag Labeling Kit RED-tris-NTA, 2nd Generation (Nanotemper) nach dem im Kit enthaltenen Standardprotokoll Fluoreszenz markiert. Konzentrationsreihe eins enthielt ausschließlich 5 μM der markierten UBX-Domäne von FAF1. Konzentrationsreihe zwei, drei und vier enthielten zusätzlich 12,5, 25 bzw. 50 μM XN. Die Fluoreszenzanisotropie Messung wurde an einem CLARIOstar UV/vis-Spektrometer der Firma BMG Labtech durchgeführt. Jede Konzentration wurde zehnmal gemessen. Die anschließende Datenauswertung erfolgte in Origin Pro.

¹⁹Der standardmäßig eingesetzte PBS-Puffer wurde in diesem Assay mit 0.1 % Tween 20 (v/v) zu Löslichkeitserhöhung von XN versetzt.

Kapitel 6

Ergebnisse & Diskussion

In diesem Hauptteil der Arbeit werden die Ergebnisse der durchgeführten Experimente gezeigt und diskutiert. Zunächst wird auf die Resultate des fragmentbasierten virtuellen Screenings und anschließend auf die Ergebnisse der biophysikalischen Methoden eingegangen. Der letzte Abschnitt befasst sich mit der Lokalisierung der Bindestellen mit Hilfe von computerbasierten Methoden, NMR-Experimenten und kristallographischen Analysen. Am Ende werden diejenigen Fragmente näher betrachtet, für die eine Optimierung als sinnvoll erachtet wird.

6.1 Virtuelles Screening an der SHP-Bindestelle von p97

6.1.1 Charakterisierung der Fragmentbibliotheken anhand physikochemischer Eigenschaften

Die für das virtuelle Screening zusammengestellte, wie auch im biophysikalischen Screening benutzte Fragmentbibliothek wurde anhand der folgenden physikochemischen Deskriptoren charakterisiert: MW, *clogP*, Anzahl Schweratome, Anzahl rotierbarer Bindungen, Anzahl Wasserstoffbrückendonatoren/akzeptoren, TPSA, Anzahl chiraler Zentren, Anzahl an Ringsystemen, Anteil an sp^3 hybridisierten Kohlenstoffatomen. Zunächst wurden die berechneten Deskriptoren hinsichtlich ihrer Verteilungen analysiert. Die UiB Fragmentbibliothek erfüllt hier weitestgehend die Kriterien der *Rule of 3*. Die virtuelle Fragmentbibliothek besteht ebenfalls hauptsächlich aus Fragmenten, die der *Rule of 3* folgen. Die Abweichungen von dieser sind den in der Bibliothek vorhandenen zusätzlichen PPI-Bibliotheken (siehe 5.1.1.1) geschuldet, die im Schnitt größere, lipophilere Moleküle enthalten (alle Verteilungen sind in Anhang A zu finden).

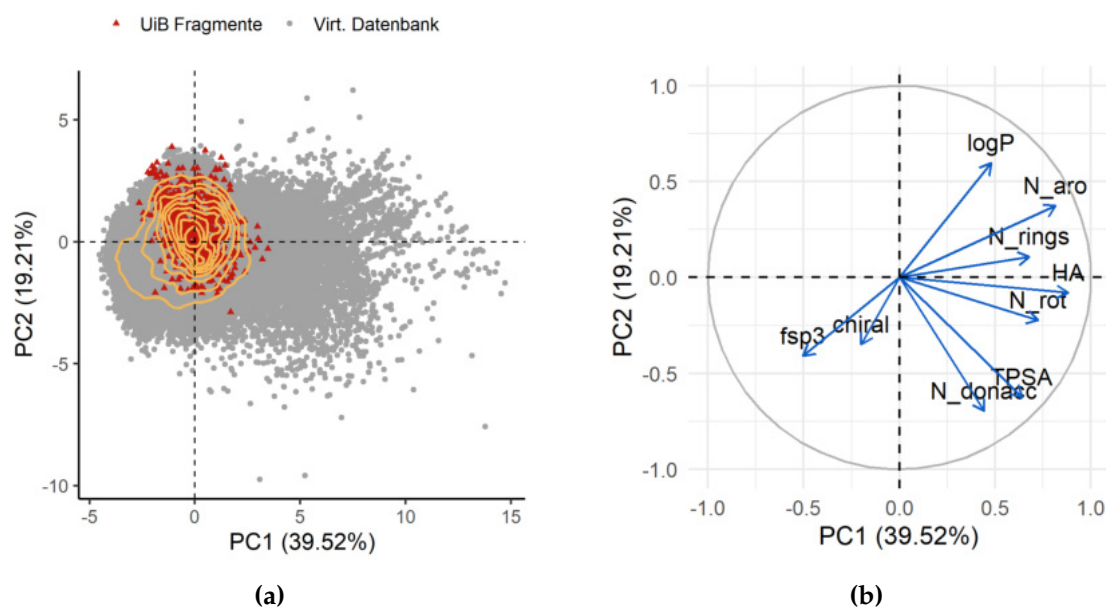


Abbildung 6.1: Chemischer Raum der verwendeten Fragmentbibliotheken

(a) Chemischer Raum der virtuellen Fragmentbibliothek (grau) und der UiB Fragmentbibliothek (rot). Die gelben Konturlinien zeigen die Dichte der Fragmente an, die zum Koordinatenursprung zunimmt. Die UiB Fragmentbibliothek deckt den chemischen Raum in diesem Bereich sehr gut ab.

(b) Beiträge der einzelnen Deskriptoren zu den Hauptkomponenten 1 und 2.

Abbildung 6.1 zeigt die Projektion der ersten beiden Hauptkomponenten einer PCA der beiden Fragmentbibliotheken auf Grundlage der oben genannten Deskriptoren. Die Hauptkomponente 1 und 2 decken hierbei mit 58.73 % über die Hälfte der gefundenen Variabilität des Datensatzes ab. In Grau sind die Fragmente der virtuellen Fragmentbibliothek und in Rot die Fragmente der UiB Fragmentbibliothek gezeigt. Die gelben Konturlinien stellen die Anzahl der vorgefundenen Moleküle dar; diese ist im Ursprung am höchsten. Die Verteilung zeigt einen deutlichen Schweif in Richtung positiver Werte der ersten Hauptkomponente. Die Beiträge der einzelnen Deskriptoren zu den Hauptkomponenten 1 und 2 illustrieren, dass dies vor allem auf Fragmente mit einer größeren Anzahl an Schweratomen (HA), Ringsystemen (N_rings bzw. N_aro) und rotierbaren Bindungen (N_rot) zurückzuführen ist. Die Fragmente der UiB Fragmentbibliothek decken mit 653 Molekülen den chemischen Raum, an dem sich die meisten Fragmente der virtuellen Fragmentbibliothek befinden (mehrere 10,000 Moleküle), sehr gut ab. Die virtuelle Fragmentbibliothek bietet für das virtuelle Screening dagegen eine einfache Möglichkeit, neue Bereiche des chemischen Raumes zu untersuchen, die mit dem biophysikalischen Screening hier nicht zugänglich gewesen wären.

6.1.2 Strukturbasierte Analyse des SHP-Bindemotivs und der Bindestelle von UFD1 und Derlin-1

Bei der in dieser Arbeit primär adressierten Region handelt es sich um die in der Nc-Subdomäne von p97 liegenden Interaktionsstellen für Kofaktoren, die über ein SHP-Motiv binden. Die PDB-Datenbank beinhaltet zwei Kristallstrukturen, die das SHP-Motiv der Kofaktoren UFD1 (PDB: 5B6C, 1.55 Å) [245] und Derlin-1 (PDB: 5GLF, 2.25 Å) [246] im Komplex mit der N-Domäne von p97 zeigen. Das SHP-Motiv von UFD1 wurde hierbei als Fusionsprotein mit p97-N kristallisiert, während für Derlin-1 ein von dem SHP-Motiv abgeleitetes Peptid benutzt wurde. Beide Strukturen bilden die Grundlage der Strukturanalysen und des darauf aufbauenden virtuellen Screenings.

6.1.2.1 Analyse der Kristallstrukturen der SHP-Motive von UFD1 und Derlin-1 im Komplex mit p97-N

Beide Strukturen besitzen mit 2.25 Å (Derlin-1 [246]) bzw. 1.55 Å (UFD1 [245]) eine gute bzw. sehr gute Auflösung, um molekulare Interaktionen untersuchen zu können. Die folgenden Sekundärstrukturelemente sind von Seite der Nc-Subdomäne an der Bindung beider Kofaktoren beteiligt (Abbildung 6.2a): Die β -Faltblatt-Stränge $\beta 7$, $\beta 10$ und $\beta 12$, die α -Helix $\alpha 4$, sowie die Schleifen L12 und L16. Beide SHP-Motive bilden eine kurze antiparallele- β -Faltblatt-Struktur mit $\beta 12$ der Nc-Untereinheit. Beide Motive werden zusätzlich durch die α -Helix $\alpha 4$ in ihrer Bindung stabilisiert. UFD1 und Derlin-1 binden beide in zwei Kavitäten von p97-N: die in dieser Arbeit als **SHP-I** bezeichnete Kavität, die von den Schleifen L12 und L16 gebildet wird; und die als **SHP-II** bezeichnete Kavität, die von Aminosäuren der $\beta 7$ -, $\beta 10$ - und $\beta 12$ -Faltblattstränge gebildet wird. In SHP-I interagieren die beiden Leucine L235^{UFD1} bzw. L248^{Derlin-1} der Consensus Sequenz (h(x)₁₋₂F/W(x)₀₋₁GxGx₂L) über hydrophobe Wechselwirkungen mit den Nc-Domänen Aminosäuren F131, F139, L140, V176, I182, sowie dem hydrophoben Teil der Seitenkette von K136. F228^{UFD1} bzw. W242^{Derlin-1} interagieren in SHP-II mit den Nc-Domänen Aminosäuren R113, I114, H115, E167 und H183. R113 wechselwirkt hier sowohl über den hydrophoben Teil der Seitenkette, als auch über Kationen- π -Wechselwirkungen mit den aromatischen Seitenketten von F228^{UFD1} bzw. W242^{Derlin-1}. Zusätzlich zeigt H183^{p97} eine T-förmige aromatische Wechselwirkung mit F228^{UFD1} und W242^{Derlin-1}. Die restlichen genannten Aminosäuren interagieren über ihre hydrophoben Teile mit UFD1 bzw. Derlin-1.

Die Bindestellen SHP-I und SHP-II liegen über den β 12-Faltblattstrang miteinander in Kontakt. Mit dem Peptidrückgrat von β 12 bilden UFD1 und Derlin-1 die folgenden Wasserstoffbrückenbindungen aus:

- zwischen O^1 -I182^{p97} und N^2 -N231^{UFD1} bzw. N^2 -F246^{Derlin-1}
- zwischen N^2 -I182^{p97} und O^1 -N231^{UFD1} bzw. O^1 -F246^{Derlin-1}
- zwischen O^1 -T180^{p97} und N^1 -L248^{Derlin-1}
- über ein Wassermolekül verbrückt zwischen O^1 -T180^{p97} und N^1 -L235^{UFD1}.

R247^{Derlin-1} bildet eine Salzbrücke mit D179^{p97}, wohingegen R234^{UFD1} in der Struktur keine Elektronendichte zeigt, sodass die Position der Seitenkette nicht bekannt ist. Des Weiteren können in beiden Komplexen π - π Interaktionen zwischen den Seitenketten von F131^{p97} und F246^{Derlin-1} bzw. N233^{UFD1} gefunden werden. F246^{Derlin-1} bzw. N233^{UFD1} zeigen in Richtung einer dritten Kavität, die in dieser Arbeit als **SHP-III** Bindestelle bezeichnet wird. Sie wird von beiden Kofaktoren nicht adressiert, enthält aber ein solvensexponiertes C184^{p97}, das für einen kovalenten Ansatz in der Masterarbeit von Stephan Böhler herangezogen wurde [280].

Mutationsstudien ergaben, dass die p97 Mutationen R113A, H115A, H183A, und F131A in Pulldown Assays eine stark reduzierte bzw. keine Bindung mehr zu den Kofaktoren zeigten [245, 246, 281]. L235A, R234A, F225A bzw. F228A-Mutationen von UFD1 verminderten bzw. zerstörten die Bindung zu p97 in Pulldown Assays und ITC Messungen [245, 281]. Für Derlin-1 zeigten L248A, R247A, G243A und G245A-Mutationen keine Bindung mehr zu p97 in Pulldown Assays [246]. Alle Mutationen stehen hier im Einklang mit den Beobachtungen in den Kristallstrukturen.

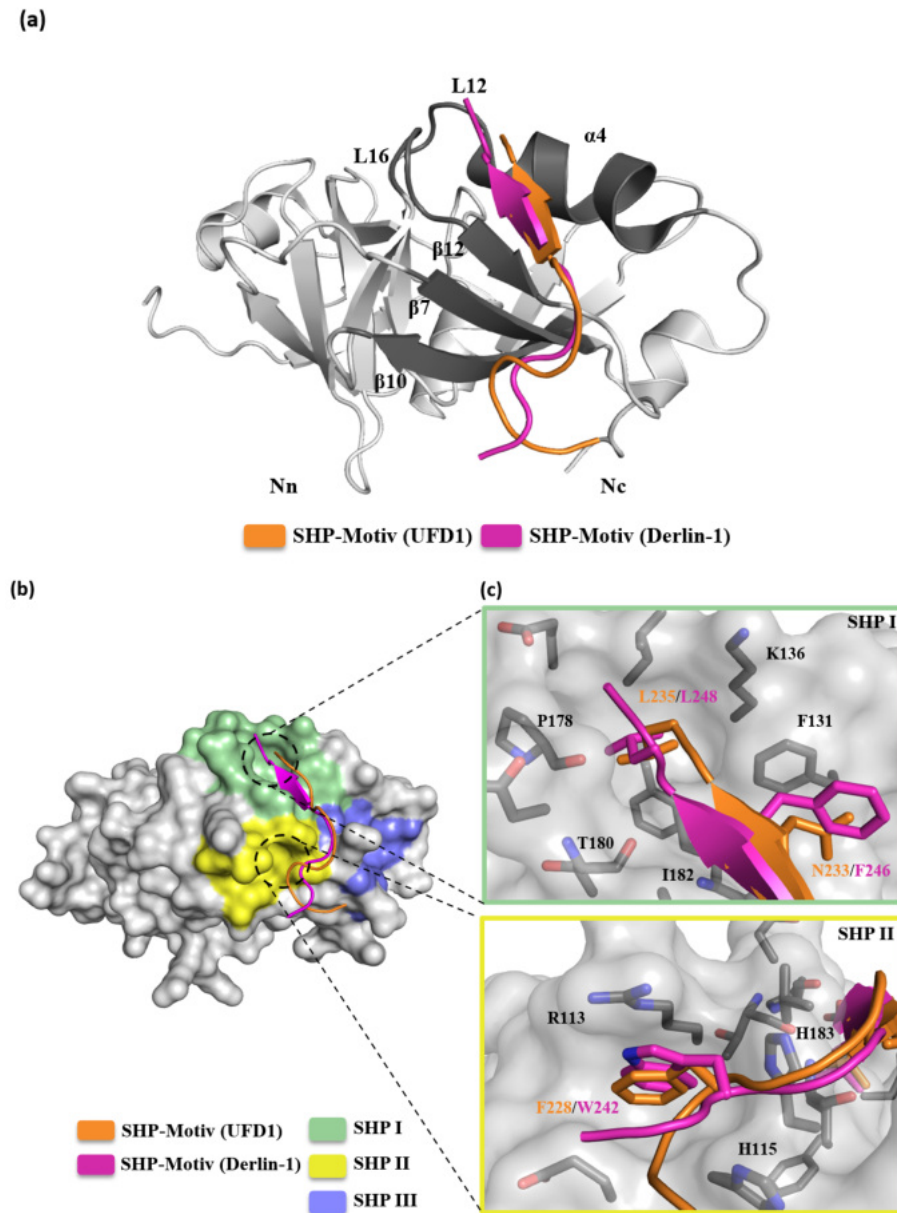


Abbildung 6.2: Die SHP-Bindestelle

(a) Darstellung der an der Interaktion mit den Kofaktoren UFD1 (PDB: 5B6C) bzw. Derlin-1 (PDB: 5GLF) beteiligten Sekundärelemente der N-Domäne.

(b) Lage der in der SHP-Binderegion gefundenen Kavitäten und deren für diese Arbeit verwendeten Bezeichnungen.

(c) Detaillierte Ansicht der beiden Kavitäten SHP-I und SHP-II mit den an Interaktionen beteiligten Aminosäuren der N-Domäne.

6.1.2.2 Interaktionspotential-Analysen für die SHP-I und SHP-II Bindestellen

Abbildungen 6.3 und 6.4 zeigen die Ergebnisse der gitterbasierten Interaktionspotential-Analyse für vier unterschiedliche chemische Sonden in der SHP-I bzw.

SHP-II Bindestelle. Die Abbildungen zeigen Isogitterdarstellungen für den jeweiligen Energiewert der Sonde, um visuell die Lage möglicher Interaktionsstellen zu verdeutlichen. Eine günstige Interaktionsstelle für Methyl- bzw. aromatische Kohlenstoff-Sonden liegt an den Positionen, die in den Komplexen von UFD1 bzw. Derlin-1 mit einer Leucinseitenkette adressiert werden und zeigen einen hydrophoben *hot-spot* für diesen Bereich an (rote Pfeile in Abbildung 6.4). Für eine Carboxyl-Sauerstoff Sonde kann ein deutlicher *hot-spot* in der Nähe von N^2 -I182 gefunden werden. Auch dieser *hot-spot* steht im Einklang mit der in den Komplexstrukturen vorgefundenen Wasserstoffbrückenbindung zwischen $N233^{UFD1}$ bzw. $F246^{Derlin-1}$ und N^2 -I182^{p97}. Ein weiterer starker *hot-spot* befindet sich oberhalb des von $L235^{UFD1}$ und $L248^{Derlin-1}$ adressierten Bereichs. In dieser Region findet sich in Kristallstrukturen und MD-Simulationen ein Netzwerk an konservierten Wassermolekülen (siehe 6.1.2.4). Ein positiv geladener sp^3 Stickstoff zeigt ein energetisches Minimum in der Nähe zu $P178^{p97}$ und $T180^{p97}$ an. Diese Bereiche fallen mit den polaren Wechselwirkungen zwischen dem Wassermolekül und $T180^{p97}$ in der UFD1-Komplexstruktur bzw. mit der beobachteten Wasserstoffbrückenbindung zwischen $L248^{Derlin-1}$ und $T180^{p97}$ zusammen. Ein weiterer Interaktionspunkt für diese Sonde liegt in Richtung der Aminosäure $K136^{p97}$, wird aber in den Komplexen nicht adressiert. Bindestelle SHP-II (Abbildung 6.3) lieferte nur für hydrophobe bzw. aromatische Sonden eine signifikante Interaktionsstelle, die in beiden Komplexstrukturen mit den aromatischen Aminosäuren Phenylalanin^{UFD1} bzw. Tryptophan^{Derlin-1} adressiert ist. Alle weiteren verwendeten Sonden bestätigten das eben Genannte und lieferten keine neuen Erkenntnisse.

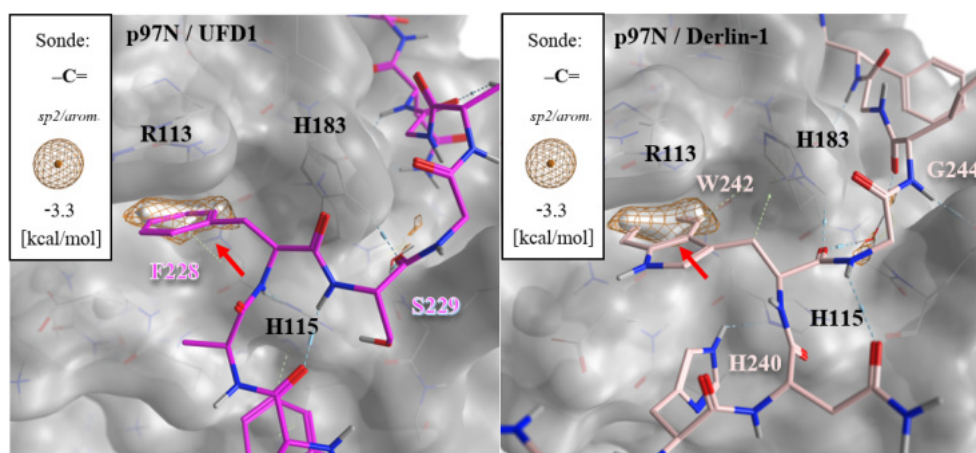


Abbildung 6.3: Interaktionspotentiale für die SHP-II Bindestelle

Interaktionspotentiale für eine aromatische Sonde in der SHP-II Region, basierend auf den beiden p97-N Komplexen mit UFD1 (violett) und Derlin-1 (pink). Für Details siehe Text.

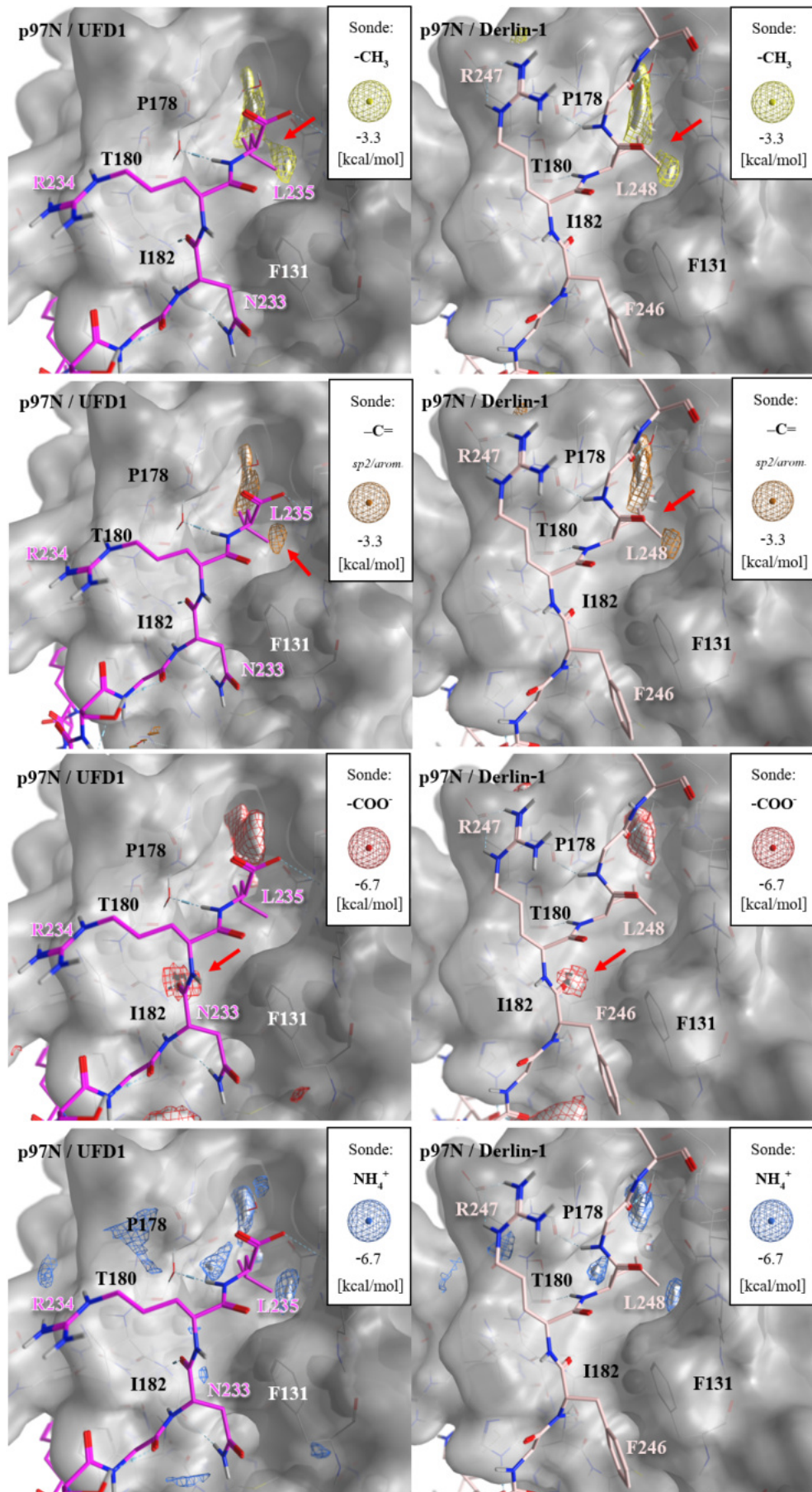


Abbildung 6.4: Interaktionspotentiale für die SHP-I Bindestelle

Interaktionspotentiale für vier verschiedene Sonden, basierend auf den beiden p97-N Komplexen von UFD1 (violett) und Derlin-1 (pink). Für Details siehe Text.

6.1.2.3 Analyse mittels des FTMap-Servers

Die Auswertung der Ergebnisse der FTMap-Methode unterstreichen die bisher gefundenen Interaktionsmöglichkeiten und zeigen gleichzeitig erste Orientierungen möglicher funktioneller Gruppen organischer Moleküle auf. In Abbildung 6.5 sind alle für die N-Domäne gefundenen Interaktionsbereiche für die beiden PDB Einträge 5B6C und 5GLF gezeigt. Für 5B6C ergaben sich insgesamt 14 Fragment-Cluster, von denen vier entweder innerhalb oder an biologisch nicht relevanten Stellen von p97 lagen. Für 5GLF ergaben sich 10 Fragment-Cluster, von denen ebenfalls vier an biologisch nicht zugänglichen Bereichen der N-Domäne lagen. Nicht biologisch relevante Stellen liegen in Bereichen die im nativen p97 durch die D1-Domäne maskiert sind und somit nicht für mögliche bindende Moleküle zugänglich sind. Alle nicht relevanten Fragment-Cluster wurden nicht weiter betrachtet. Die biologisch relevanten Fragment-Cluster lagen allesamt in oder in der Nähe der beiden Kofaktor Bindestellen für die UBX/L-Domänen bzw. das SHP-Motiv. Sie sind auf den Spalt zwischen der Nn- und Nc-Subdomäne bzw. nur auf die Nc-Subdomäne begrenzt. Innerhalb der SHP-Bindestelle konnte nur die SHP-I, nicht aber die SHP-II Bindestelle als mögliche Interaktionsstelle ausfindig gemacht werden. Neben den Stellen, die mit den Kofaktor Binderegionen zusammenfallen, konnten vereinzelte zusätzliche Interaktionsstellen lokalisiert werden. Hierzu gehört die SHP-III Bindestelle in direkter Nachbarschaft zu C184, die ausschließlich mit hydrophoben Fragmenten besetzt ist, sowie die hier als **La/b** bezeichneten Regionen, welche im funktionalen p97 in der Verbindungsregion von N- und D1-Domäne liegen.

Im Weiteren soll der Fokus auf den Ergebnissen liegen, die später in die Erstellung der Pharmakophormodelle für die SHP-I Bindestelle eingeflossen sind. Abbildung 6.6 zeigt im Detail die Interaktionen einiger Fragmente innerhalb der SHP-I Bindestelle. Insgesamt wurden die gleichen Interaktionsbereiche gefunden, die bereits durch die Interaktionspotential-Analysen in MOE identifiziert wurden. So liegen hydrophobe Fragmente im bereits gefundenen hydrophoben Bereich der SHP-I Bindetasche und sind mit den nativen Binderegionen von L235^{UFD1} bzw. L248^{Derlin-1} identisch. Des Weiteren konnte die für die Bindung wichtig erachtete Wasserstoffbrückenakzeptorfunktion zur Ausbildung einer polaren Wechselwirkung mit N²-I182 wieder gefunden werden. Die Wasserstoffbrückendonatorfunktionen im Bereich von O¹-T180 und O¹-P178 (wie auch O¹-A177) konnten

ebenfalls verifiziert werden. Zusätzlich zeigten die Fragmente mögliche Interaktionen in dem durch die Interaktionspotential-Analysen nur grob umrissenen Bereich zu Y143 hin. Hier wurde die Möglichkeit zur Ausbildung einer Wasserstoffbrückenbindung zu N^2 -E141, wie zu O^1 -L176 (nicht in Abbildung gezeigt) identifiziert. Eine visuelle Analyse der Überlagerung der Fragmente zeigt, dass die Bindung eines Moleküls mit zwei über einen Linker verknüpften Ringsystemen in dieser doch recht kleinen Bindetasche sterisch möglich wäre.

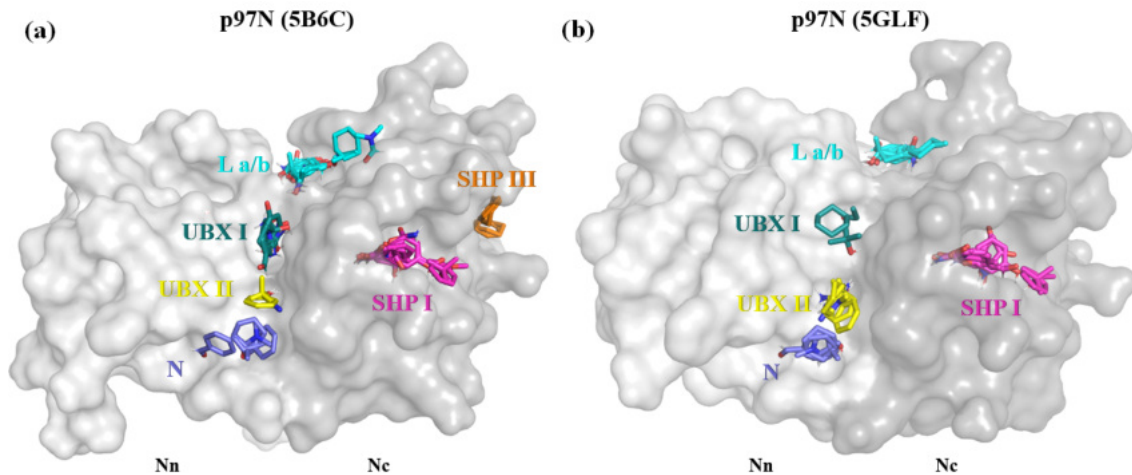


Abbildung 6.5: Mittels FTMap-Server vorhergesagte *hot-spots*

Gefundene bevorzugte Interaktionsbereiche basierend auf den Kristallstrukturen 5B6C (a) und 5GLF (b). Die *hot-spots* konzentrieren sich auf die Nc-Subdomäne und den Bereich zwischen der Nn- und Nc-Subdomäne. Sie stimmen mit den natürlichen Bindestellen für die UBXL-Domänen, VIM/VBM-Motive und dem SHP-Motiv überein. Interessanterweise wurden keine Bindestellen in der Nn-Subdomäne gefunden, wohingegen die Bindestellen La/b und SHP-III von den bisher bekanntem Kofaktoren nach derzeitigem Kenntnisstand nicht adressiert werden.

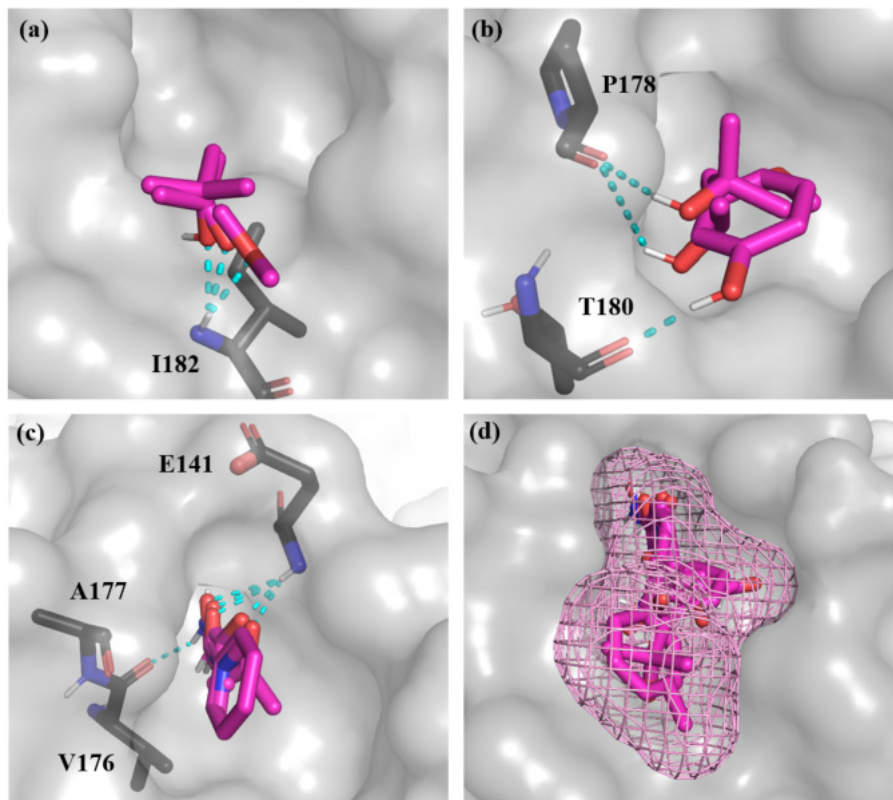


Abbildung 6.6: Gefundene Wechselwirkungen der FTMap-Sonden innerhalb der SHP-I Bindestelle

- (a) Wasserstoffbrückenbindungen zu N^2 -I182
 (b) Wasserstoffbrückenbindungen zu O^1 -T180 und O^1 -P178
 (c) Wasserstoffbrückenbindungen zu N^2 -E141 und O^1 -V176
 (d) Überlagerung aller Fragmente in der SHP-I Bindestelle, die ein qualitatives Bild von möglichen Liganden ergibt

6.1.2.4 Einfluss von Wassermolekülen innerhalb der SHP-I Region

Die Kristallstruktur von UFD1 zeigt ein koordiniertes Wasser zwischen O^1 -T180^{P97} und N^1 -L235^{UFD1}. Vor diesem Hintergrund sollen im Folgenden die möglichen Einflüsse von Wassermolekülen auf die SHP-I Bindestelle näher untersucht werden. Abbildung 6.7 zeigt die Überlagerung von neun PDB-Einträgen (4KDI, 3QQ7, 3QQ8, 3TIW, 3QC8, 3QWZ, 5EPP, 5B6C, 5GLF) auf die SHP-I Bindestelle. Bei allen Strukturen (mit Ausnahme von 5B6C und 5GLF) handelt es sich um apo-Strukturen bezüglich der SHP-I Bindestelle, in denen keine Kristallkontakte vorhanden sind. Die Überlagerung zeigt deutlich zwei hoch-konservierte Bereiche, die mit Wassermolekülen besetzt sind (**S1** und **S2**). Ein dritter Bereich (**S3**) liegt oberhalb von S2, weist allerdings eine deutlich höhere Variabilität der Positionen der Wassermoleküle auf. An diesen drei Stellen findet sich in allen Strukturen Wasser. Daneben existieren weniger besetzte Bereiche, wie **S4** in der Nähe von

P178 und T180. Hier finden sich bei mehr als der Hälfte der Strukturen Wassermoleküle. Auf Grundlage der Kristallstrukturen lag die Vermutung nahe, dass es sich vor allem bei S1, S2 und S3 um strukturell wichtige Wassermoleküle handeln könnte, die möglicherweise für die Konformation der SHP-I Bindestelle und für die Interaktion mit kleinen Molekülen in diesem Bereich von Bedeutung sind.

Um die Eigenschaften des Wassers und dessen Bedeutung für die Bindestelle SHP-I näher zu untersuchen, wurde ausgehend von auf der apo-Struktur von 5B6C eine 100 ns MD-Simulation in AMBER [216] durchgeführt. Die erhaltene Trajektorie wurde mittels des GIST Moduls in *cyptraj* (AmberTools18) analysiert. Abbildung 6.8 zeigt die Ergebnisse dieser Simulation. Rote Isogitterflächen zeigen Gitterpunkte an, an denen eine 15-fach höhere Dichte an Sauerstoffatomen von Wassermolekülen im Vergleich zum „Bulk“-Wasser zu finden ist. Weiße Isogitterflächen zeigen Gitterpunkte an, an denen eine 12-fach höhere Dichte an Wasserstoffatomen von Wassermolekülen im Vergleich zum „Bulk“-Wasser zu finden ist. Die gezeigten Dichten fallen mit den Regionen zusammen, an denen auch in den Kristallstrukturen Elektronendichten von möglichen Sauerstoffatomen zu finden sind. Unter Zuhilfenahme der aus der Simulation erhaltenen Wasserstoffdichten lässt sich für einige Bereiche eine Orientierung der Wassermoleküle zeigen. So zeigen die Dichten bei P178 und T180 eine deutliche Orientierung in Richtung der Carbonylgruppen der genannten Aminosäuren. Ähnliches findet man auch für Dichten in der Nähe der Aminosäuren Y143 und E141. Zusammengekommen lassen die Dichten ein Wassernetzwerk innerhalb der Bindetasche erkennen. Um dieses hier besser zu zeigen, wurden zufällige Momentaufnahmen aus der Trajektorie der MD-Simulation ausgewählt. Die Momentaufnahmen wurden in MOE mittels des Amber14:ETH Kraftfeldes mit einem Gradienten von $0.0001 \text{ kcal mol}^{-1} \text{ \AA}$ und gesetztem Tethering von 0.5 \AA energieminiert. In den Konfigurationen konnten für sämtliche Dichtemaxima Wassermoleküle lokalisiert werden. Diese zeigen das soeben, anhand der Dichten postulierte Wassernetzwerk das von den Wassermolekülen S1 bis S3 gebildet wird und exemplarisch in Abbildung 6.8 anhand einer Momentaufnahme gezeigt ist. Möglicherweise hat dieses Wassernetzwerk einen Einfluss auf die Struktur der Bindestelle SHP-I, gerade im Hinblick auf die Konformation der Schleifen L12 und L16. Durch überbrückte Wasserstoffbrückenbindungen zwischen dem Rückgrat der Peptidbindungen der Schleifen L12 und L16 könnte die Konformation der SHP-I Bindestelle hier stabilisiert werden. Zur Bestätigung dieser These bedarf es zusätzlicher MD-Simulationen, wobei der konkrete Einfluss des Wassers über *mixed-solvent* Simulationen geklärt werden könnte [282].

Für die Erstellung von Pharmakophormodellen, welche für das später durchgeführte virtuelle Screening verwendet wurden, wurden die durch GIST gefundenen und bewerteten Wassercluster mit der Kristallstruktur 5B6C überlagert, um eine Entscheidung treffen zu können, welche der in der Kristallstruktur enthaltenen Wassermoleküle mit in die Modelle aufgenommen werden sollten. Abbildung 6.8 zeigt, dass sich in der Nähe der Wassercluster der MD-Simulation stets Wassermoleküle der Kristallstruktur wiederfinden (Cluster S1, S2, S3, S4). Die Bereiche der Cluster **S6** und **S7** sind in der Kristallstruktur durch UFD1 besetzt. Es fällt auf, dass S7 genau mit der Position der Wasserstoffbrückenakzeptorfunktion an N²-I182 zusammenfällt. Die Region um S6 ist mit der Seitenkette von L235 besetzt. Interessanterweise findet man hier nur in zwei der sieben apo-Kristallstrukturen Wassermoleküle, was zu der Vermutung führt, dass durch die eingeschränkte MD-Simulation dieser Bereich möglicherweise überbewertet wird.

Der Cluster **S3*** stellt eine nur in der MD-Simulation gefundene, besonders günstige Bindestelle für Wasser dar. Durch eine Wasserstoffbrückenbindung zwischen der Hydroxylgruppe von Y143 mit O⁵-E141 wird E141 so positioniert, dass eine sehr günstige Koordinationsstelle für ein Wassermolekül zwischen N²-E141, O⁵-E141 und einem weiteren Wassermolekül an Position S3 entsteht. Da in keiner apo-Strukturen diese Konformation mit einer Wasserstoffbrücke zwischen Y143 und E141 beobachtet werden konnte, wird S3* als ein Artefakt der MD-Simulation angesehen.

Für die in der Kristallstruktur vorhandenen Wassermoleküle wurden folgende Schlussfolgerungen gezogen: Wassermolekül S4 stellt ein interessantes Wassermolekül dar, da man es entweder mit einem organischen Molekül verdrängen oder an dieses binden könnte. Beide Möglichkeiten finden sich auch in den Komplexstrukturen UFD1 (koordiniert) und Derlin-1 (nicht koordiniert). S1 scheint zwar am leichtesten verdrängbar zu sein, liegt aber so tief im Protein vergraben, dass es sterisch nur schwer für ein Molekül erreichbar scheint. Das Gleiche gilt für Wassermolekül S2. S3 stellt das bestbewertete Wassermolekül dar. Es könnte trotz der möglicherweise energetisch ungünstigen Desolvatisierung dieser Stelle interessant sein, mittels eines Fragments diese Interaktion zu imitieren, da dieses hierdurch sehr vergraben zwischen den Schleifen L16 und L12 liegen würde. Als kritischer Punkt sei hier angemerkt, dass die Simulation eines Konformationen-Ensembles (wie z.B. aller bekannten Kristallstrukturen) der SHP-Bindestelle vermutlich ein detaillierteres Bild der Wasser Eigenschaften, gerade auch bezüglich der Energiewerte zuließe. Solche Simulationen, die in Summe im μ s-Bereich liegen würden, wären mit dem zum Zeitpunkt der Studie verfügbaren Computer-Ressourcen extrem langwierig gewesen, weshalb auf diese

verzichtet wurde. Trotzdem findet man ein größtenteils stimmiges Bild zwischen der Simulation und den verfügbaren Kristallstrukturen vor.

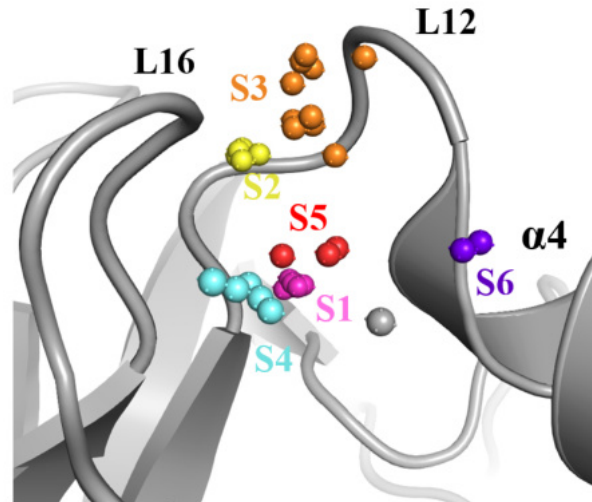


Abbildung 6.7: Wassermoleküle in bekannten p97-N Kristallstrukturen

Die Abbildung zeigt das Ergebnis der Überlagerung von neun in der PDB hinterlegte Kristallstrukturen hinsichtlich der Wassermoleküle in der SHP-I Bindestelle. Für Details siehe Text.

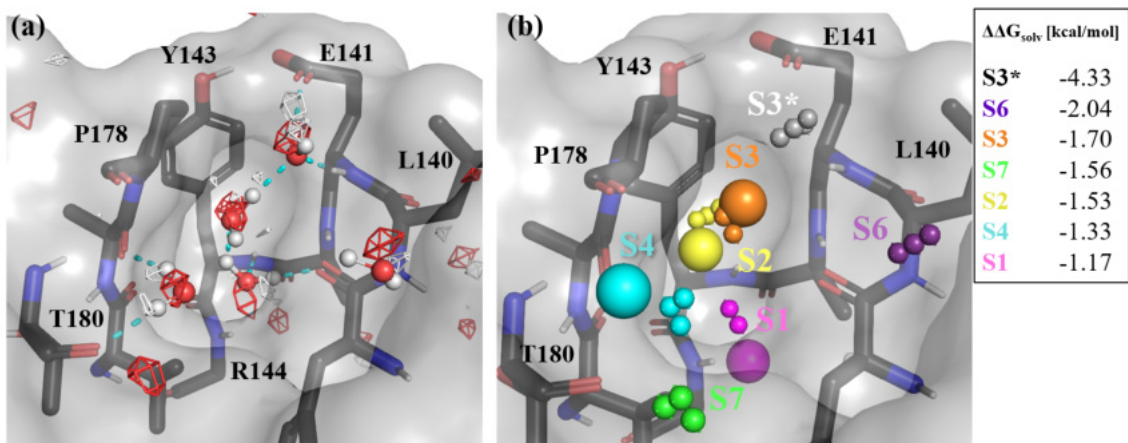


Abbildung 6.8: Gefundene Wasserstrukturen auf Grundlage einer 100 ns MD-Simulation

(a) Gefundene Dichten für Wassermolekül-Sauerstoffatome (rote Gitterdarstellung, 15-fach höhere Dichte im Vergleich zum Bulk-Wasser) und Dichten für Wassermolekül-Wasserstoffatome (weisse Gitterdarstellung, 12-fach höhere Dichte im Vergleich zum Bulk-Wasser). Anhand der Dichten lassen sich deutlich bevorzugte Orientierungen für Wassermoleküle erkennen. Überlagert sind die Dichten mit einem zufällig aus der Simulation ausgewählte Momentaufnahme. Die Wassermoleküle liegen hier deutlich im Bereich der vorgefundenen Dichten und zeigen ein Netzwerk an koordinierten Wassermolekülen. (b) Mittels GIST gefundene Wassercluster (kleine Sphären) und deren Energiewerte, überlagert mit den in der Kristallstruktur 5B6C gefundenen Wassermolekülen (große Sphären).

6.1.2.5 Zusammenfassung der Strukturanalysen

Zusammenfassend lassen sich folgende Schlussfolgerungen für mögliche Interaktionen anhand der oben durchgeführten Analysen festhalten:

1. die Bindestelle SHP-I ist größer und enthält mehr potentielle Interaktionsstellen als die Bindestelle SHP-II. SHP-II wurde zudem mittels des FTMap-Algorithmus nicht als eine mögliche Bindestelle identifiziert;
2. die SHP-I Bindestelle kann in einen eher hydrophoben Teil in Richtung F131 und einen eher polaren Teil in Richtung Y143 eingeteilt werden;
3. die Wasserstoffbrückenakzeptorfunktion zu N^2 -I182 scheint wichtig für die Bindung der beiden von UFD1 bzw. Derlin-1 abgeleiteten Peptide zu sein. Zudem konnte sie in der Interaktionspotential-Analyse und den Ergebnissen des FTMap-Algorithmus als ein eindeutiger *hot-spot* wiedergefunden werden;
4. eine Wasserstoffbrückendonorfunktion zu O^1 -P178 bzw. O^1 -T180 stellt eine weitere interessante Interaktionsstelle dar, die entweder über ein verbrückendes Wasser oder direkt adressiert werden könnte;
5. Wassermoleküle scheinen in der SHP-I Kavität eine wichtige Rolle zu spielen und ein Wassernetzwerk innerhalb des polaren Teils der Kavität auszubilden. Dieses Wassernetzwerk könnte als dreidimensionale Vorlage für polare Wechselwirkungen von organischen Molekülen dienen.

Anhand dieser Ergebnisse wurde ein virtuelles Fragmentenscreening an der SHP-I Bindestelle durchgeführt.

6.1.3 Virtuelles Screening an der SHP-I Bindestelle

6.1.3.1 Pharmakophormodelle und Pharmakophor-Screenings

Auf Basis der eben vorgestellten Strukturanalyse wurden verschiedene Pharmakophormodelle entwickelt, mit denen die virtuelle Fragmentbibliothek für das spätere Docking vorselektiert wurde. Vier verschiedene Pharmakophormodelle wurden letztendlich für die SHP-I Bindestelle erstellt, deren schematischer und struktureller Aufbau in den Abbildungen 6.9 bzw. 6.10 gezeigt wird. Alle Modelle enthielten zwei Sphären mit aromatischen/hydrophoben Eigenschaften an den Positionen der Atome $C^{5/5'}$ -L235^{UFD1} bzw. L248^{Derlin-1} der Kofaktoren. Von dieser hydrophoben „Ankerregion“ ausgehend wurden zwei Modellstränge erstellt: der eine Modellstrang enthielt die in den Komplexen und Interaktionsanalysen gefundene Wasserstoffbrückenakzeptorfunktion zu N^2 -I182, der andere Modellstrang eine Wasserstoffbrückenakzeptorfunktion an der Position des Wassermoleküls S3. Beide Modellstränge bekamen als dritte Funktion entweder eine Wasserstoffbrückendonorfunktion zum Wassermolekül S4, mit der die überbrückte Bindung im p97-N-UFD1 Komplex imitiert werden sollte, oder eine Wasserstoffbrückendonorfunktion an der Position des Wassermoleküls S4, mit der die Bindung von Derlin-1 nachgeahmt werden sollte.

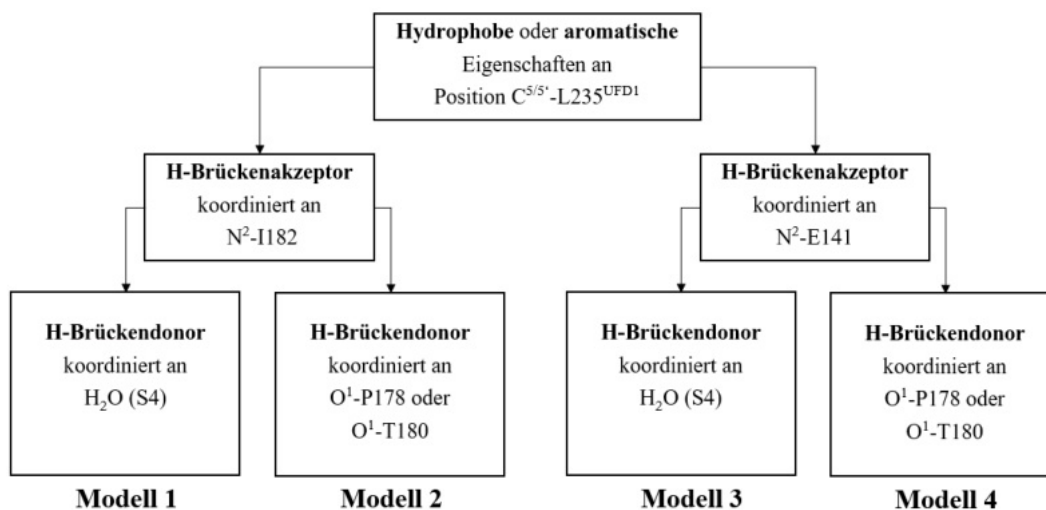


Abbildung 6.9: Schematischer Aufbau der Pharmakophormodelle 1 bis 4

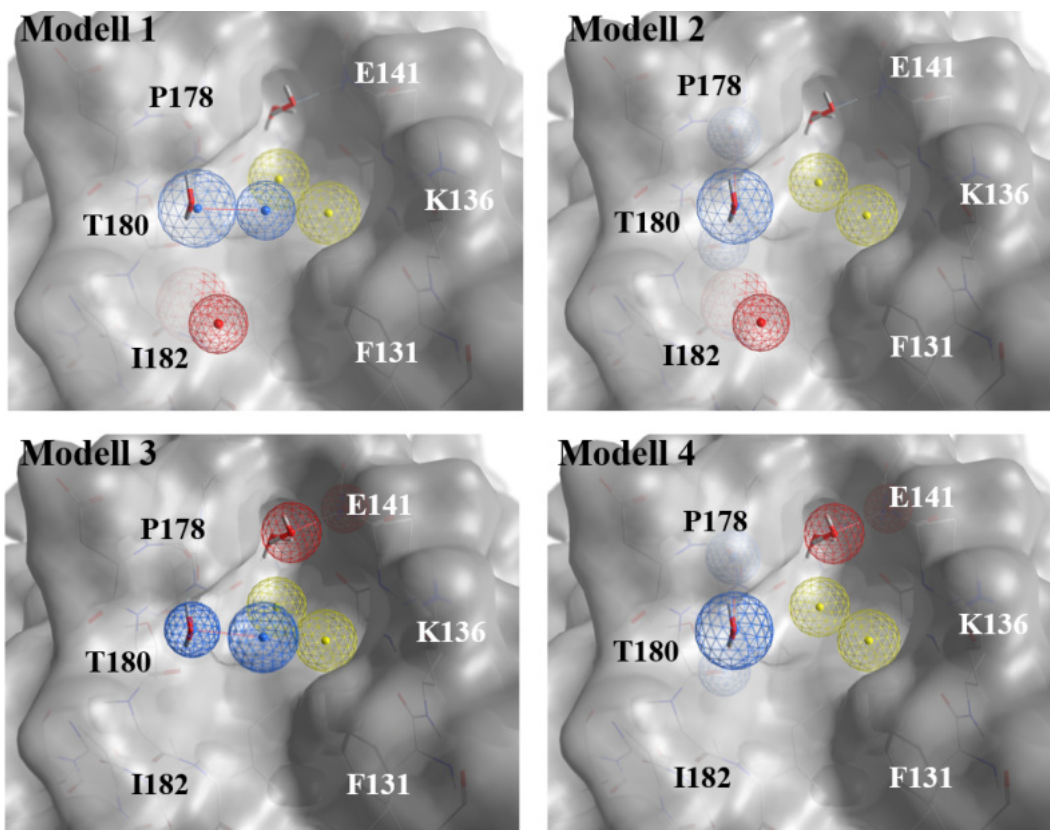


Abbildung 6.10: Struktureller Aufbau der Pharmakophormodelle

Sphären in Blau stellen Wasserstoffdonatorfunktionen, in Rot Wasserstoffakzeptorfunktionen und in Gelb hydrophobe bzw. aromatische Funktionen dar.

Zusätzliche Pharmakophoreigenschaften ergaben keinen sichtbaren Nutzen, da die in der Bibliothek vorhandenen Moleküle größtenteils nicht die nötige Größe hatten, um noch mehr Pharmakophorfunktionen räumlich abdecken zu können. Im weiteren Verlauf wurden zum Verfeinern der Modelle einige Volumenausschließende Sphären definiert (d. h. Bereiche, an denen kein Ligandatomben liegen durfte), um der Form der Bindetasche Rechnung zu tragen. Die Ergebnisse des Pharmakophor-Screenings mit der jeweiligen Anzahl an Treffern pro Modell sind in Tabelle 6.1 gezeigt. Von 86,660 Molekülen wurden für alle Modelle zusammengerechnet 15,225 (17.57%) Moleküle gefunden, die den Pharmakophormodellen entsprachen.

Tabelle 6.1: Ergebnisse des Pharmakophor-Screenings

Modell	1	2	3	4
Treffer	2595 (2.99%)	8954 (10.33%)	2819 (3.25%)	857 (0.99%)

6.1.3.2 Docking

Die Ergebnisse des virtuellen Screenings mit den Darstellungen der Zwischenschritte und Auswahlkriterien sollen in dem folgenden Abschnitt erörtert werden. Für die jeweiligen Hitlisten der einzelnen Pharmakophormodelle wurde, ein Docking mit dem Programm GOLD [100] durchgeführt. Nach dem ersten Dockinglauf ergaben sich die in Abbildung 6.11 gezeigten Verteilungen der $Score_{HA}$ Werte der jeweiligen Scoringfunktionen. Hierbei ist für die ChemPLP-Funktion eine Bindepose besser, je höher der $Score_{HA}$ ist, während für die ASP (und auch später gezeigte DSX)-Funktion eine Bindepose besser bewertet wird, je negativer der $Score_{HA}$ ist. Die Verteilung der ASP-Scoringfunktion zeigt qualitativ den Verlauf einer normalverteilten Gaußkurve, mit einer nur minimal ersichtlichen Schwanzbildung (*tailing*) hin zu negativeren $Score_{HA}$ -Werten. Alle Modelle erzielen in etwa den gleichen Mittelwert (siehe Tabelle 6.2). Die Verteilungen der ChemPLP-Scoringfunktion zeigen hingegen ein deutlicheres *tailing* hin zu höheren $Score_{HA}$ -Werten. Dies trifft besonders für die Hitlisten der Modelle 3 und 4 zu.

Für das zweite Docking wurden diejenigen Fraktionen an Molekülen ausgewählt, die rechts des gesetzten Grenzwertes zu höheren $Score_{HA}$ -Werten lagen. Als Grenzwert wurden die jeweils 10% besten Fragmente anhand ihrer $Score_{HA}$ -Werten ausgewählt (Tabelle 6.2). Die so erhaltenen Fragmente wurden anschließend, wie in Kapitel 5.1.5.3 beschrieben, mit der PLIF-Funktion selektiert. Einige Fragmente zeigten hier Wechselwirkungen mit sehr solvensexponierten Aminosäuren oder gar keine polaren Interaktionen und wurden daher aussortiert. Nur eine kleinere Anzahl an Fragmenten wurde in zwei oder mehr Modellen gefunden. Auch der Anteil an Fragmenten die beiden Scoringfunktionen hoch bewertet waren, war relativ klein (Tabelle 6.3). Die ausgewählten Fragmente (Tabelle 6.3) wurden einem zweiten Docking mit einer fünffach höheren Anzahl an generierten Posen unterzogen. Mit dieser höheren Anzahl an generierten Posen sollte eine Konvergenz der erhaltenen besten Posen sichergestellt werden. Abbildung 6.12 zeigt die Verteilung der $Score_{HA}$ -Werte für den zweiten Docking-Lauf. Die höhere Posengeneration ergab hierbei eine leichte Verschiebung der $Score_{HA}$ -Werte hin zu höher positiven bzw. negativeren Werten. Eine höhere Posen Generierung ergab demnach noch besser bewertete Bindemodi der Fragmente, die im ersten Durchlauf nicht gefunden wurden.

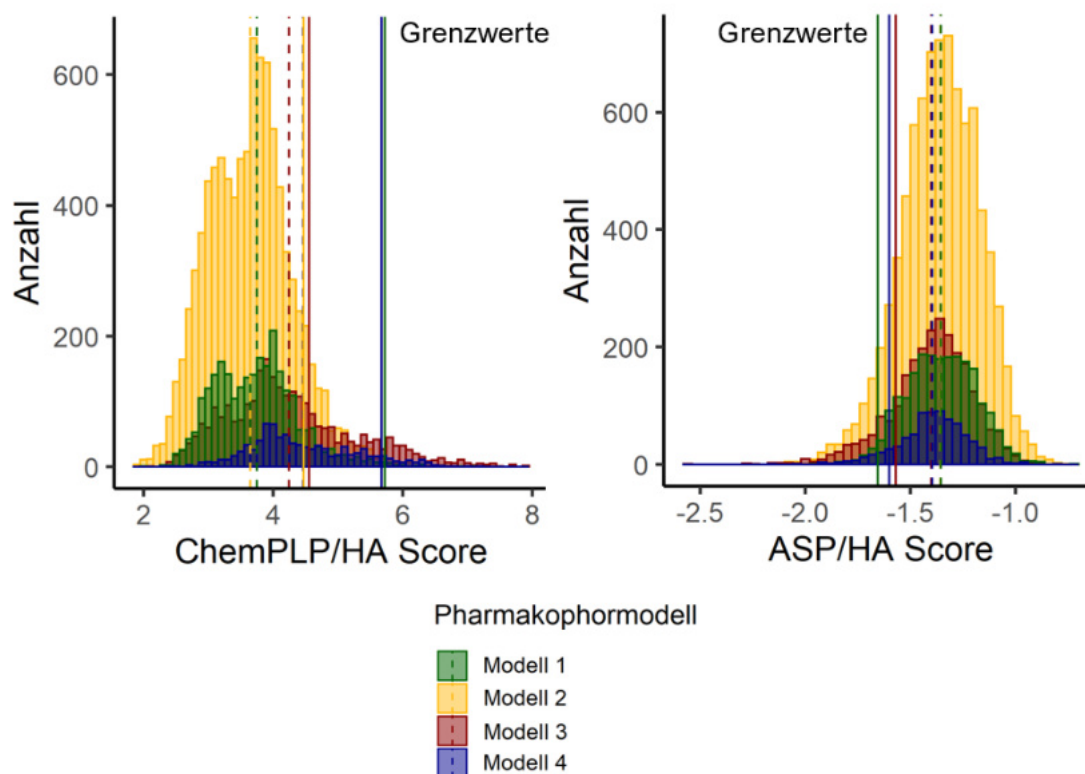


Abbildung 6.11: Verteilung der $Score_{HA}$ Werte für das erste Docking

Links die Ergebnisse für die ChemPLP-Scoringfunktion und rechts für die ASP-Scoringfunktion. Die gestrichelten Linien zeigen die Lage der $Score_{HA}$ -Mittelwerte an, die durchgehenden Linien die gesetzten Grenzwerte. Alle Moleküle links (ChemPLP) bzw. rechts (ASP) des jeweiligen Grenzwertes wurden verworfen.

Tabelle 6.2: Mittelwerte und Grenzwerte basierend auf den $Score_{HA}$ Werten des ersten Dockinglaufs

Modell	1	2	3	4
Mittelwert (ASP)	-1.354	-1.357	-1.401	-1.395
Grenzwert (ASP)	-1.571	-1.600	-1.655	-1.602
Mittelwert (ChemPLP)	1.354	1.357	1.401	1.395
Grenzwert (ChemPLP)	4.552	4.471	5.723	5.671

Tabelle 6.3: Anzahl Moleküle für den zweiten Dockinglauf nach Selektion durch die PLIF-Analyse

Die Zahlen rechts der vertikalen Linie gibt an, wie oft gleiche Fragmente in mehreren Modellen gefunden wurden. In der letzten Zeile ist die Anzahl an Fragmenten die in beiden Scoringfunktionen pro Modell gefunden wurden gezeigt.

Modell	1	2	3	4	in 2	in 3	in 4
ChemPLP	118	185	128	46	30	2	0
ASP	136	252	51	38	21	1	0
In beiden Scoringfunktionen	16	49	0	5			

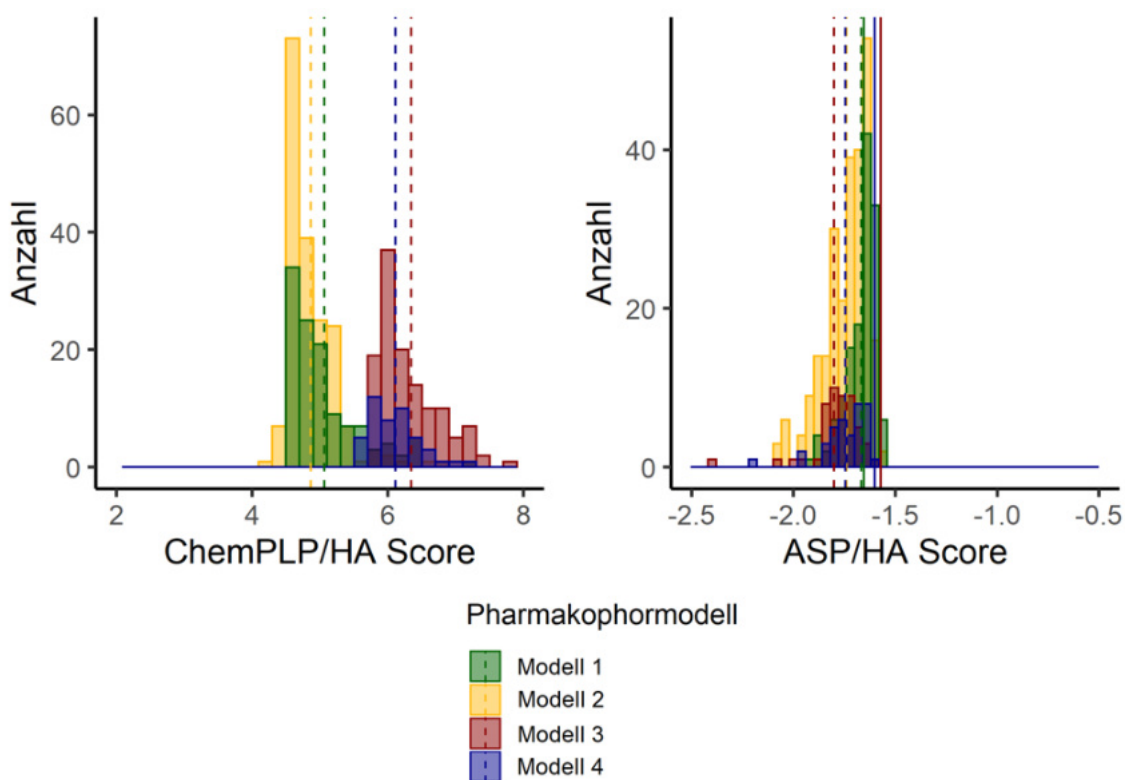


Abbildung 6.12: Verteilung der $Score_{HA}$ Werte für das zweite Docking

Links die Ergebnisse für die ChemPLP-Scoringfunktion und rechts für die ASP-Scoringfunktion. Die gestrichelten Linien zeigen wieder die Lage der $Score_{HA}$ Mittelwerte an.

Nach dem durchgeführten Rescoring der Bindeposen mittels DSX [109] wurden die jeweils best-bewerteten Bindeposen (auf Grundlage des $Score_{HA}^{DSX}$) eines jeden Fragmentes bezüglich der Fragestellung analysiert, ob die aufgestellten Pharmakophormodelle sich auch in den Docking Ergebnissen wiederfinden ließen. Die visuelle Inspektion der Dockingposen zeigte für alle Pharmakophormodelle eine hydrophobe Interaktion an der Stelle der Wechselwirkungen der Leucinseitenketten der Kofaktoren. Des Weiteren zeigten die Bindeposen der ASP-Scoringfunktion (nach DSX-Rescoring) eine höhere Präferenz, aromatische Systeme über eine π - π Wechselwirkung mit der Seitenkette von F131 wechselwirken zu lassen, als dies in den Bindeposen der ChemPLP-Scoringfunktion (nach DSX-Rescoring) der Fall war. Nach dieser visuellen Inspektion der Ergebnisse wurden die Bindeposen quantitativ mittels des PLIF-*tools* in MOE auf ihre polaren Interaktionen mit p97-N hin untersucht. Abbildungen 6.13 und 6.14 zeigen die vorgefundenen polaren Interaktionen der einzelnen Dockingläufe.

Für das erste Pharmakophormodell zeigten die Bindeposen der ChemPLP-Scoringfunktion besser die Interaktionen des Modells, als die Bindeposen der ASP-Scoringfunktion. So fand sich bei den ChemPLP Ergebnissen immerhin bei 37 % die favorisierte Interaktion mit dem Wasser S4. Zusätzlich zeigten die Bindeposen der Fragmente neben der direkten Koordination an das Wassermolekül auch zusätzliche direkte Interaktionen mit den Aminosäuren P178 und T180. Die ASP-Scoringfunktion lieferte hier nur einen verschwindend geringen Anteil der gewünschten Interaktionen. Die als essentiell erachtete Wasserstoffbrückenakzeptorfunktion zu I182 fand sich dagegen mit 87% in fast jeder der ASP-Posen, während sie bei den ChemPLP-Posen nur mit 32% vertreten war.

Die favorisierten Wechselwirkungen des zweiten Pharmakophormodells fanden sich sowohl in den ChemPLP als auch in den ASP-Posen wieder, wobei auch hier die Interaktionen in ChemPLP-Posen prozentual häufiger vorgefunden wurden als in den ASP-Posen. Bezüglich des dritten Pharmakophormodells zeigten die ChemPLP-Posen eine höhere Anzahl direkter Interaktionen mit den Aminosäuren A177, P178 und T180, als mit dem favorisierten Wassermolekül. Die ASP-Posen schnitten hier noch schlechter ab, zeigten aber in fast der Hälfte der Posen die Interaktion zu N^1 -E141.

Das vierte Pharmakophormodell lieferte für ASP nur in etwa 20% aller Bindeposen die gewünschte Wasserstoffbrücke zu N^2 -E141; bei ChemPLP-Posen war dies mit 4% vernachlässigbar. Interessanterweise zeigen die Posen beide Scoringfunktionen Interaktionen mit I182, wobei in den ASP-Posen bevorzugt eine Wasserstoffbrückenakzeptorfunktion, in den ChemPLP-Posen eher eine Wasserstoffbrückendonorfunktion zu I182 zu finden ist.

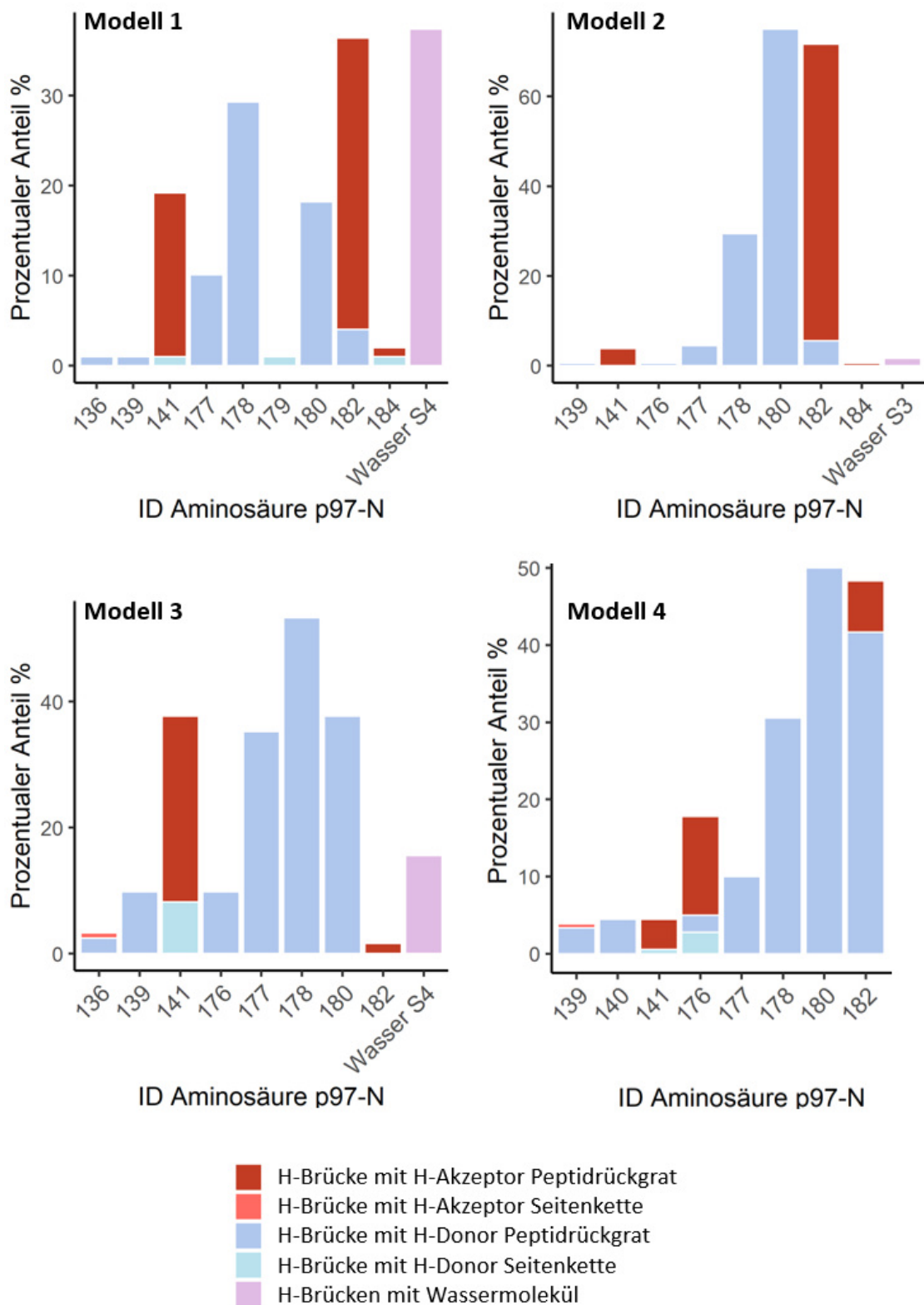


Abbildung 6.13: ChemPLP-Ergebnisse der PLIF-Analysen im Hinblick auf die verwendeten Pharmakophormodelle

Die Zahlen geben die relative Häufigkeit gefundener Interaktionen in Bezug zur Anzahl aller im zweiten Dockinglauf untersuchten Moleküle wieder.

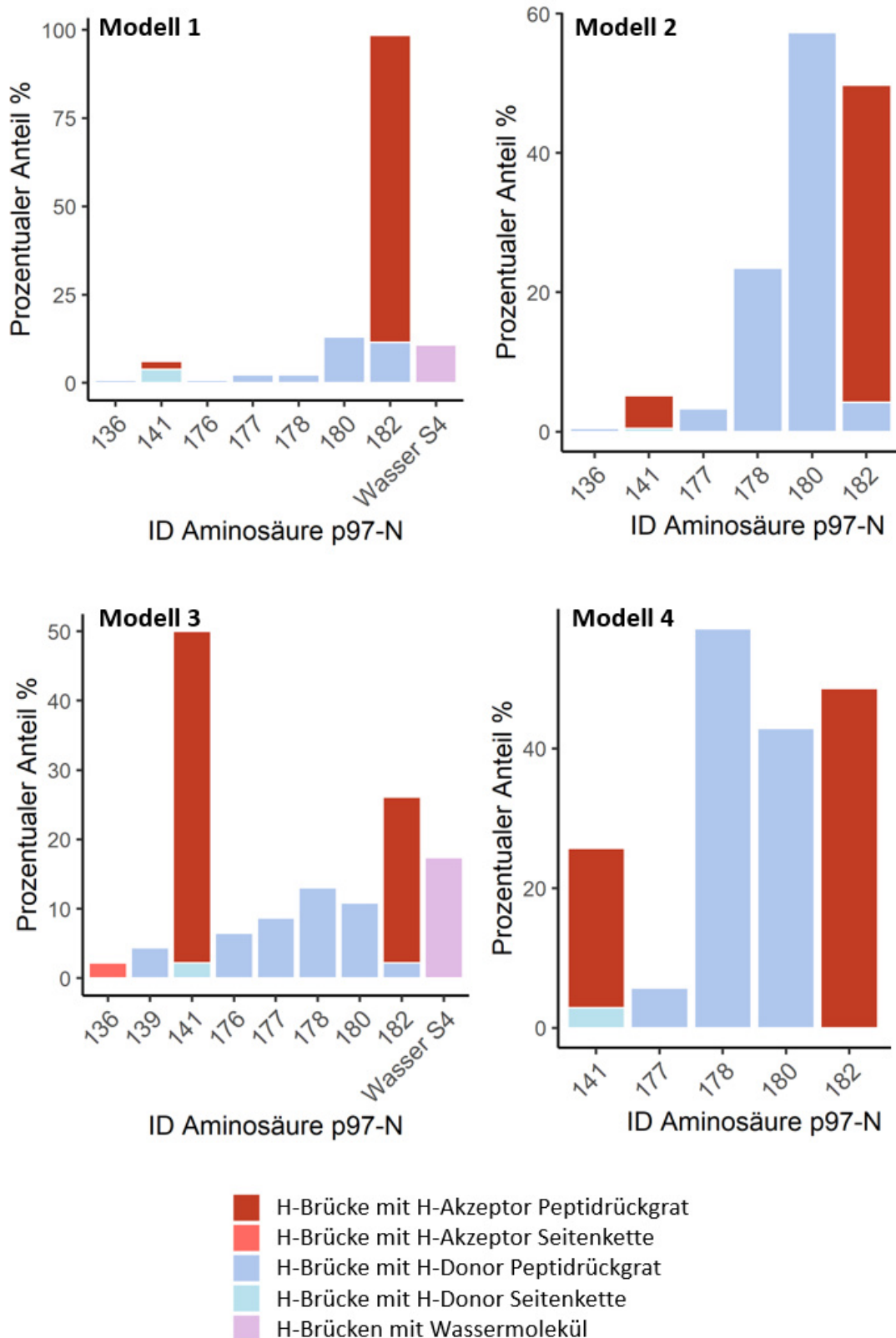


Abbildung 6.14: ASP-Ergebnisse der PLIF-Analysen im Hinblick auf die verwendeten Pharmakophormodelle

Die Zahlen geben die relative Häufigkeit gefundener Interaktionen in Bezug zur Anzahl aller im zweiten Dockinglauf untersuchten Moleküle wieder.

In allen Dockingläufen konnten keine weiteren dominanten Interaktionen in den Bindeposen gefunden werden, die nicht schon in den Pharmakophormodellen berücksichtigt wurden. Ausgenommen sind hier Interaktionen mit den solvens-exponierten Seitenketten von K136 oder E141, die einerseits aufgrund ihrer ungünstigen Lage, andererseits wegen nicht eindeutiger Elektronendichten in den bekannten Kristallstrukturen und der damit verbundenen Unsicherheit in der Position der Seitenketten nicht weiterverfolgt wurden.

Die weitere Auswahl an Molekülen für anschließende MD-Simulationen wurde auf Basis der Pharmakophormodelle getroffen. Eine alleinige Auswahl anhand der Scoring-Werte wurde nicht vorgenommen, da die erhaltenen Scoring-Werte sich nicht stark genug unterschieden und besonders gut bewertete Fragmente nicht die favorisierten Interaktionen zeigten.

Zur Auswahl wurden die Moleküle erneut auf Grundlage der Pharmakophormodelle analysiert, mit dem Kriterium, dass mehrere Interaktionen gleichzeitig in einer Bindepose zu finden sein mussten. Tabelle 6.4 gibt eine Übersicht, welche Kriterien im Einzelnen für ein Molekül erfüllt sein mussten, um nicht aussortiert zu werden. Nach diesem Filterschritt lag die Anzahl der Moleküle in einem sinnvollen Bereich, für die darauffolgenden MD-Simulationen. Visuell wurde nun mittels folgender Kriterien geprüft auf:

- Sinnvolle tautomere und protomere Zustände für pH 7. Geprüft wurde hier unter Zuhilfenahme des Programms MoKa [226–228].
- Energetisch sinnvolle Konformationen der Moleküle. So sollten konjugierte Systeme planar und nicht gegeneinander verdreht sein. Ungünstige inter- und intramolekulare Kontakte sollten vermieden werden (geprüft wurde dies anhand der visuellen Kontrolle der ermittelten Paarpotentiale des DSX-Scores in Pymol).

Tabelle 6.4: Selektionskriterien für die PLIF-Analyse**d** = Wasserstoffbrückendonator, **a** = Wasserstoffbrückenakzeptor**&** = Logischer Operator *UND*, **→** = *Interaktion an Atom X***,** = Logischer Operator *ODER*

Modell	Kriterien	Nr. ChemPLP	Nr. ASP
1	a → N-I182 & (d → O-A177, O-P178, O-T180, O-S4)	20	34
2	a → N-I182 & (d → O-A177, O-P178, O-T180)	105	57
3	a → N-E141 & (d → O-A177, O-P178, O-T180, O-S4)	13	22
4	(a → N-E141, N-I182) & (d → O-A177, O-P178, O-T180)	43	35
Summe^a		181	148

^a Die Anzahl an Fragmenten beinhaltet einen geringen Anteil an Fragmenten die in beiden Scoringfunktionen bzw. in zwei Modellen gefunden wurden. Aufgrund des geringen Anteils unter 10% und weil dies kein späteres Auswahlkriterium darstellte, wird hier auf eine Aufschlüsselung verzichtet.

6.1.3.3 MD-Simulationen

Letztendlich wurden 29 Treffer der ChemPLP-Posen und 20 Treffer der ASP-Posen für 25 ns MD-Simulationen ausgewählt. Diese sollten zeigen, ob die im Docking gefundene Bindepose unter expliziten Solvens-Bedingungen stabil sind. Tabellen 6.5 und 6.6 zeigen die Docking-Ergebnisse der für die Simulationen ausgewählten Fragmente. Aufgelistet sind hier der beste gefundene $Score_{HA}$ der eingesetzten Scoringfunktion, sowie der beste gefundene $Score_{HA}$ Wert der DSX-Scoringfunktion eines Fragments. Die beste Bindepose auf Grundlage der DSX-Scoringfunktion wurde für die MD-Simulation als Startpose herangezogen. Während für die ChemPLP-Posen nur die Pharmakophormodelle 1 bis 3 vertreten waren, fand sich das Pharmakophormodell vier relativ häufig in den ASP-Posen.

Die ausgewählten Fragmente wurden für 25 ns simuliert und anschließend anhand ihrer RMSD-Werte (basierend auf den Schweratomen des Liganden) in Bezug zur Startkonfiguration der Simulation analysiert. Abbildung 6.15 zeigt eine Übersicht aller simulierten Moleküle. Moleküle mit stabilen Bindeposen zeigen RMSD-Werte mit kleinen Spannweiten zwischen Minimal- und Maximal-Wert, sowie engen Spannweiten zwischen dem ersten und dritten Quartil. Eine zusätzliche Analyse der RMSD-Werte in Bezug auf ihre zeitliche Entwicklung während der Simulation ergab, dass von 29 Molekülen der ChemPLP-Posen zehn Moleküle dauerhaft die Bindestelle verließen, sieben eine fluktuierende Bindepose zeigten und zwölf eine stabile Pose einnahmen. Von diesen zwölf stabilen Posen behielten acht ihre anfängliche Startpose bei, während vier eine neue Bindungspose

einnahmen. Von den 20 Molekülen der ASP-Posen verließen fünf Moleküle dauerhaft die SHP-I Bindestelle, vier nahmen eine fluktuierende Bindepote ein, und elf Moleküle nahmen eine stabile Bindepote ein, von denen acht ihre Startpose behielten und drei einen neuen Bindemodus einnahmen.

Im nächsten Schritt wurden alle oben als stabil identifizierten Fragmente für weitere 50 ns simuliert. Abbildung 6.16 zeigt analog die RMSD-Werte der Moleküle nach diesen weiteren 50 ns Simulation. 4 Moleküle (2 ASP, 2 ChemPLP) verließen die Bindetasche innerhalb der Simulationszeit. Die restlichen Moleküle konnten in solche unterteilt werden, die weiterhin über die ganze Zeit eine definierte Bindepote hielten, und solche, die stark fluktuierende Bindepotes zeigten.

Tabelle 6.5: Ausgewählte Fragmente (ChemPLP-Scoringfunktion) für die MD-Simulationen (25 ns)

Die angegebenen Scores sind die jeweils besten für jedes Fragment sortiert nach dem $Score_{HA}^{ChemPLP}$

Insgesamt bester gefundener $Score_{HA}^{ChemPLP}$: 8.12

Insgesamt bester gefundener $Score_{HA}^{DSX}$: -8.05

ID	$Score_{HA}^{DSX}$	$Score_{HA}^{ChemPLP}$	Modell
Z1079127074	-7.68	6.93	Modell 3
Z803145266	-5.63	6.10	Modell 3
Z1702891974	-5.31	6.03	Modell 3
Z1350606470 (SHP12)	-4.78	5.92	Modell 3
Z2488948504	-5.46	5.87	Modell 3
Z1622680304	-5.98	5.67	Modell 1
Z2004547665	-5.05	5.53	Modell 1
Z1270399998	-5.41	5.35	Modell 4
Z1130340537	-5.07	5.19	Modell 1
Z1754517015	-5.32	5.05	Modell 1
Z2439799034	-5.18	5.02	Modell 2
Z1949076379F	-5.41	4.98	Modell 1
Z104639526 (SHP11)	-5.58	4.97	Modell 2
Z31231940 (SHP9)	-5.11	4.96	Modell 2
Z19650415	-5.02	4.85	Modell 2
Z1696822760	-3.8	4.80	Modell 4
Z1583117108	-5.56	4.77	Modell 1
Z54747399	-5.29	4.73	Modell 1
Z18499270	-4.53	4.71	Modell 2
Z1824025439	-5.10	4.70	Modell 1
Z2289798421	-4.46	4.64	Modell 2
Z1357823981	-4.15	4.63	Modell 1
Z1359569066	-4.15	4.63	Modell 1
Z54738025	-4.74	4.62	Modell 1
Z416877372	-5.01	4.59	Modell 1
Z19649968	-4.56	4.59	Modell 2
F5319-0013	-4.89	4.54	Modell 2
Z1832147660	-4.23	4.51	Modell 2
Z32368274 (SHP10)	-4.35	4.5	Modell 2

Tabelle 6.6: Ausgewählte Fragmente (ASP-Scoringfunktion) für die MD-Simulationen (25 ns)

Die angegebenen Scores sind die jeweils besten für jedes Fragment sortiert nach dem $Score_{HA}^{ASP}$

Insgesamt bester gefundener $Score_{HA}^{ASP}$: 2.55

Insgesamt bester gefundener $Score_{HA}^{DSX}$: -7.68

ID	$Score_{HA}^{DSX}$	$Score_{HA}^{ASP}$	Modell
Z2310677974	-5.61	-1.96	Modell 4
Z409946896	-5.30	-1.84	Modell 2
Z415728082 (SHP2)	-4.21	-1.83	Modell 3
Z1565851368 (SHP5)	-4.29	-1.81	Modell 1
Z1511494600	-6.60	-1.80	Modell 4
Z2513335412	-4.55	-1.79	Modell 1
Z1694842016 (SHP3)	-4.50	-1.78	Modell 2
F5890-0168	-4.27	-1.78	Modell 2
Z1381789635	-5.51	-1.74	Modell 4
Z1856567394 (SHP4)	-4.34	-1.73	Modell 1
Z1145207891 (SHP6)	-5.20	-1.72	Modell 1
Z1696822760	-5.15	-1.70	Modell 4
Z32368274 (SHP10)	-4.65	-1.66	Modell 4
Z1689632848	-4.67	-1.66	Modell 2
F5890-0092	-4.37	-1.64	Modell 2
Z446358664 (SHP1)	-4.96	-1.63	Modell 2
F5890-0143	-3.96	-1.62	Modell 2
Z1270399998	-5.90	-1.62	Modell 4
Z237932184	-4.64	-1.62	Modell 1
Z1381790187	-5.15	-1.58	Modell 2

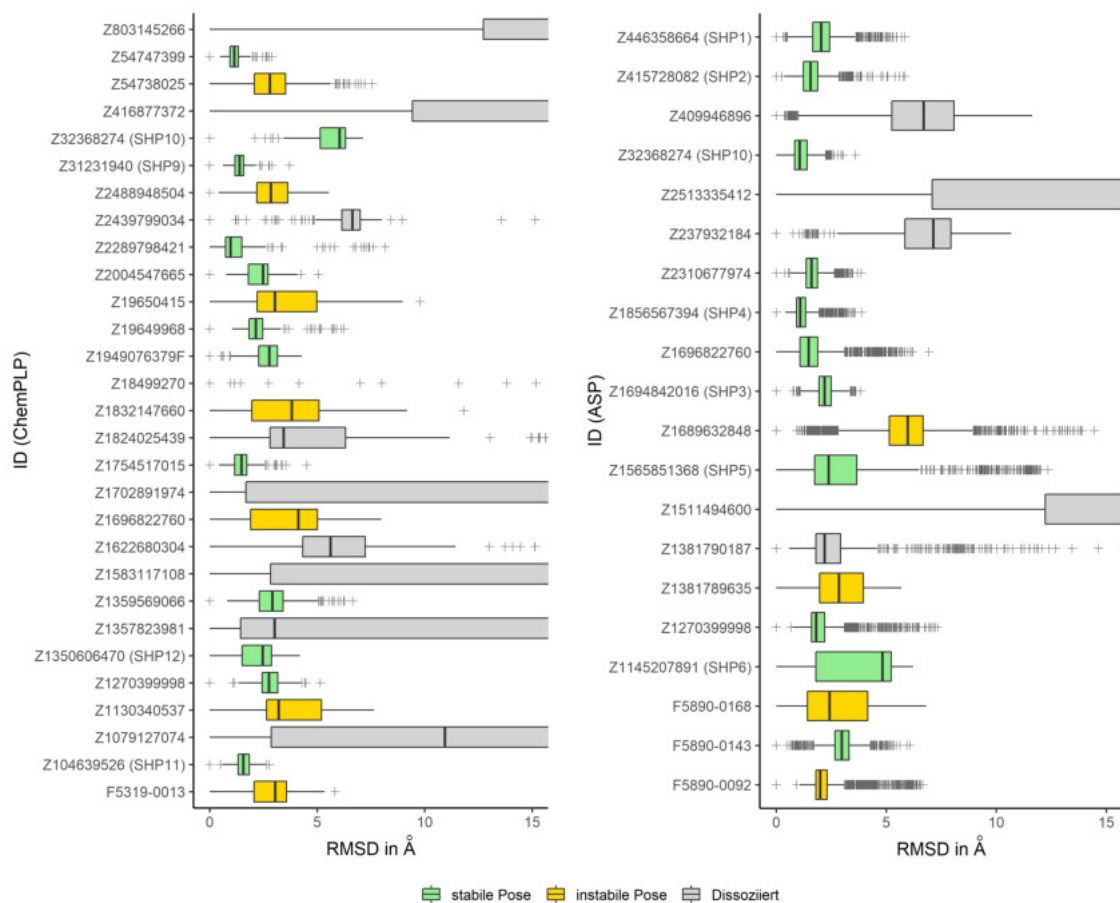


Abbildung 6.15: Ergebnisse der 25 ns MD-Simulationen

Boxplots der Liganden RMSD-Werte in Bezug zur Startkonfiguration der Simulation. Auf der linken Seite sind die ausgewählten Fragmente basierend auf den ChemPLP-Posen, auf der rechten basierend auf den ASP-Posen zu sehen. Die Anordnung der Fragmente folgt der ID-Nummer.

Grün: Moleküle die eine stabile Bindepese in der SHP-I Bindestelle zeigten und für eine weitere MD-Simulation von 50ns ausgewählt wurden.

Gelb: Moleküle die eine instabile Bindepese zeigten.

Grau: Moleküle, die die Bindestelle verließen.

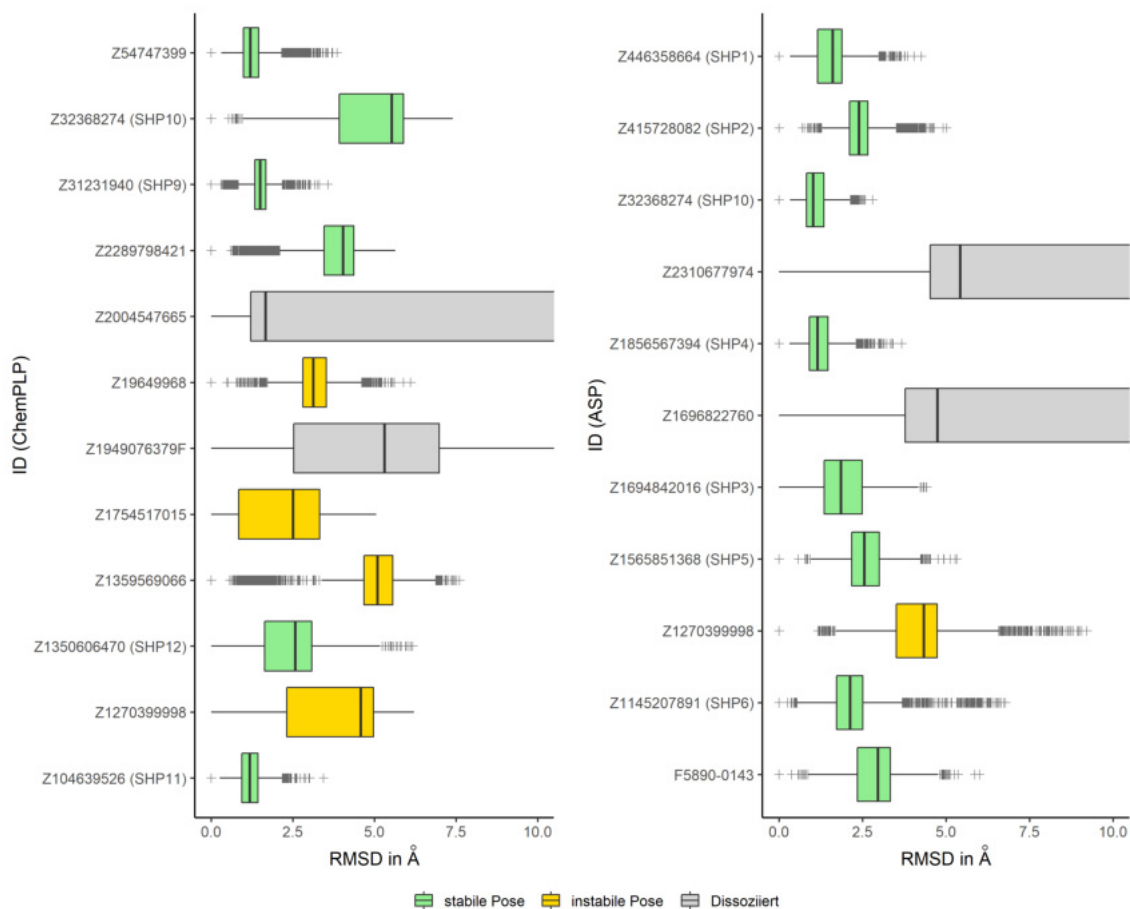


Abbildung 6.16: Ergebnisse der 50 ns MD-Simulationen ausgewählter Fragmente aus den 25 ns Simulationen

Boxplots der Liganden RMSD-Werte in Bezug zur Startkonfiguration der 50 ns Simulation. Auf der linken Seite sind die Fragmente basierend auf den ChemPLP-Posen, auf der rechten basierend auf den ASP-Posen zu sehen. Die Anordnung der Fragmente folgt der ID-Nummer.

Grün: Moleküle die eine stabile Bindepose in der SHP-I Bindestelle zeigten in die nähere Auswahl kamen.

Gelb: Moleküle die eine instabile Bindepose zeigten.

Grau: Moleküle, die die Bindestelle verließen.

6.1.3.4 Auswahl von 10 Fragmenten für experimentelle Untersuchungen

Basierend auf der Analyse der Trajektorien und der Verfügbarkeit der Fragmente bei ihren Anbietern, wurden die in Tabelle 6.7 gezeigten 10 Fragmente für experimentelle Untersuchungen ausgewählt. Die jeweiligen Docking-Posen, MD-Posen, sowie der zeitliche Verlauf der RMSD-Werte dieser Fragmente sind in Anlage B zu finden. Die Stabilität der Bindeposen innerhalb der 50 ns war bei allen Fragmenten gegeben, und die Trajektorien zeigten nur geringe Fluktuationen der Bindepose, mit Ausnahme von SHP10, das eine anfängliche Umorientierung der Pose innerhalb der 50 ns zeigte, und SHP12, das eine vergleichsweise hohe

Fluktuation aufwies. Beide wurden aufgrund ihrer als interessant erachteten Interaktionen dennoch mit ausgewählt. Tendenziell zeigten die ASP-Ergebnisse etwas geringere Fluktuationen in den RMSD-Werten als die ChemPLP-Ergebnisse. SHP10 war das einzige Fragment, das in den Ergebnissen beider Scoringfunktionen vertreten war, wobei die ASP-Pose die stabilere war. Alle ausgewählten Fragmente zeigten hydrophobe oder aromatische Systeme im hydrophoben Bereich der Bindestelle. Eine stabile, direkte oder über Wassermoleküle vermittelte indirekte Wasserstoffbrückenbindung zu T180 oder P178 konnte ebenfalls bei allen Fragmenten in den Simulationen vorgefunden werden. Die in den Kofaktor-Komplexen gefundene Wasserstoffbrücke zu I182 konnte in sechs Trajektorien beobachtet werden. In den später durchgeführten STD-NMR- und BLI-Experimenten (siehe Kapitel 6.3) konnten, SHP3 und SHP9 experimentell als Liganden bestätigt werden. Die MD-Ergebnisse dieser beiden Fragmente sollen daher hier im Detail gezeigt werden. Auf eine detailliertere Beschreibung der nicht bindenden Fragmente soll hier verzichtet werden, nicht zuletzt auch deshalb, weil kein unmittelbares Merkmal als Erklärung für das Nicht-Binden gefunden werden konnte.

Tabelle 6.7: Übersicht der für die experimentellen Tests ausgewählten Fragmente.

Alle Verbindung wurden von dem Anbieter Enamine bezogen.

ID Intern	ID Enamine	Modell	Scoringfunktion	DSX Score/HA	Score/HA
SHP1	Z446358664	Modell 2	ASP	-4.96	-1.63
SHP2	Z415728082	Modell 3	ASP	-4.21	-1.83
SHP3	Z1694842016	Modell 2	ASP	-4.45	-1.78
SHP4	Z1856567394	Modell 1	ASP	-4.34	-1.73
SHP5	Z1565851368	Modell 1	ASP	-4.29	-1.81
SHP6	Z1145207891	Modell 1	ASP	-5.20	-1.72
SHP9	Z31231940	Modell 2	ChemPLP	-5.11	4.96
SHP10	Z32368274	Modell 4	ASP	-4.65	-1.66
SHP11	Z104639526	Modell 2	ChemPLP	-5.58	4.97
SHP12	Z1350606470	Modell 3	ChemPLP	-4.78	5.92

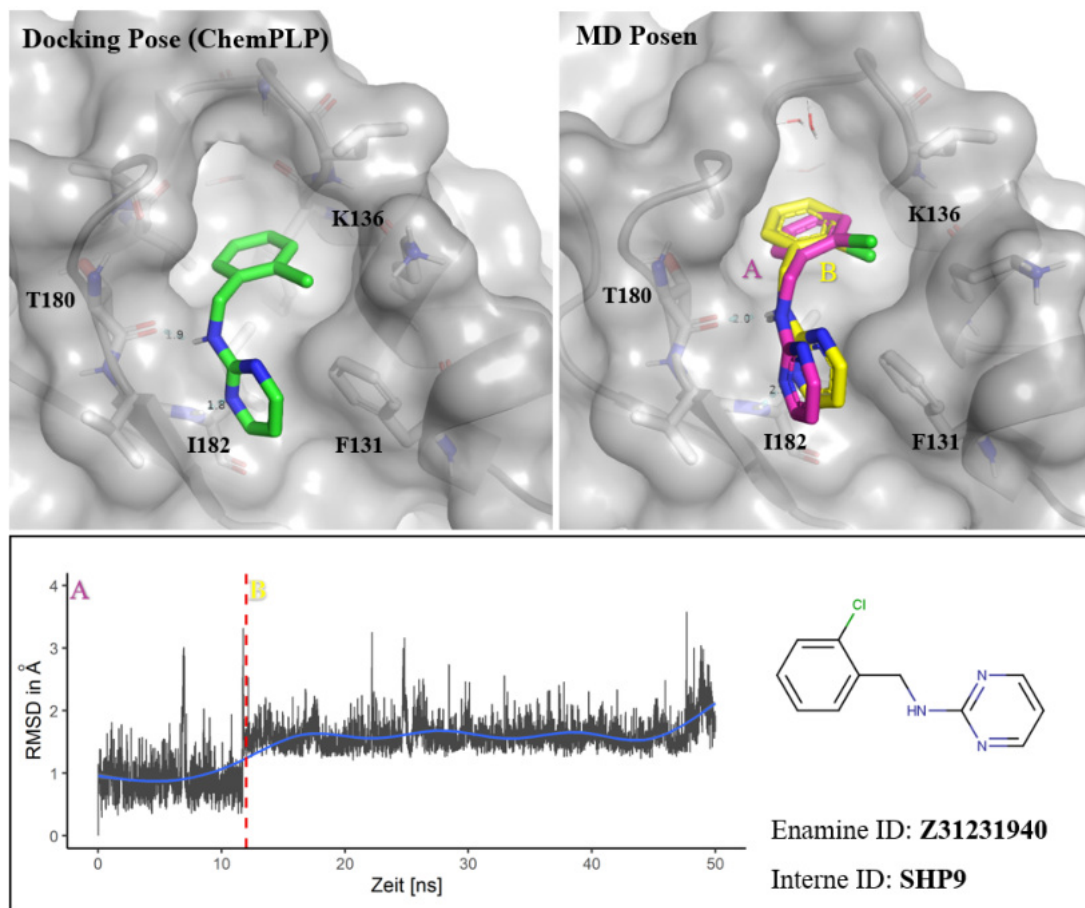


Abbildung 6.17: Docking und MD-Profil der als Ligand bestätigten Verbindung SHP9
Gezeigt ist die gefundene Dockingpose der verwendeten Scoringfunktion, die gefundenen MD-Posen der 50 ns Simulation, sowie die zeitliche Entwicklung des RMSD-Wertes während der Simulation. Kleine Zahlen zeigen die H-Brücken Abstände in Å (Schweratom zu H-Atom). Für detaillierte Erläuterungen siehe Text.

Das Molekül SHP9 zeigte in beiden MD-Simulationen (25 ns und 50 ns) eine ausgesprochen stabile Pose. Die in Abbildung 6.17 gezeigte Docking-Pose in Grün weicht nur geringfügig von den später in der MD-Simulation eingenommenen Posen ab. In der 25 ns Simulation wurde die gleiche, hier als Pose **A** bezeichnete Pose vorgefunden. Die zwei in der 50 ns MD-Simulation gefundenen Posen **A** und **B** weichen nur leicht voneinander ab. Hierbei bewegt sich der Pyrimidinring von Pose **A** zu Pose **B** so, dass dieser besser über π - π -Wechselwirkungen mit dem Phenylring von F131 interagieren kann. Der 2-Chlorphenylring liegt im hydrophoben Bereich der SHP-I Bindestelle, mit Orientierung des Chloratoms zum hydrophoben Teil der K136 Seitenkette. Die Amin-Funktion wechselwirkt über eine Wasserstoffbrückenbindung mit O^1 -T180, und ein Stickstoffatom des Pyrimidinringes fungiert als Wasserstoffbrückenakzeptor für N^2 -I182. Beide Interaktionen imitieren somit die in den UFD1 und Derlin-1 Komplexen gefundene Interaktion des Peptidrückgrates.

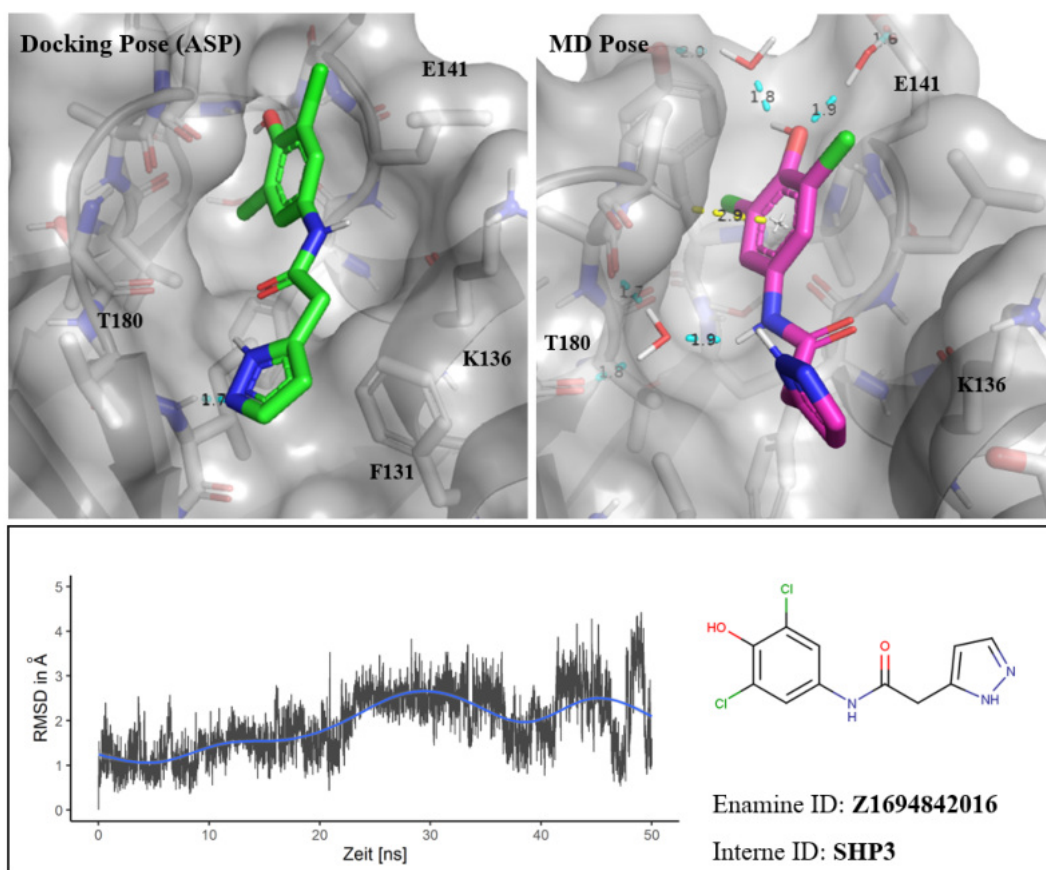


Abbildung 6.18: Docking und MD-Profil der als Ligand bestätigten Verbindung SHP3
Gezeigt ist die gefundene Dockingpose der verwendeten Scoringfunktion, die gefundene MD-Pose der 50 ns Simulation, sowie die zeitliche Entwicklung des RMSD-Wertes während der Simulation. Kleine Zahlen zeigen die H-Brücken Abstände in Å (Schweratom zu H-Atom). Für detaillierte Erläuterungen siehe Text.

In den MD-Simulationen von SHP3 konnte ein stabiles Wassermolekül gefunden werden, das, wie in der Komplexstruktur von p97-N und UFD1 überbrückende Wasserstoffbrückenbindungen zu O^1 -T180 und O^1 -P178 zeigte. Daneben fanden sich kurzzeitig auch direkte Wasserstoffbrücken zwischen der Amid-Bindung und O^1 -T180. Die Amidbindung orientierte sich innerhalb der ersten Nanosekunden der 25 ns MD-Simulation zu sehende neue Pose. Die in der Docking-Pose dargestellte Interaktion zwischen dem Pyrazolring und I182 war in der MD-Simulation nicht stabil. Es wurde nur die gezeigte tautomere Form des Pyrazolrings simuliert. Die Fluktuationen gegen Ende der Trajektorie sind auf den nicht stabilen Bindemodus des Pyrazolrings zurückzuführen. Der Dichlorphenolring zeigt in die Kavität der SHP-I Bindestelle hinein. Der Protonierungszustand der Hydroxylgruppe wurde in den später durchgeführten NMR-Messungen als deprotoniert vorgefunden. Die hier gezeigte protonierte Form interagiert mit den in der Bindetasche vorhandenen Wassermolekülen.

6.1.3.5 Zusammenfassung und Diskussion des virtuellen Screenings an der SHP-I Bindestelle in der p97-N Subdomäne

Anhand der strukturellen Analysen der SHP-Bindestellen in den p97-N Komplexstrukturen mit den Kofaktoren UFD1 und Derlin-1, sowie der Analysen der apo-Strukturen von p97-N durch die Interaktionspotentiale in MOE und den FTMap-Algorithmus, konnten mögliche Interaktionspunkte vor allem in der SHP-I Bindestelle identifiziert werden. Die in mehreren Kristallstrukturen gefundenen Wassermoleküle innerhalb der SHP-I Bindestelle konnten in einer MD-Simulation und anschließenden Analyse durch die GIST-Methode bestätigt werden. Sie legen den Schluss nahe, dass innerhalb der SHP-I Bindestelle ein Wassernetzwerk ausgebildet wird und dieses möglicherweise einen Einfluss auf die konformationelle Stabilität der Bindestelle hat. Basierend auf diesen Ergebnissen wird die These eines Einflusses der dort gefundenen Wassermoleküle auf die Bindung kleinen Moleküle aufgestellt.

Anhand dieser Ergebnisse wurden vier unterschiedliche Pharmakophormodelle entwickelt. Diese vier Modelle bildeten die Grundlage des Pharmakophor-Screenings, der Selektionen im Docking und der MD-Simulationen. Letztendlich wurden zehn Fragmente anhand ihrer RMSD-Werte und Interaktionen für die spätere experimentelle Untersuchung ausgewählt. Die ausgewählten Moleküle folgten hierbei den Interaktionen der jeweiligen Pharmakophormodelle.

Ein großes Problem des Screenings war das Fehlen einer kristallographischen Komplexstruktur eines niedermolekularen Verbindungen in diesem Bereich, mit dem die verwendeten Scoringfunktionen durch ein Redocking validiert hätten werden können. Des Weiteren stellt die SHP-Bindestelle durch ihre starke (für PPI typische) Exponiertheit mit nur kleinen Kavitäten (SHP-I, SHP-II) eine große Herausforderung für das Design eines Inhibitors dar.

Die Selektion der Docking Ergebnisse anhand der zugrunde gelegten Pharmakophormodelle stellt hier nur einen möglichen Weg dar, bei dem nicht ausgeschlossen werden kann, dass durch eine andere Wahl an Selektionskriterien möglicherweise andere Strukturen hätten gefunden werden können. Auch stellt sich die Frage, ob das hier durchgeführte DSX-Rescoring nach dem zweiten Dockinglauf zur alleinigen Auswahl der Posen wirklich zielführend war, oder ob ein eher Consensus-basierter Ansatz in der Form, dass nur diejenigen Posen selektiert werden, die in beiden Scoringfunktionen hohe Werte erzielen, ein zusätzlicher Ansatz gewesen wäre (siehe hierzu Anhang C).

Alle Fragmente wurden in den MD-Simulationen nur einmal simuliert. Eine größere Aussagekraft über die Stabilität der Fragmente wäre sicherlich durch die Simulation mehrerer Replikate gegeben gewesen ¹

Wie bei *Giordanetto et al.*[283] beschrieben, zeigt eine Analyse von 489 bekannten Kristallstrukturen von Fragment-Protein-Komplexen, dass bei 73% der Fragmente mehr als 80% ihrer lösemittelzugänglichen Fläche in Kontakt mit dem Protein stehen und die Fragmente demnach als sehr im Protein „vergraben“ bezeichnet werden können. Eine solche Analyse der „Vergrabenheit“ der Fragmente wäre sicherlich ein interessantes zusätzliches Selektionskriterium im virtuellen Screening gewesen. Alle hier verwendeten Scoringfunktionen (empirisch wie auch wissensbasiert) beruhen auf experimentellen bzw. strukturellen Daten von Protein-Ligand-Komplexen. Diese Trainingsdatensätze enthalten aber vor allem größere, wirkstoffähnlichere Moleküle. Auch finden sich in den Trainingsdaten überwiegend Liganden, die die katalytischen Zentren von Proteinen adressieren. Diese Umgebungen unterscheiden sich deutlich von den eher flachen Exponierten Stellen, an denen PPI stattfinden [284, 285]. Daher stellt sich die Frage, inwiefern diese Scoringfunktionen mögliche Bindepasen von Fragmenten, insbesondere im Kontext von PPI-Interaktionen, adäquat vorhersagen können [286]. Auch wenn die Scoringfunktion ChemPLP bereits für das Scoring von Fragmenten Verwendung fand [287, 288], könnten hier Ansätze, wie kraftfeldbasierte Scoringfunktionen eine bessere Beschreibung der Interaktionen liefern. So wäre eine mögliche Strategie, die Fragmente kurzen MD-Simulationen zu unterziehen und die Affinitäten über die *Molecular Mechanics Generalized Born Surface Area* (MM-GBSA)-Methode abzuschätzen² [289]. Als letzter Punkt sei genannt, dass bei dem hier durchgeführten virtuellen Screening nur mit *einer* Konformation der SHP-I Bindetasche gearbeitet wurde. Interessant wäre sicherlich, im Kontext eines Ensemble Docking mit mehreren Konformationen der SHP-I Bindestelle zu arbeiten [290].

Das durchgeführte virtuelle Screening ist demnach als eine erste „Pionier“-Arbeit zu sehen, und neue experimentelle Daten im Rahmen dieses Projektes werden sicherlich neue Strategien aufzeigen und weitere Möglichkeiten eröffnen.

¹Das im Rahmen dieses Projektes durchgeführte virtuelle Screening zur Identifikation von möglichen kovalent bindenden Fragmenten griff diese Punkte auf. In dem dort durchgeführten Docking wurde ein Konsensus-basierter Ansatz aus drei verschiedenen Scoringfunktionen gewählt und alle anschließenden MD-Simulationen als Triplikate durchgeführt (Masterarbeit von Stephan Böhler [280].)

²Dieser Ansatz wurde ebenfalls in das neue virtuelle Screening für kovalente Fragmente integriert. (Masterarbeit von Stephan Böhler citeBoehler.)

6.2 Biophysikalische Messungen an p97

In diesem Kapitel sollen die Ergebnisse der biophysikalischen Messungen an p97 vorgestellt werden, die im Rahmen dieser Arbeit durchgeführt wurden. Hierbei wird zunächst auf die BLI-Messungen und STD-NMR-Messungen des virtuellen und biophysikalischen Fragmentscreenings eingegangen, bevor diese zusammengefasst diskutiert werden.

6.2.1 Validierung des Biolayer-Interferometrie Assays

Die durchgeführten BLI-Messungen stellen einen Hauptteil dieser Arbeit dar. Daher soll hier im Detail auf die Ergebnisse der Etablierung und Validierung des BLI-Assays eingegangen werden. Eine geeignete Positivkontrolle stellt hier (da es sich bei p97 um eine ATPase handelt) Adenosindiphosphat (ADP) dar, dessen Bindung an das p97-ND1-Konstrukt zur Etablierung des Assays genutzt wurde. Die in der Literatur berichtete Affinität von ADP zu der Nukleotidbindestelle der ND1-Domäne von p97 ist mit K_D -Werten von 200 nM [61] (ND1^{aa 1-458}), 203 nM (ND1^{aa 1-458}), 98.3 nM (ND1^{aa 1-480}) und 123 nM (p97^{FL}) [291] angegeben. Alle diese K_D Werte wurden durch SPR-Spektroskopie bestimmt und die Proteine mittels eines Avi-Tag am C-Terminus biotinyliert und am Sensor immobilisiert. Da es sich bei den hier durchgeführten BLI-Messungen und dem gewählten Versuchsaufbau um eine gleichwertige Methode handelt, sollte sich der K_D -Wert für ADP für das hier eingesetzte Konstrukt entsprechend reproduzieren lassen.

Abbildung 6.19 zeigt zwei Sensorgramme der durchgeführten ADP-Validierungsmessungen. Die Kurven weisen das für SPR- oder BLI-Messungen typische, an eine „Haifischflosse“ erinnernde Profil auf. Zu sehen sind deutliche exponentielle Anstiege in der Assoziationsphase, das Erreichen von Plateaus, sowie exponentiell abfallende Kurven in der Dissoziationsphase [292]. Die Plateauphasen zeigen das Erreichen eines chemischen Gleichgewichts an und lassen eine Auswertung mit Hilfe des Langmuir-Modells zu.

Auffallend ist für die höchsten ADP Konzentrationen (100 μ M), dass die Verschiebung³ nicht wieder den Wert der Basallinie am Ende der Dissoziation erreicht, sondern parallel zu dieser verläuft. Dieser Sachverhalt lässt sich auch schwächer bei den geringeren Konzentrationen beobachten. Der Grund lag hier

³Im Zusammenhang mit den durchgeführten BLI-Messungen, sei hier nochmal darauf hingewiesen, dass es sich bei den erhaltenen Messsignalen um eine Verschiebung λ der Minima und Maxima im elektromagnetischen Spektrum handelt. Daher werden alle Messsignale bei BLI-Messungen auch als Verschiebungen bezeichnet.

vermutlich bei der Verschleppung von ADP zwischen dem Assoziations-Napf und dem Dissoziations-Napf, wodurch das gebundene ADP nachdem dissoziieren wieder an p97 binden konnte. Dieser Sachverhalt führte bei der kinetischen Auswertung zu Problemen, weswegen hier auf die beiden höchsten Konzentrationen verzichtet wurde. Der für die Validierung errechnete Z-Faktor betrug 0.91 und zeigte, dass dieser Assay hervorragend zwischen einer positiven und einer negativen Probe unterscheiden konnte.

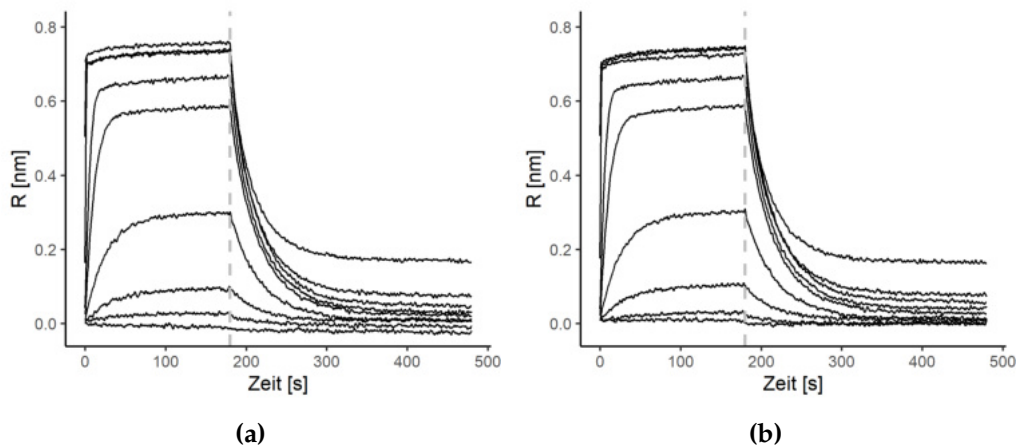
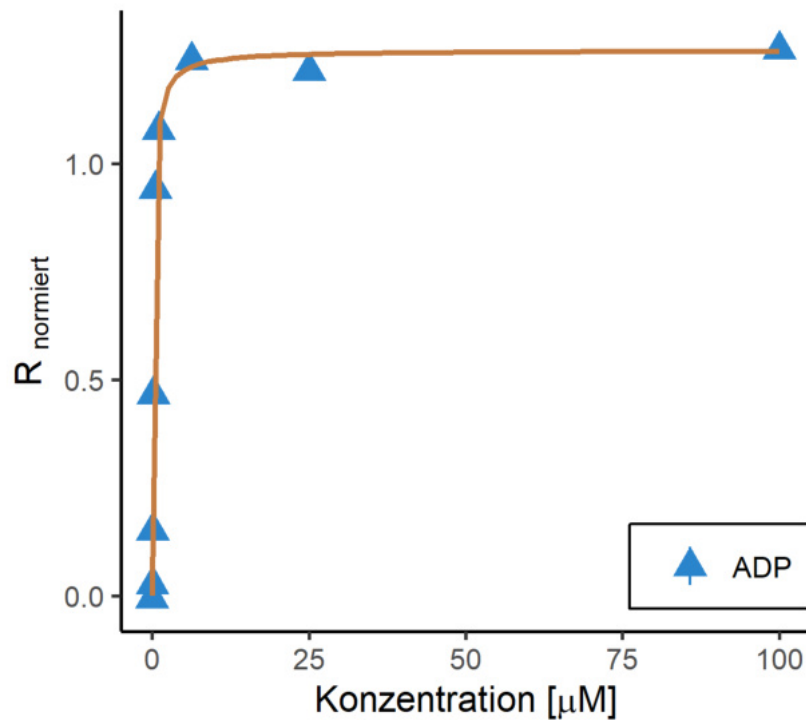


Abbildung 6.19: Sensorgramme für ADP

Zwei Sensorgramme (a und b) der sechs durchgeführten Validierungsmessungen für unterschiedliche ADP-Konzentrationen.

6.2.1.1 Auswertung mittels des Langmuir-Modells

Abbildung 6.20 zeigt die Gleichgewichtsauswertung der Messergebnisse. Um die Messungen von leicht unterschiedlich mit Protein beladenen Sensoren miteinander vergleichen und auswerten zu können, wurden alle Messwerte mit Hilfe der gemessenen Verschiebung einer 500 nM ADP Lösung normiert. Die Messwerte der sechs Messungen zeigen eine sehr geringe Streuung und lassen sich mit hoher Korrelation durch ein 1:1-Langmuir-Modell beschreiben. Der so erhaltene K_D von 200.6 ± 14.8 nM (Fehler des Regressionsmodells, $\Delta_{Reg.}$) ist mit den in der Literatur berichteten Werten praktisch identisch.



K_D	$\Delta_{Reg.}$	R_{max}	R^2
[nM]			
200.6	14.8	1.279	0.996

Abbildung 6.20: Langmuir-Modell

1:1-Langmuir-Modell basierend auf allen sechs Messungen. Die Daten wurden mittels der gemessenen Verschiebung einer 500 nM ADP Probe normiert.

6.2.1.2 Kinetische Auswertung

Abbildung 6.21 zeigt die Überlagerung eines 1:1-Bindungsmodells (basierend auf einer globalen exponentiellen Regression) mit den realen Messwerten für die Assoziation bzw. die Dissoziation der Messungen mit einem Sensor. Das ermittelte Modell beschreibt in großer Übereinstimmung die erhaltenen Daten. In den residualen-Graphen zeigen sich nur leichte Abweichungen, die im Vergleich zum Messsignal vernachlässigbar sind⁴. Eine Zusammenfassung aller so erhaltenen kinetischen Daten der Einzelmessungen ist in Tabelle 6.8 gegeben. Alle erhaltenen kinetischen Werte (k_a, k_d) zeigen bei allen sechs Messungen nur eine geringe Streuung. Der so ermittelte K_D -Wert liegt mit 238.6 nM nur geringfügig über dem Literaturwert.

⁴Der Maximalwert der Abweichung betrug 6.7%. Der von Sartorius für eine sehr gute Messung angegebene Grenzwert liegt bei 10%.

Zur Überprüfung wurden die erhaltenen Daten über den bei einem 1:1-Modell vorliegenden linearen Zusammenhang zwischen der Konzentration $[L]$ des Liganden und der empirische Geschwindigkeitskonstante k_{obs} ausgewertet. Für jede Konzentration der sechs Messungen wurde die empirische Geschwindigkeitskonstante k_{obs} durch eine exponentielle Regression der Assoziationsphase bestimmt und gegen die ADP Konzentration aufgetragen. Wie in Abbildung 6.22 zu sehen, zeigen die Daten einen eindeutigen linearen Zusammenhang zwischen den empirischen Geschwindigkeitskonstanten und den ADP Konzentrationen. Durch lineare Regression sind somit die Geschwindigkeitskonstanten k_d (als Schnittpunkt der Geraden mit der y-Achse) und k_a (als Steigung der Geraden) zugänglich.

Tabelle 6.8: Kinetische Auswertung der ADP-Messungen

Nr.	k_d		R^2 Ass.	k_a		R^2 Diss.	K_D	$\Delta_{Reg.}$
	$[s^{-1}]$	$[L \mu\text{mol s}^{-1}]$		$[\mu\text{mol L}^{-1}]$				
1.	0.0287	$2.77 \cdot 10^{-5}$	0.9994	$1.26 \cdot 10^{-4}$	$5.23 \cdot 10^{-7}$	0.9985	227.8	1.2
2.	0.0283	$2.44 \cdot 10^{-5}$	0.9994	$1.24 \cdot 10^{-4}$	$5.41 \cdot 10^{-6}$	0.9988	227.7	1.0
3.	0.0285	$2.56 \cdot 10^{-5}$	0.9993	$1.20 \cdot 10^{-4}$	$5.78 \cdot 10^{-7}$	0.9988	236.8	1.4
4.	0.0296	$2.59 \cdot 10^{-5}$	0.9991	$1.18 \cdot 10^{-4}$	$6.38 \cdot 10^{-7}$	0.9988	250.9	1.6
5.	0.0287	$2.88 \cdot 10^{-5}$	0.9993	$1.20 \cdot 10^{-4}$	$5.69 \cdot 10^{-7}$	0.9984	239.7	1.4
6.	0.0295	$2.69 \cdot 10^{-5}$	0.9992	$1.19 \cdot 10^{-4}$	$6.06 \cdot 10^{-7}$	0.9987	248.9	1.5
	\bar{k}_d	σ		\bar{k}_a	σ		\bar{K}_D	σ
	0.0289	$4.86 \cdot 10^{-4}$		$1.21 \cdot 10^{-4}$	$3.0 \cdot 10^{-6}$		238.6	9.1

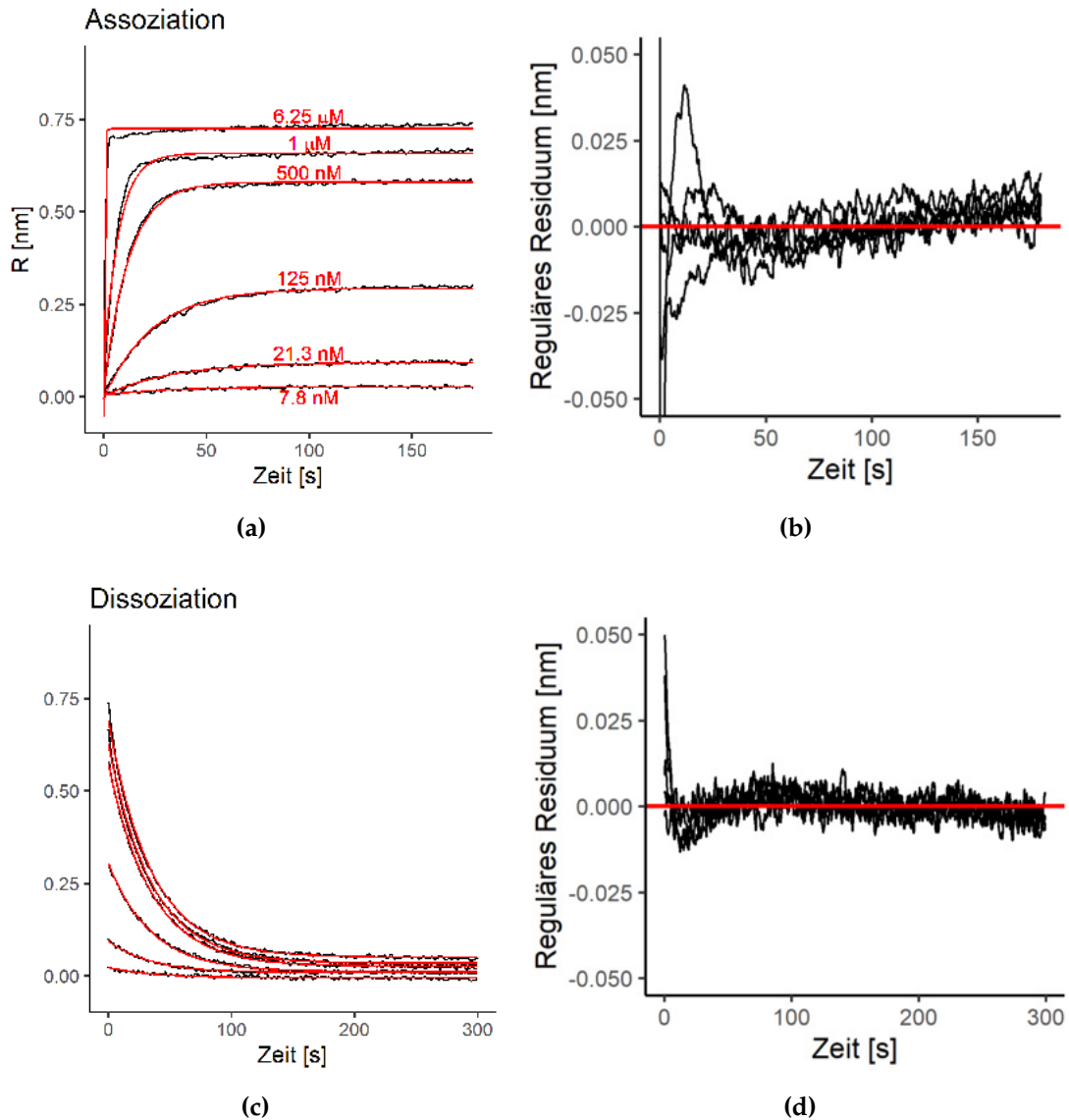


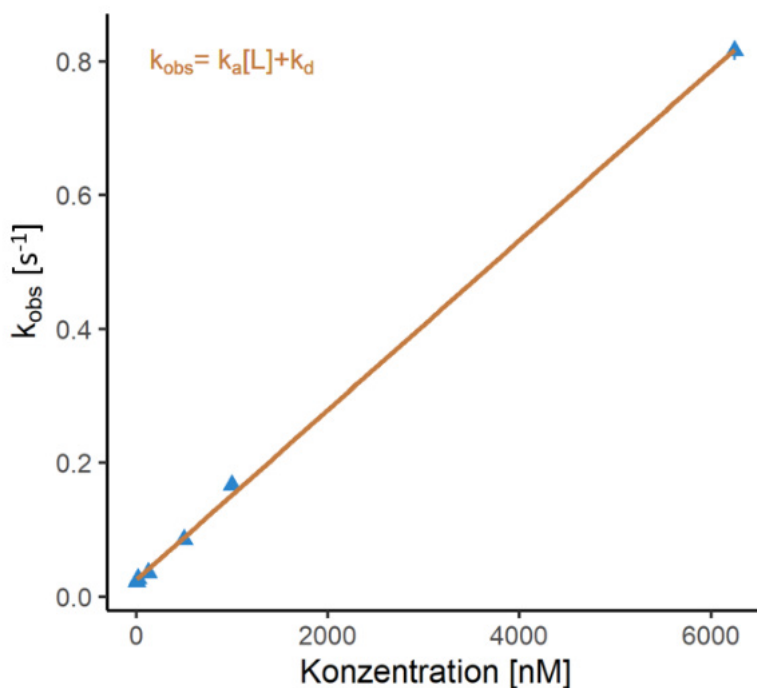
Abbildung 6.21: Ergebnisse der kinetischen Auswertung unter Annahme eines 1:1-Bindungsmodells

(a) 1:1-Bindungsmodell (rot), basierend auf einer globalen exponentiellen Regression der Assoziationsphase für einen Sensor (6 unterschiedliche Konzentrationen).

(b) Reguläre Residuen der globalen exponentiellen Regressionen für das oben gezeigte Modell der Assoziationsphase.

(c) 1:1-Bindungsmodell (rot), basierend auf einer globalen exponentiellen Regression der Dissoziationsphase für einen Sensor (6 unterschiedliche Konzentrationen).

(d) Reguläre Residuen der globalen exponentiellen Regressionen für das oben gezeigte Modell der Dissoziationsphase.



k_d	$\Delta_{Reg.}$	k_a	$\Delta_{Reg.}$	K_D	$\Delta_{Reg.}$
	$[s^{-1}]$	$[L \mu mol s^{-1}]$		$[nmol L^{-1}]$	
0.0249	0.0018	$1.27 \cdot 10^{-4}$	$6.92 \cdot 10^{-7}$	197.0	15.2

Abbildung 6.22: Kinetische Auswertung mittels k_{obs}

Kinetische Auswertung auf Basis der empirischen Geschwindigkeitskonstanten und den dazugehörigen Konzentrationen. Zu sehen ist ein eindeutiger linearer Zusammenhang der für eine 1:1-Bindung von ADP an p97-ND1 spricht.

6.2.1.3 Diskussion der Validierung

Eine Gegenüberstellung der erhaltenen Affinitäten der unterschiedlichen Auswertungen zeigt, dass sich die bekannten Affinitäten mit dem BLI-Assay reproduziert lassen, und die hier durchgeführten Messungen eine hohe Präzision zeigen. Auf Grundlage dieser Ergebnisse wurde die Messung einer $0.5 \mu M$ ADP Probe als Positivkontrolle für alle weiteren Messungen mit p97-ND1 gewählt. Das gewählte 1:1-Modell liefert eine adäquate Beschreibung der Bindung von ADP an p97.

Auffallend sind die recht hohen Verschiebungen R , von bis zu 0.8 nm. Das in SPR Experimenten häufig zugrunde gelegte lineare Modell zur Ermittlung der maximalen beobachtbaren Winkeländerung, ist in Gleichung 6.1 gegeben [61][293]:

$$R_{max} = \frac{R_{Beladung\ Sensor}}{MW_{Protein}} MW_{Ligand} \quad (6.1)$$

Legt man dieses Modell für BLI-Messungen zugrunde, wie es bei Sartorius [274] für die Bindung kleiner Moleküle an die Carbanhydrase II berichtet wird, so würde man bei einer mittleren Proteinbeladung von 5 nm und einem Molekulargewicht (MW) von 427 Da für ADP bzw. 55.2 kDa (Monomer) für die ND1-Domäne, eine Verschiebung von 0.039 nm bei 100% Besetzung aller Bindestellen erwarten; bei Annahme einer Stöchiometrie von $n = 6$ für das Hexamer somit 0.234 nm.

Die Verschiebung von ADP bei Sättigung liegt mit rund 0.8 nm, etwa 3.5 mal höher als erwartet. Die Schichtdickenänderung bzw. Dichteänderung der Sensoroberfläche lässt sich demnach nicht ausschließlich mit der Bindung des ADP über dessen Molekulargewicht erklären. Des weiteren lässt sich die Übertragung des SPR-Modells, das einen linearen Zusammenhang zwischen der maximalen Winkeländerung und der Masse eines Liganden besitzt, nur bedingt auf die BLI-Methode übertragen (wie bei [274] berichtet). Hier kommt es, aufgrund von zusätzlichen Protein- und/oder Ligandeneinflüssen, häufig zu nichtlinearen Zusammenhängen, die im Detail nicht komplett verstanden sind⁵.

Da p97 eine ADP-, sowie eine ATP-Konformation einnehmen kann, wäre eine mögliche Erklärung, dass das Messsignal zusätzlich von der Konformationsänderung innerhalb der ND1-Domäne abhängt. In der Literatur findet man gegenwärtig keine breiten Informationen zu gemessenen Konformationsänderungen mithilfe der BLI-Methode [294]. Um zu beweisen, dass die Verschiebung eine Konformationsabhängigkeit besitzt, wurde ein Versuch mit ADP und dem ATP Analogon AMP-PNP durchgeführt. Die Überlegung war hierbei, dass, wenn die Verschiebung nur von der Masse des bindenden Liganden abhängig wäre, beim Wechsel aus einer 1 mM ADP (427 Da) Lösung in eine 1 mM AMP-PNP (506 Da) Lösung keine Verschiebung zu sehen sein sollte. Die Massendifferenz beider Liganden läge hierbei unterhalb des Detektionslimits (> 160 Da) des Gerätes. Abbildung 6.23 zeigt das Ergebnis eines solchen Experiments. Zu sehen sind die Rohdaten der Messung. Der Sensor startet in Assaypuffer und taucht bei Blau in eine 1 mM ADP-Lösung. Wie zu erwarten steigt die Verschiebung massiv an (> 0.7 nm). Bei Rot taucht der Sensor in eine 1mM AMP-PNP-Lösung ein, was

⁵Persönliche Korrespondenz mit Dr. Hans Wünsche, Sartorius

zu einer negativen Verschiebung führt. Dies zeigt, dass das Messsignal von ADP nicht alleine vom Einfluss der Masse des Liganden, sondern auch von der Konformation von p97 abzuhängen scheint. Ein negatives Signal wurde bei *Martin et al.* [273] ebenfalls mit der Konformationsänderung einer Schleifenstruktur bei Bindung eines Liganden an eine Kinase in Verbindung gebracht. Inwiefern hier tatsächlich die Bewegung der N-Domäne beobachtet wird, oder ob es sich nur um die Konformationsänderung von Aminosäuren innerhalb der Nukleotid-Bindestelle handelt, kann nicht beurteilt werden. Auf der Grundlage der guten Reproduzierbarkeit der in der Literatur berichteten Affinität von ADP an die p97-ND1 Domäne, wurde eine 500 nM Lösung als Positivkontrolle bei allen p97-ND1 BLI-Messungen eingesetzt.

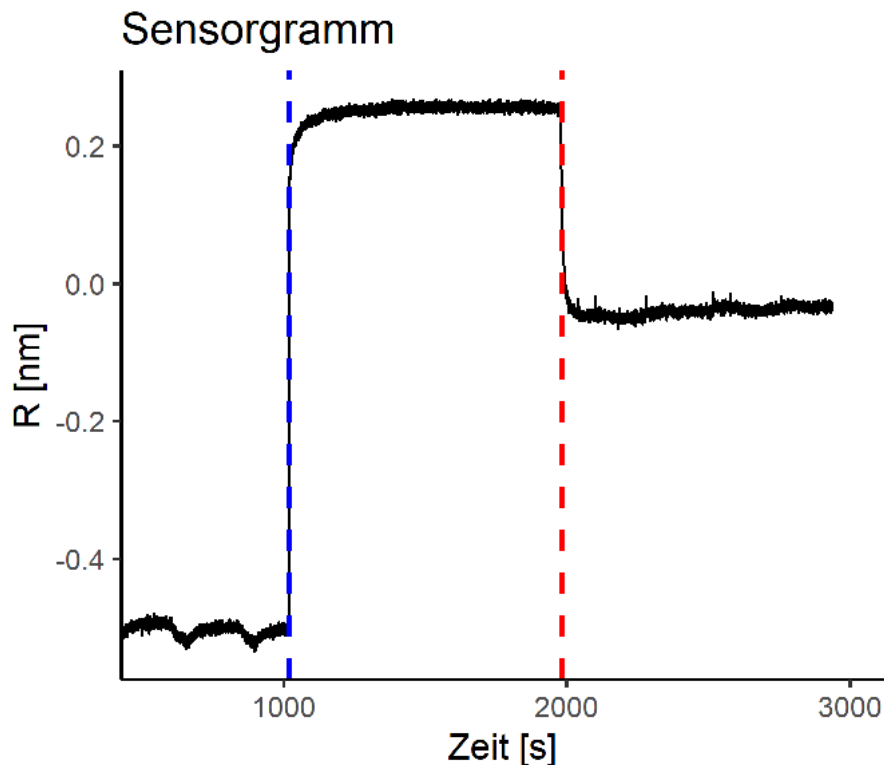


Abbildung 6.23: Abhängigkeit des Messsignals von der Konformation von p97

Rohverschiebungen der Messungen mit ADP und AMP-PNP. Von einer Nukleotid-freien Lösung taucht der Sensor bei blau in einen Napf mit einer 1 mM ADP Lösung. Bei rot taucht er in einen Napf mit einer 1 mM AMP-PNP Lösung. Wäre das Messsignal nur von der Masse des Liganden abhängig, sollte hier keine Signaländerung zu sehen sein. Das negative Signal deutet daher darauf hin, dass die beobachtete Verschiebung u.a. von Konformationsänderungen in p97 beeinflusst wird.

6.2.2 Der p97-Inhibitor Xanthohumol

Bisher wurde in der Literatur nur der Naturstoff Xanthohumol (XN) als p97-Inhibitor mit einer Bindestelle in der N-Domäne beschrieben. Die bei *Sasazawa et al.* [69] durchgeführten Experimente ließen allerdings keine Aussage über die Affinität von XN zu. Die hier mit XN durchgeführten BLI-Messungen dienen daher der Beantwortung der Folgenden zwei Fragestellungen: Der Bestätigung der Bindung von XN an p97, insbesondere an die N-Domäne, sowie der Messung der Affinitäten von XN für die N- und ND1-Domäne.

6.2.2.1 Messungen an der isolierten N-Domäne

Insgesamt wurden 6 Messungen mit einer 1:1-Verdünnungsreihe von 20 bis 0.625 μM mit drei Sensoren durchgeführt (2 Messungen pro Sensor). Aufgrund der schlechten Wasserlöslichkeit von XN sind Messungen im höheren Konzentrationsbereich ($>50 \mu\text{M}$) nicht möglich. In Bereichen von 50 bis 200 μM zeigten sich nicht auswertbare Sensorgramme, durch Aggregatbildung und ein sichtbares Ausfallen von XN.

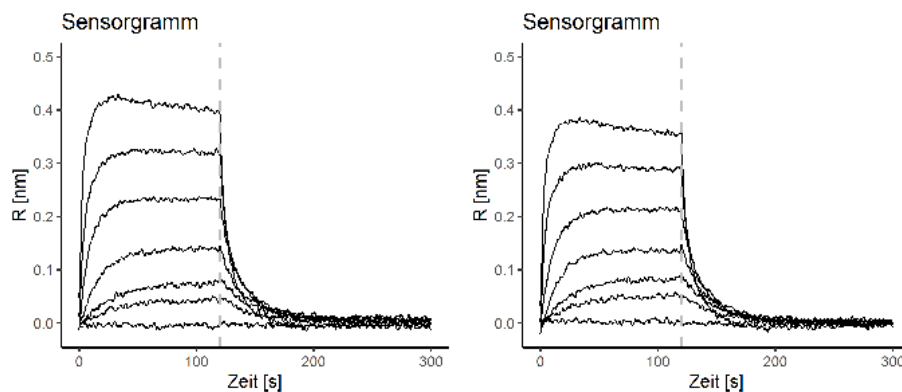


Abbildung 6.24: Sensorgramme zweier XN-Messungen

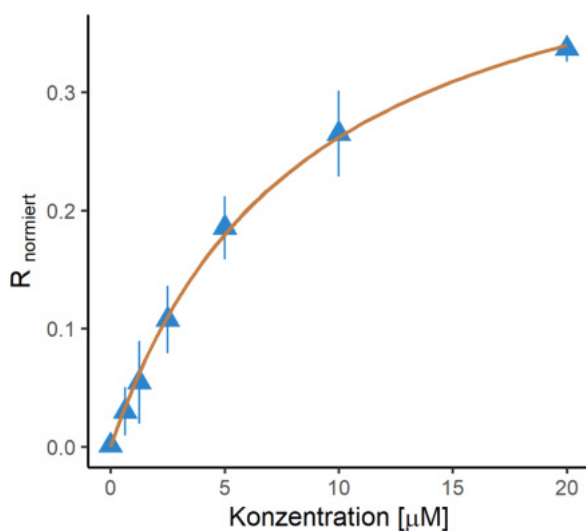
Gemessene Verschiebungen für sechs unterschiedliche XN-Konzentrationen zweier verschiedener Sensoren.

Abbildung 6.24 zeigt beispielhaft zwei Sensorgramme für zwei unterschiedliche Sensoren (sämtliche Sensorgramme der Messungen sind in Anlage D zu finden). Die Kurven zeigen (ähnlich wie bei ADP), exponentielle Anstiege in der Assoziation, deutliche Plateauphasen und exponentiell abfallenden Kurven in der Dissoziation, mit Wiedererreichen der Basislinie. Das in XN enthaltene Michael-System bietet Anlass über einen kovalenten Mechanismus der Bindung zu spekulieren. Aufgrund der vollständigen Dissoziation bei allen Konzentrationen, ist jedoch

von einer nicht-kovalenten Bindung auszugehen. Von Stephan Böhler durchgeführte massenspektrometrische Untersuchungen zeigten ebenfalls keine Modifikation der Cysteine in der N-Domäne von p97 (nicht veröffentlichte Daten). Auf Grundlage dieser Ergebnisse wird deshalb davon ausgegangen, dass die Bindung von XN nicht auf einem kovalenten Mechanismus beruht.

6.2.2.2 Auswertung des Gleichgewichtszustandes mittels des Langmuir-Modells

Die eindeutigen Plateauphasen in den Sensorgrammen ließen eine Auswertung durch ein Langmuir-Modell zu. Das in Abbildung 6.25 gezeigte auf allen sechs Messungen basierende Modell, lieferte eine sehr gute Korrelation mit einem R^2 Wert von 0.99 und einer Affinität von rund $8 \mu\text{M}$. Zu beachten ist hier allerdings, dass die gezeigten Sensorgramme schon bei $20 \mu\text{M}$ eine leicht negative Drift während der Plateauphase zeigten. Diese negativen Drifts könnten auf eine beginnende Aggregatbildung von XN hindeuten, mit dem Resultat, dass die hier ermittelte Affinität zu positiv geschätzt wird und der K_D -Wert höher liegen könnte.



K_D	$\Delta_{Reg.}$	R_{max}	R^2
$[\mu\text{mol L}^{-1}]$			
8.3	0.5	0.477	0.999

Abbildung 6.25: Langmuir-Modell für die Auswertung der XN-Messungen

1:1-Langmuir-Modell auf Grundlage aller sechs Messungen. Die Daten der unterschiedlichen Sensoren wurden mittels des Beladungssignals gegeneinander normiert.

6.2.2.3 Kinetische Auswertung der Messungen

Für die kinetische Auswertung wurden die Daten durch ein 1:1-Modell simuliert. Abbildung 6.26 zeigt die Modelle in Überlagerung mit den Messdaten für eine Messung. Das 1:1-Modell beschreibt die Assoziations-, wie auch die Dissoziationsphasen in guter Übereinstimmung. Nur für kleinere Konzentrationen folgte der An- bzw. Abstieg der Kurven nicht perfekt einem 1:1-Modell. Die Abweichungen in den Residualgraphen liegen nur für die beiden kleinsten Konzentrationen oberhalb des oben genannten 10% Grenzwertes. Eine Optimierung der Assaybedingungen (wie beispielsweise Pufferzusammensetzung, Detergenzien oder Optimierung der Beladungsdichte der Sensoren) könnten diese leichten Abweichungen beseitigen. Massentransport-Effekte können hier anhand der Kurvenverläufe ausgeschlossen werden. Auf Grundlage der eben genannten Beobachtungen, sowie der R^2 -Werte der einzelnen Modelle (Tabelle 6.10), wurden nur die Konzentrationen 2.5 bis 20 μM für eine globale exponentielle Regression herangezogen. Tabelle 6.9 zeigt die erhaltenen Geschwindigkeitskonstanten k_d und k_a , sowie die darauf beruhenden Affinitäten K_D aller Messungen. Die mittlere Affinität wurde mit 13 μM bestimmt und wich nur geringfügig von dem Ergebnis des Langmuir-Modells ab.

Der lineare Zusammenhang zwischen der Konzentration von XN und der empirischen Geschwindigkeitskonstanten k_{obs} wurde ebenfalls untersucht. Abbildung 6.33 zeigt die erhaltene Regressionsgerade, basierend auf den k_{obs} -Werten der sechs Messungen. Auch wenn die k_{obs} -Werte eine deutlich stärkere Streuung als bei ADP zeigten, ist dennoch ein linearer Zusammenhang zu erkennen. Die so erhaltene Affinität liegt mit 3.37 μM etwas höher, aber in der selben Größenordnung.

Tabelle 6.9: Ergebnisse der kinetischen Analyse mittels globaler exponentieller Regression der XN-Messungen

Nr.	k_d	$\Delta_{Reg.}$	k_a	$\Delta_{Reg.}$	K_D	$\Delta_{Reg.}$
		$[\text{s}^{-1}]$		$[\text{L } \mu\text{mol s}^{-1}]$		$[\mu\text{mol L}^{-1}]$
1a	0.0935	$5.6 \cdot 10^{-4}$	$7.0 \cdot 10^{-3}$	$9.9 \cdot 10^{-5}$	13.4	0.3
b	0.1227	$6.2 \cdot 10^{-4}$	$7.1 \cdot 10^{-3}$	$1.0 \cdot 10^{-4}$	17.3	0.3
2a	0.0513	$1.9 \cdot 10^{-4}$	$5.8 \cdot 10^{-3}$	$4.8 \cdot 10^{-5}$	8.8	0.1
b	0.0659	$2.4 \cdot 10^{-4}$	$5.8 \cdot 10^{-3}$	$4.0 \cdot 10^{-5}$	11.3	0.1
3a	0.0903	$5.6 \cdot 10^{-4}$	$7.8 \cdot 10^{-3}$	$1.2 \cdot 10^{-4}$	11.7	0.3
b	0.1232	$6.7 \cdot 10^{-4}$	$7.5 \cdot 10^{-3}$	$1.0 \cdot 10^{-4}$	16.4	0.3
	\bar{k}_d	σ	\bar{k}_a	σ	\bar{K}_D	σ
	0.0912	0.0267	$6.8 \cdot 10^{-3}$	$7.8 \cdot 10^{-4}$	13.2	3.0

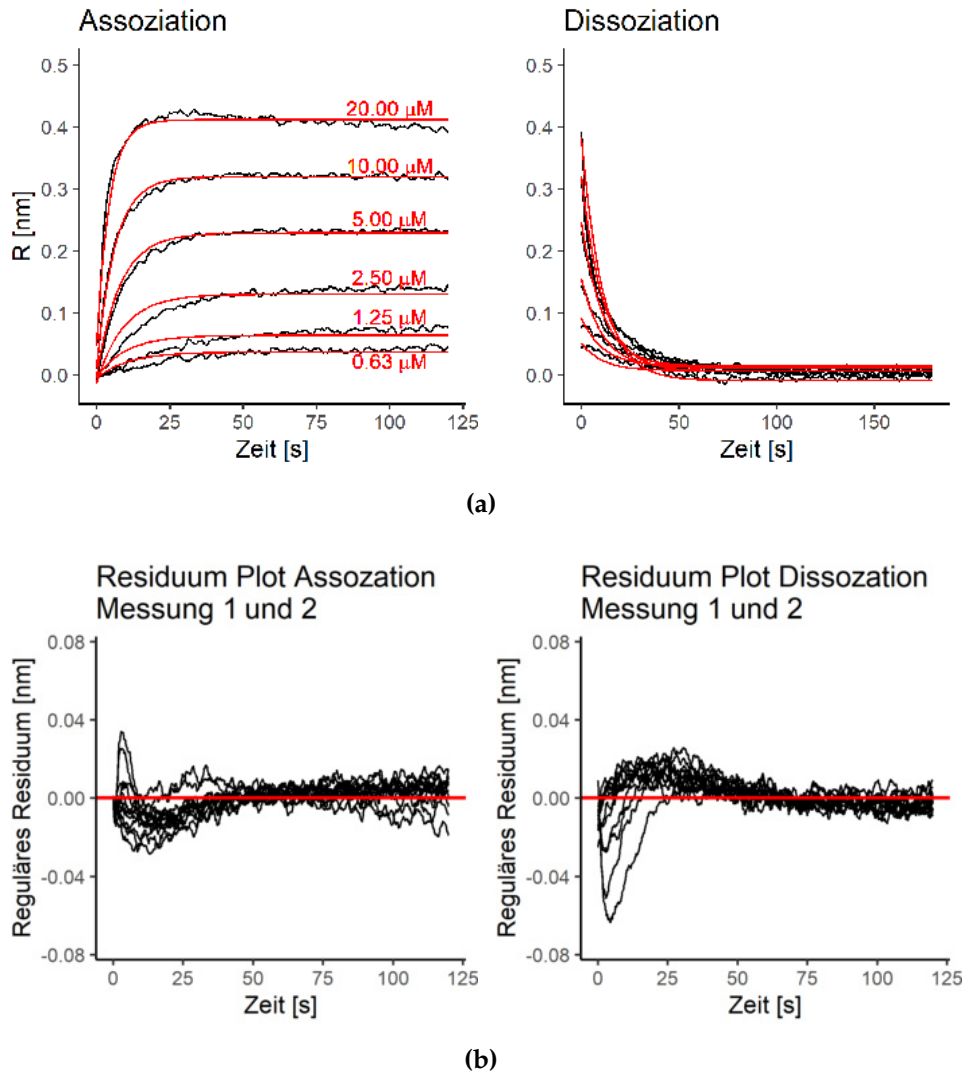
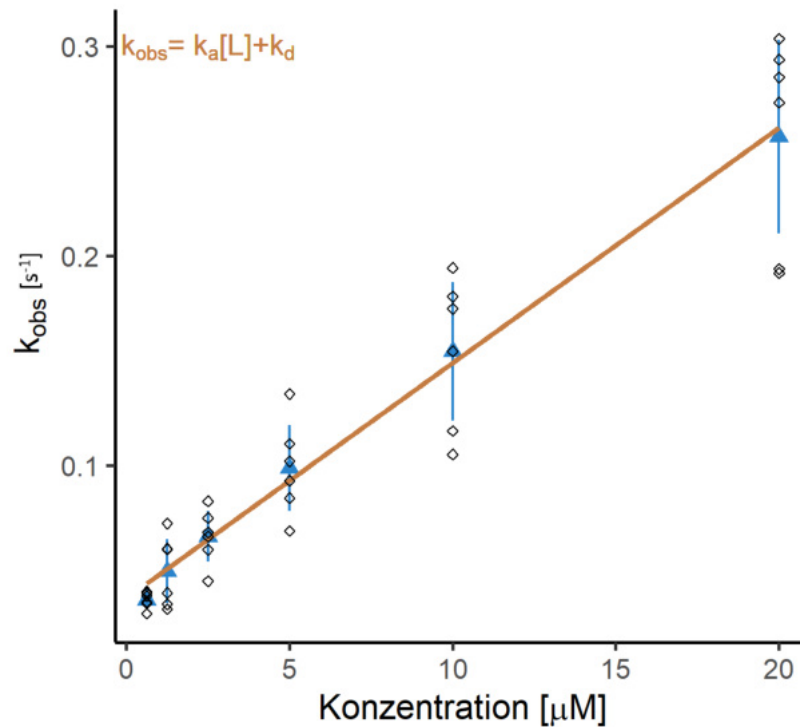


Abbildung 6.26: Ergebnisse der kinetischen Auswertung für einen Sensor der XN-Messungen

(a) Globale Auswertung der Messung 1, (b) Residuum Kurven für Messungen 1 und 2 (für Details siehe Text).

Tabelle 6.10: R^2 Werte der einzelnen Regressionen für ein 1:1-Modell

Konzentration in [μM]	0.625	1.25	2.5	5.0	10.0	20.0
Messung						
Assoziation 1. Messung R^2	0.598	0.759	0.859	0.973	0.977	0.976
Assoziation 2. Messung R^2	0.701	0.707	0.875	0.978	0.989	0.956
Dissoziation 1. Messung R^2	0.747	0.814	0.880	0.944	0.967	0.925
Dissoziation 2. Messung R^2	0.633	0.774	0.861	0.949	0.964	0.962



k_d	$\Delta_{Reg.}$	k_a	$\Delta_{Reg.}$	K_D	$\Delta_{Reg.}$	R^2
	$[s^{-1}]$	$[L \mu mol s^{-1}]$		$[\mu mol L^{-1}]$		
0.0367	0.0063	0.011	$6.69 \cdot 10^{-4}$	3.37	0.76	0.889

Abbildung 6.27: Kinetische Auswertung mittels empirischer Geschwindigkeitskonstanten k_{obs}

Deutlich ist ein linearer Zusammenhang zwischen der Konzentration an XN und den empirischen Geschwindigkeitskonstanten zu sehen, der somit für eine 1:1-Bindung von XN an die N-Domäne spricht.

6.2.2.4 Messungen an der ND1-Domäne

Bei der Immobilisierung von Proteinen besteht die Gefahr, dass artifizielle Bindestellen für kleine Moleküle entstehen können. Um sicherzugehen, dass die beobachtete Bindung an die N-Domäne nicht auf einem solchen Effekt beruht, wurden die Messungen mit dem ND1-Konstrukt wiederholt. Durch Verwendung eines komplett anderen Konstruktes und der somit anderen Beschaffenheit der entstehenden Sensoroberfläche können solche möglichen falsch-positiven Signale identifiziert werden [292]. Aufgrund von Problemen mit den für die Messungen genutzten SSA-Sensoren seitens des Herstellers (siehe Kapitel 5.5.2.1), wurden die Messungen zunächst an mit p97-ND1 beladenen SAX-Sensoren durchgeführt. Diese Sensoren sind laut Hersteller aufgrund ihrer geringen Streptavidindichte

auf dem Sensor nicht sensitiv genug, um die Bindung von kleinen Molekülen zu detektieren. Es wurde daher versucht, die Sensoren so stark wie möglich mit p97-ND1 zu beladen, um dennoch ein Messsignal zu erhalten. Die maximal erreichbare Beladungsstärke von p97-ND1 stellte hier eine Verschiebung von ca. 8 nm dar. Die Positivkontrolle mittels einer 500 nM ADP-Lösung ergab hier immerhin im Mittel eine Verschiebung von 0.2 nm⁶. Mit dieser Beladungsstärke war es schlussendlich möglich, schwache Signale für die Bindung von XN an p97 zu erhalten. Abbildung 6.28 zeigt zwei Sensorgramme der insgesamt 8 durchgeführten Messungen (4 unterschiedliche Sensoren mit je 2 Messungen, siehe auch Anlage D). Wie zu erwarten, lagen die Verschiebungen größtenteils unter 0.05 nm. Bei höheren XN-Konzentrationen zwischen 12.5 und 25 μ M kam es zu stark artifiziellen Kurven, mit atypischen Kurvenverläufen und nicht wieder Erreichen der Basislinie. Eine vermutliche Ursache hierfür ist hier die Wechselwirkung der Sensoren mit bereits einsetzender Aggregatbildung des XN. Der SAX-Sensor scheint hierbei empfindlicher auf Aggregate zu reagieren als der zuvor benutzte SSA-Sensor.

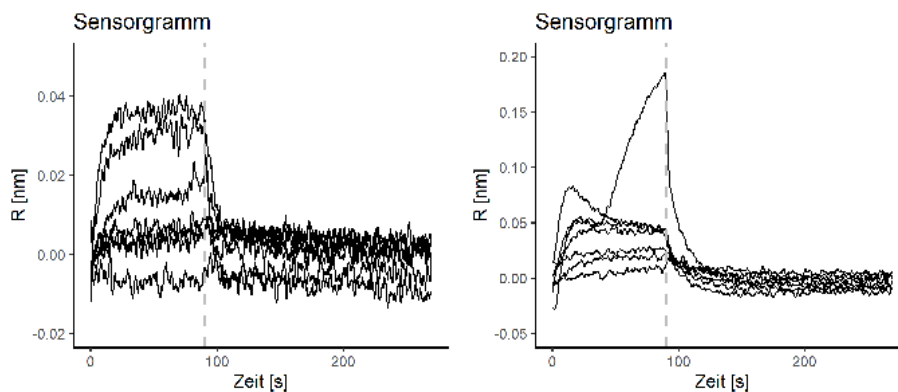


Abbildung 6.28: Sensorgramme zweier XN-Messung unter Verwendung von SAX-Sensoren

Im späteren Verlauf der Arbeit konnten zusätzliche Messungen unter Verwendung von SSA-Sensoren durchgeführt werden. Wie zu erwarten, verbesserte sich hierbei die Signalqualität erheblich und bestätigte noch einmal die Bindung von XN an p97 (Abbildung 6.29).

⁶Zum Vergleich: Die an der UiB in Bergen gemessenen Signalen zeigten Verschiebungen von 0.6 nm mit SSA-Sensoren.

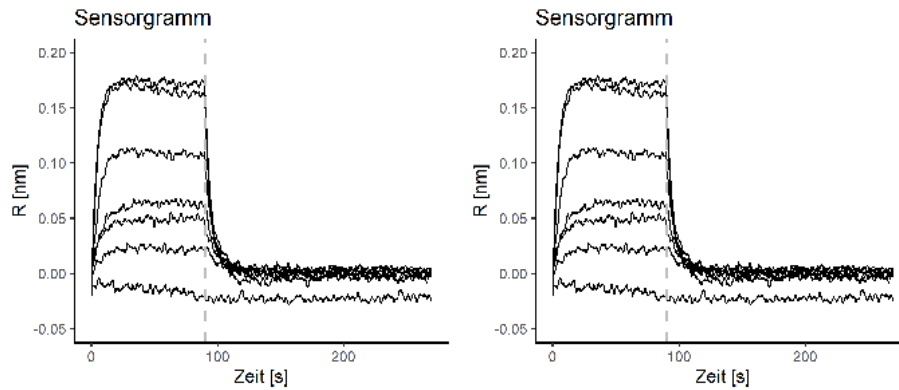
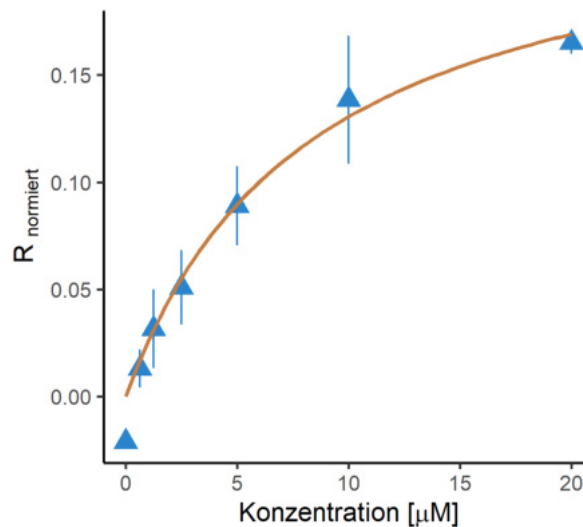


Abbildung 6.29: Sensorgramme zweier XN+Messungen unter Verwendung von SSA-Sensoren

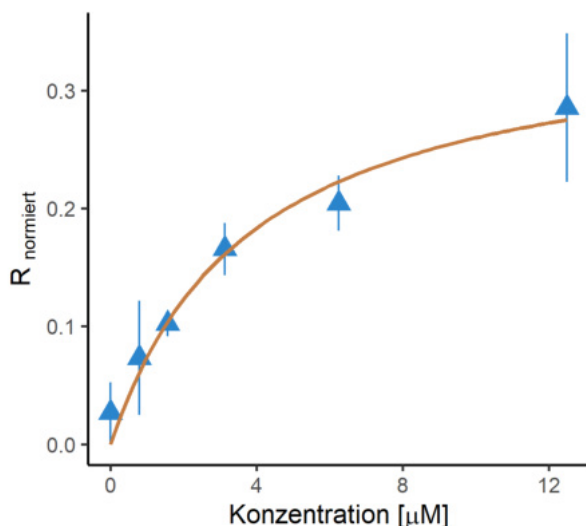
Die erhaltenen Messdaten wurden analog zu der p97 N-Domäne unter Annahme verschiedener Modelle ausgewertet. Abbildungen 6.30 und 6.31 zeigen die Ergebnisse der Auswertung unter Verwendung eines Langmuir-Modells. Die so erhaltenen Affinitäten, 8,6 μM (SSA) bzw. 3,4 μM (SAX), sind mit den für die N-Domäne erhaltenen Werten sehr gut vergleichbar.



K_D	$\Delta_{Reg.}$	R_{max}	R^2
[$\mu\text{mol L}^{-1}$]			
8.6	0.9	0.237	0.999

Abbildung 6.30: Langmuir-Modell für die XN-Messungen an der ND1-Domäne

1:1-Langmuir-Modell basierend auf vier Messungen (SSA-Sensoren). Die Verschiebungen wurden gegen die Signale einer 500 nM ADP-Lösung normiert.



K_D	$\Delta_{\text{Reg.}}$	R_{max}	R^2
[$\mu\text{mol L}^{-1}$]			
3.4	0.8	0.33	0.951

Abbildung 6.31: Langmuir-Modell für die XN-Messungen an der ND1-Domäne

1:1-Langmuir-Modell basierend auf acht Messungen (SAX-Sensoren). Der Messpunkt bei einer Konzentration von 12.5 μM wurde nur anhand von vier Messungen gewonnen. Der Zeitpunkt lag für diese Werte bei 15 Sekunden, bevor eine Drift im Signal zu erkennen war. Die Verschiebungen wurden gegen die Signale einer 500 nM ADP-Lösung normiert.

Für die erhaltenen Daten wurden ebenfalls kinetische 1:1-Modelle erstellt. Auch wenn bei den SAX-Sensoren das Rauschen aufgrund der geringen Verschiebung relativ hoch ist, zeigen die Daten beider Sensortypen eine sehr gute Übereinstimmung mit einem 1:1-Modell (6.32). Im Gegensatz zu den Messungen mit der N-Domäne zeigten hier auch die Konzentrationen von 0.625 und 1.25 μM eine sehr gute Korrelation. Die kinetischen Daten der Auswertungen sind in den Tabellen 6.11 und 6.12 zusammengefasst.

Die Auswertung des linearen Zusammenhangs zwischen der empirischen Geschwindigkeitskonstante k_{obs} und der XN-Konzentration wurde aufgrund des viel schlechteren S/N-Verhältnisses der SAX-Sensoren, nur für die Messungen mit SSA-Sensoren durchgeführt. Abbildung 6.33 zeigt eine sehr starke Streuung der k_{obs} Werte für kleinere Konzentrationen. Die stark abweichenden Werte stammen hier vom ersten SSA-Sensor, während der zweite SSA-Sensor einen sehr guten linearen Zusammenhang zeigt. Für die Konzentrationen 0.625 bis 2.5 μM

wurden daher die Werte des ersten SSA-Sensors als Ausreißer behandelt und die lineare Regression mit den Werten des zweiten SSA-Sensors durchgeführt. Der so erhaltene Affinitätswert ist mit den bisher ermittelten Affinitätswerten nahezu identisch.

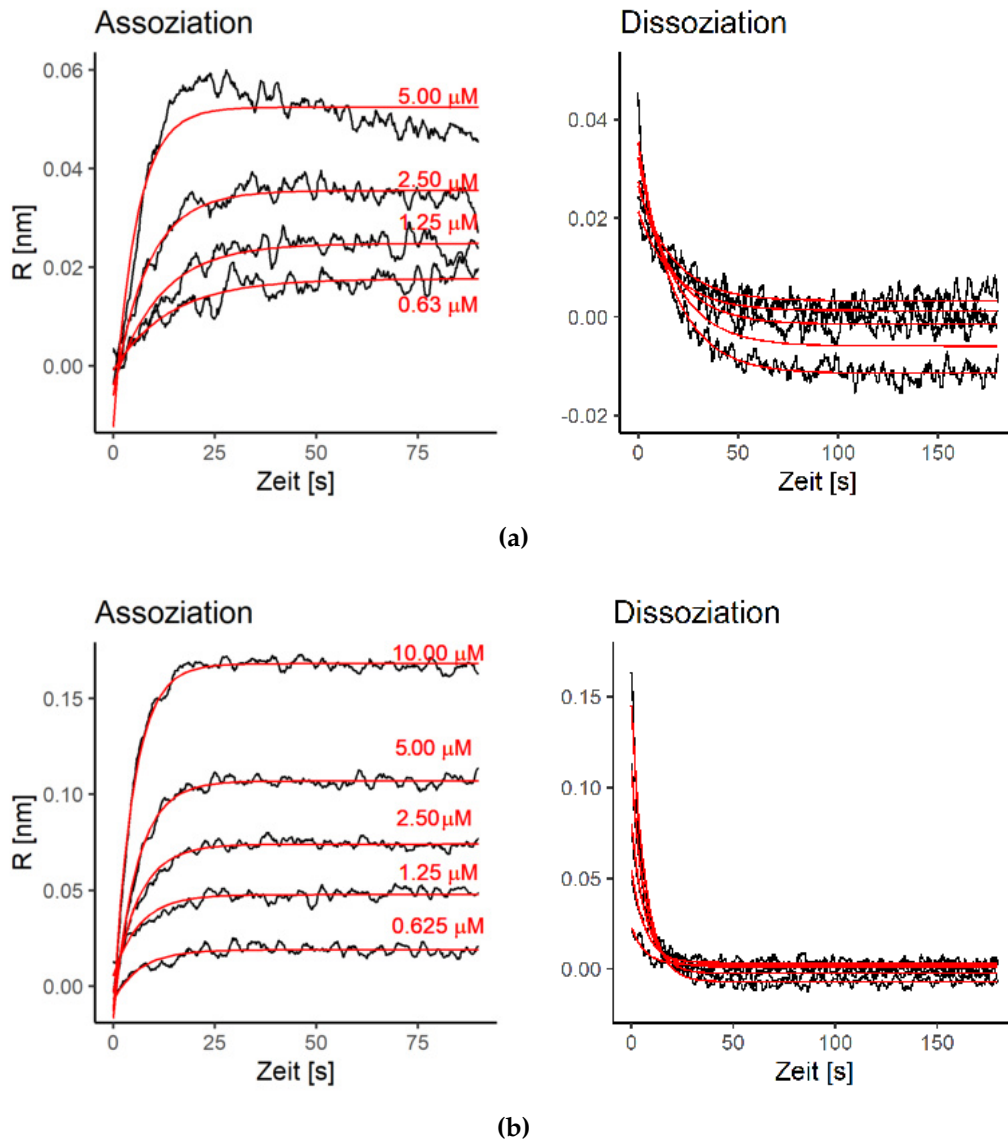


Abbildung 6.32: Ergebnisse der kinetischen Auswertung von XN an der ND1-Domäne für einen jeweils einen SAX- und SSA-Sensor

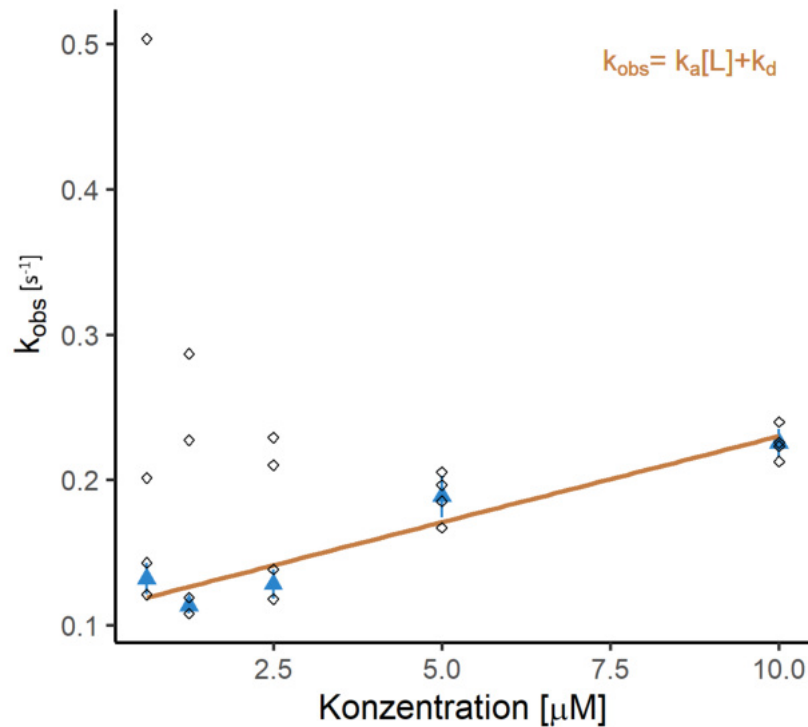
(a) Globale Auswertung der XN-Messungen mit einem SAX-Sensor, (b) Globale Auswertung der XN-Messungen mit einem SSA-Sensor (für Details siehe Text).

Tabelle 6.11: Ergebnisse der kinetischen Analyse mittels globaler exponentieller Regression für XN an der ND1-Domäne (Messungen mit SSA-Sensoren)

Nr.	k_d	$\Delta_{Reg.}$	k_a	$\Delta_{Reg.}$	K_D	$\Delta_{Reg.}$
		$[s^{-1}]$		$[L \mu mol s^{-1}]$		$[\mu mol L^{-1}]$
1a	0.1483	$8.8 \cdot 10^{-4}$	$5.7 \cdot 10^{-3}$	$1.6 \cdot 10^{-4}$	25.9	0.9
b	0.1516	$9.9 \cdot 10^{-4}$	$6.8 \cdot 10^{-3}$	$1.7 \cdot 10^{-4}$	22.2	0.7
2a	0.1644	$1.5 \cdot 10^{-3}$	$7.6 \cdot 10^{-3}$	$3.1 \cdot 10^{-4}$	21.8	1.1
b	0.1668	$1.4 \cdot 10^{-3}$	$6.1 \cdot 10^{-3}$	$3.0 \cdot 10^{-4}$	27.2	1.6
	\bar{k}_d	σ	\bar{k}_a	σ	\bar{K}_D	σ
	0.1578	$8.0 \cdot 10^{-3}$	$6.6 \cdot 10^{-3}$	$7.0 \cdot 10^{-4}$	24.3	2.3

Tabelle 6.12: Ergebnisse der kinetischen Analyse mittels globaler exponentieller Regression für XN an der ND1-Domäne (Messungen mit SAX-Sensoren)

Nr.	k_d	$\Delta_{Reg.}$	k_a	$\Delta_{Reg.}$	K_D	$\Delta_{Reg.}$
		$[s^{-1}]$		$[L \mu mol s^{-1}]$		$[\mu mol L^{-1}]$
1	0.0542	$1.8 \cdot 10^{-4}$	$6.4 \cdot 10^{-3}$	$6.5 \cdot 10^{-5}$	8.5	0.1
2	0.0767	$4.3 \cdot 10^{-4}$	$5.0 \cdot 10^{-3}$	$1.1 \cdot 10^{-4}$	15.3	0.4
3	0.0567	$4.1 \cdot 10^{-4}$	$2.0 \cdot 10^{-3}$	$3.6 \cdot 10^{-4}$	2.9	0.1
4	0.0568	$4.7 \cdot 10^{-4}$	$8.6 \cdot 10^{-3}$	$1.5 \cdot 10^{-4}$	6.6	0.2
	\bar{k}_d	σ	\bar{k}_a	σ	\bar{K}_D	σ
	0.0611	$9.1 \cdot 10^{-3}$	$10.0 \cdot 10^{-3}$	$5.8 \cdot 10^{-3}$	8.3	4.5



k_d	$\Delta_{Reg.}$	k_a	$\Delta_{Reg.}$	K_D	$\Delta_{Reg.}$	R^2
	[s^{-1}]	[$L \mu mol s^{-1}$]		[$\mu mol L^{-1}$]		
0.1039	0.0085	0.012	0.0019	8.38	1.33	0.910

Abbildung 6.33: Kinetische Auswertung mittels empirischer Geschwindigkeitskonstanten für die Bindung von XN an die ND1-Domäne (SSA-Sensoren)

Deutlich ist ein linearer Zusammenhang für einen der beiden SSA-Sensoren zwischen der Konzentration an XN und den empirischen Geschwindigkeitskonstanten zu sehen. Dieser spricht für eine 1:1-Bindung von XN an p97.

6.2.2.5 Untersuchung einer konformationsbedingten Abhängigkeit der Bindung von XN an p97-ND1

Um den Einfluss der Nukleotid-abhängigen Konformation von p97 auf die Bindung von XN zu klären, wurden Messungen in Abwesenheit und Anwesenheit von 1 mM ADP bzw. 1 mM AMP-PNP durchgeführt. In 6.34 (b) sind die Rohdaten der Messungen zu sehen. In Schwarz ist der Referenzsensor beladen mit p97-ND1 ohne XN dargestellt, der die jeweilige Nukleotid-Konformation und Bindungen anzeigt. In Blau ist der mit p97-ND1 beladene Sensor gezeigt, der

jeweils viermal die Verschiebung eines zusätzlich mit $5 \mu\text{M}$ XN beschickten Napfes misst. Die Verschiebung der Basislinie am Anfang der Messungen war systembedingt und betraf, unabhängig voneinander, beide Sensoren, wodurch dies als Störsignal durch die Referenzierung der Sensoren herausgerechnet werden konnte.

Die so erhaltenen Sensorgramme wurden mittels eines 1:1-Modells kinetisch ausgewertet und die jeweiligen Affinitäten von XN zu p97 bestimmt. Abbildung 6.34 (a) zeigt die Affinität von XN zu p97 in Abhängigkeit der Nukleotidkonformation. Wie zu sehen ist, ergab sich kein signifikanter Unterschied in der Affinität von XN an p97-ND1 für die unterschiedlichen Konformationen von p97 und die Affinität ist somit von der p97 Konformation unabhängig.

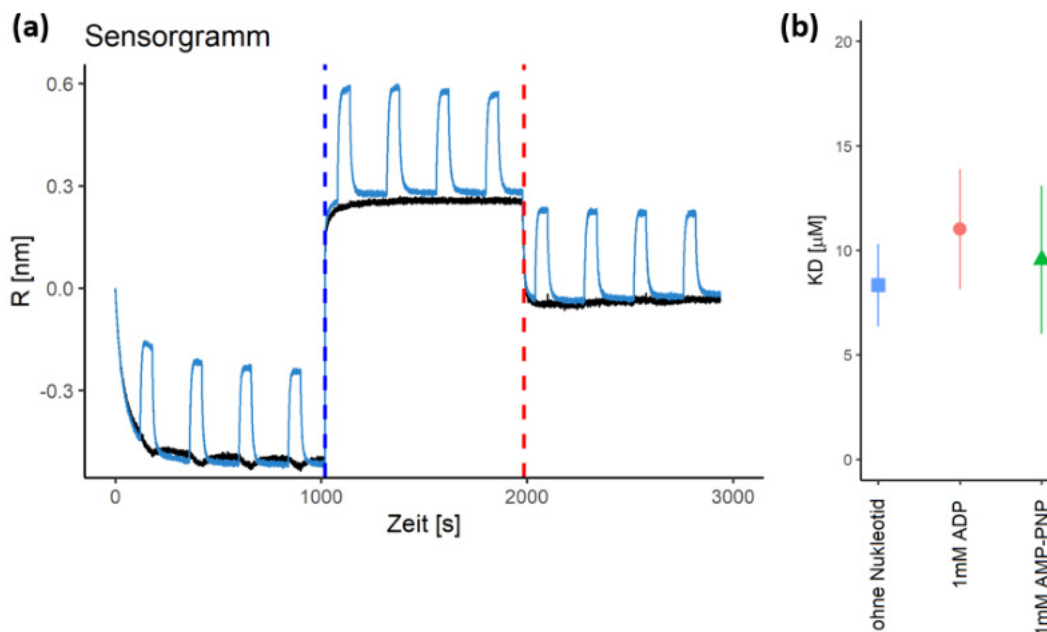


Abbildung 6.34: Ergebnisse der Untersuchung einer Konformationsbedingten Abhängigkeit der Bindung von XN an p97

(a) Rohsignale der Messungen. Zu Beginn sind die Verschiebung für vier XN Messungen im Nukleotid-freie, ab Blau in Gegenwart von 1 mM ADP und ab Rot in Gegenwart von 1 mM AMP-PNP zu sehen.

(b) Ergebnisse der kinetischen Auswertungen der jeweils vier Messungen. Für alle drei Nukleotid Zustände wurde der gleiche K_D -Wert für XN erhalten.

6.2.2.6 Zusammenfassung und Diskussion

Tabelle 6.13: Zusammenfassung aller Messungen an der N- und ND1-Domäne

Modell	Konstrukt	K_D	Fehler	k_d	k_a
			[$\mu\text{mol L}^{-1}$]	[s^{-1}]	[$\text{L } \mu\text{mol s}^{-1}$]
Glob. Exp. Reg.	N	13.2	3.0 (σ_{KD})	0.0912	0.0068
	ND1 _(SSA)	24.3	2.3 (σ_{KD})	0.1578	0.0066
	ND1 _(SAX)	8.3	4.5 (σ_{KD})	0.0611	0.01
Lineare Reg. k_{obs}	N	3.4	0.8 ($\Delta_{Reg.}$)	0.0367	0.011
	ND1 _(SSA)	8.3	1.3 ($\Delta_{Reg.}$)	0.1039	0.012
Langmuir	N	8.3	0.5 ($\Delta_{Reg.}$)		
	ND1 _(SSA)	8.6	0.9 ($\Delta_{Reg.}$)		
	ND1 _(SAX)	3.4	0.8 ($\Delta_{Reg.}$)		

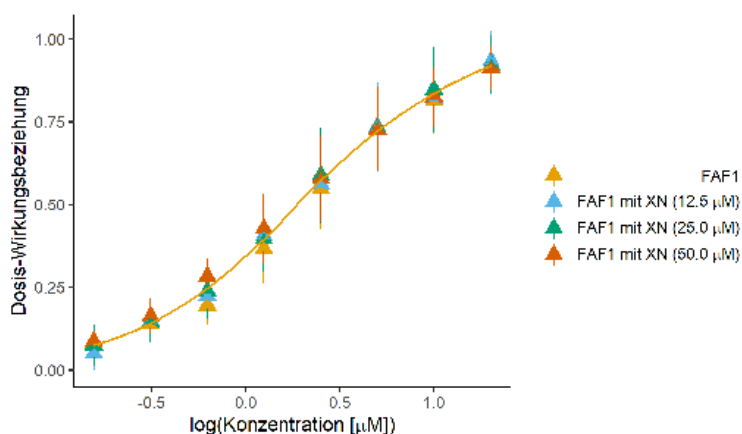
Tabelle 6.13 fasst die Ergebnisse der XN Messungen zusammen. Generell konnte die Bindung von XN an p97 bestätigt werden. Die durch unterschiedliche Modelle errechneten Affinitäten liegen alle in einem ähnlichen einstelligen bis unteren zweistelligen μM Bereich. Die Messungen mit zwei unterschiedlichen Sensortypen (SSA und SAX), sowie zwei unterschiedlichen Konstrukten bestätigen, dass es sich bei der gemessenen Bindung nicht um ein durch die Immobilisierung artifizielles Signal handelt⁷. Die als gleich einzustufenden Affinitäten an die N-, bzw. ND1-Domäne zeigen des Weiteren, dass sich die Bindestelle in der N-Domäne von p97 befinden muss. Die gute Übereinstimmung mit einem 1:1-Modell weist auf eine definierte Bindestelle und Orientierung hin. Da keine Veränderung in der Affinität von XN bei der An- oder Abwesenheit der D1-Domäne zu sehen war, sowie die Tatsache, dass die Affinität von XN unabhängig von der Nukleotid Konformation von p97 ist, lassen den Schluss zu, dass sich die Bindestelle auf der, der D1-Domäne abgewandten Seite von p97-N befinden könnte. Trotzdem wäre hier zur Untermauerung der Bindestelle in der N-Domäne eine zusätzliche Messung mit einem p97-D1D2-Konstrukt von Interesse gewesen. Hier sollte kein Signal für XN erwartet werden. Ein entsprechendes Konstrukt stand allerdings zum Zeitpunkt der Messungen nicht zur Verfügung.

⁷Dennoch sollten hier Messungen mit einer anderen Methode die Bindung von XN noch einmal bestätigen. Hierzu sind ^1H - ^{15}N -HSQC Experimente in Planung, die zusätzlich auch bei der Aufklärung der Bindestelle helfen sollen.

Um die mögliche Binderegion weiter einzugrenzen wurde ein Fluoreszenz-anisotropie-Kompetitionsassay mit fluoreszenzmarkierten FAF1-UBX durchgeführt. FAF1 gehört zu der Gruppe der UBX-Domänen enthaltene Proteine. Der Kofaktor ist in den Fas-Apoptose Weg involviert, wo es an die Fas-Todes-Domäne bindet und Apoptose induziert. Über eine UBX- sowie eine UBA-Domäne ist FAF1 mit dem Ubiquitin-Proteasom-System verknüpft. Für die Bindung von FAF1-UBX wurde in der Literatur via ITC Messungen eine Affinität von $3.9 \mu\text{M}$ mit einer 1:1-Stöchiometrie an p97(FL) berichtet [26, 295]. Abbildung 6.35 zeigt die Ergebnisse dieses Assays. Für die Bindung von FAF1-UBX an die ND1-Domäne konnte eine Affinität von $2.1 \mu\text{M}$ erhalten werden, was in Einklang mit der von [295] berichteten Affinität ist. Es konnte kein Einfluss von XN auf die Bindung von FAF1-UBX an die ND1-Domäne beobachtet werden. Eine mögliche Bindung von XN zwischen die Nn- und Nc-Subdomäne konnte daher nicht bestätigt werden. Versuche einen ähnlichen Assay mit dem SHP-Bindemotiv von UFD1 durchzuführen, und somit eine Bindung in die SHP-Region nach zuweisen, schlugen aufgrund der geringen Affinität des Peptids im dreistelligen μM Bereich fehl. Die Einflüsse der verwendeten p97-Konzentrationen auf die Viskosität der Lösung waren zu hoch und somit die gemessenen Fluoreszenzanisotropien nicht mehr eindeutig auswertbar. Ob der nicht beobachtete Einfluss von XN auf die FAF1-UBX Bindung auf der Tatsache beruht, dass XN nicht in den Spalt zwischen Nn- und Nc-Subdomäne bindet, lässt sich hier nicht abschließend klären. Ein höherer Überschuss an XN wäre hierfür nötig, ist aber durch die geringe Löslichkeit von XN nicht möglich. Eine Lokalisation der N-Domänen-Bindestelle ist anhand der Datenlage dieser Arbeit leider nicht möglich. Im Fortgang des Projektes sollen bereits begonnene $^1\text{H},^{15}\text{N}$ -HSQC NMR-Messungen die Frage der Bindestelle von XN abschließend klären. Auch Ko-Kristallisations- und Soaking-Experimente mit p97-ND1 werden versuchen die strukturellen Fragen der Bindung zu klären.

Problematisch ist allerdings die schlechte Löslichkeit von XN, gerade für die Kristallographie. Eine Lösung wäre, die Löslichkeit von XN, durch die Einführung polarer Gruppen zu erhöhen, auch wenn hier stets die Gefahr des Verlusts der Affinität besteht, gerade wenn keinerlei Strukturdaten vorliegen. Eine Möglichkeit zur Verbesserung der Löslichkeit wäre zum Beispiel die Hydratisierung der Prenylseitenkette. Die Einführung einer Hydroxygruppe durch Addition an die Doppelbindung der Prenylseitenkette könnte durch eine Oxymercurierung durchgeführt werden [296]. Daneben könnten auch Ansätze, wie sie im Kontext von *Prodrugs* zur Löslichkeitsverbesserung zum Einsatz kommen, von Interesse sein. Die in XN vorhandenen oder durch Hydratisierung eingeführten

Alkoholgruppen könnten mit löslichkeitsverbessernden Gruppen (z.B. Carbonsäuren, Aminosäuren) verestert werden [297]. Die erhaltenen Ester wären hierbei nur für *in vitro* Experimente geeignet, da sie *in vivo* keine ausreichende Stabilität besitzen. Die Einführung von Halogenen, wie Brom oder Iod (z.B. durch Bromierung der Prenylseitenkette) wäre ein weiterer interessanter Ansatz für die Kristallographie, auch wenn diese sich negativ auf die Löslichkeit auswirken. Das anomale Signal eines Brom- oder Iodatoms könnte in Röntgenstruktorexperimenten die Lokalisierung von XN deutlich erleichtern. Unabhängig vom Ausgang der strukturellen Experimente wären Struktur-Wirkungsbeziehungen sicherlich ein zweiter zu verfolgender Ansatz. In der Literatur findet man eine Reihe von XN-Analoga, die durch Totalsynthesen hergestellt wurden [296, 298]. Auf Grundlage dieser Synthesen könnten XN-Analoga weitere Aufschlüsse geben, welche Bereiche von XN für die Bindung vorwiegend relevant sind und somit das Bindeepitop von XN darstellen.



Messung	K_D	$\Delta_{Reg.}$
		$[\mu\text{mol L}^{-1}]$
FAF1	2.1	0.13
FAF1 mit 12.5 μM XN	2.0	0.10
FAF1 mit 25.0 μM XN	1.9	0.03
FAF1 mit 50.0 μM XN	1.7	0.04

Abbildung 6.35: Fluoreszenzanisotropie Wettbewerbsassay

Acht Konzentrationen einer p97-ND1 1:1-Verdünnungsreihe startend bei 20 μM wurden gegen 5 μM fluoreszenzmarkiertem FAF1 titriert, in Präsenz von 0, 12.5, 25 und 50 μM XN.

6.2.3 Fragmentscreenings

Im Folgenden sollen die Ergebnisse der beiden initialen BLI-Screenings an der ND1- und N-Domäne, der anschließenden BLI-Messungen zur Konzentrationsabhängigkeit, sowie der STD-NMR Messungen dargestellt werden.

6.2.3.1 Initiales-Screening mit p97-ND1

Zunächst wurde zur Etablierung des Screenings der UiB Fragmentbibliothek mittels BLI ein initiales-Screening mit der ND1-Domäne durchgeführt. Anhand der oben genannten Ergebnisse der ADP Messungen lag für p97-ND1 eine valide Positivkontrolle vor, mit deren Hilfe die Qualität und Aussagekraft des Screenings beurteilt werden konnte. Abbildung 6.36 zeigt die Z-Faktoren für jede gemessene Screeningplatte (UiB-Bibliothek Platten A2 bis I2; die 10 Fragmente aus dem virtuellen Screenings sind als „SHP“ bezeichnet). 10 der 13 Platten spiegeln laut Tabelle 5.6 mit einem mittleren Z-Faktor von 0.76 qualitativ sehr gute Assays wider, bei denen eine hervorragende Trennung zwischen positiven und negativen Signalen gegeben war. Zwei Screeningplatten (D2, I2) liegen mit im Werten von 0.49 und 0.41 im Grenzbereich zu einem Assay mit schwacher Aussagekraft. Nur Screeningplatte AB3 weist mit einem Z-Faktor von 0.22 einen äußerst niedrigen Z-Faktor auf. Die Sensoren der Platte AB3 wurden bereits zum Vermessen zweier Screeningplatten eingesetzt und anschließend für sechs Stunden zwischengelagert, bevor sie erneut eingesetzt wurden. Diese Lagerung wirkte sich offensichtlich negativ auf die Beladung mit p97 aus. Aufgrund der nur sehr begrenzten Messzeit konnte die Platte nicht erneut vermessen werden. Zusammenfassend lässt sich die Qualität des durchgeführten Screenings, mit einem mittleren Z-Faktor von 0.67 als sehr gut einstufen.

Abbildung 6.37 zeigt eine Übersicht aller gemessenen Verschiebungen der 688 Fragmente⁸ nach Prozessierung mittels der Doppelreferenzmethode. Zusätzlich sind die Verschiebungen der durchgeführten Positiv- und Negativkontrollen zu sehen. Die Abbildung spiegelt die zu erwartende Verteilung eines Screenings wider. Die meisten Fragmente zeigen keine Signale und bilden eine deutliche Basislinie (grau), in der auch die Negativkontrollen (schwarz) zu finden sind. Die Positivkontrollen (rot) heben sich deutlich hiervon ab. Im dazwischen liegenden Bereich zeigen sich Fragmente mit deutlichen Verschiebungen, die mögliche Treffer darstellen. Die Positivkontrollen der eingesetzten Sensoren zeigen eine deutliche Streuung, was auf unterschiedliche Beladungszustände der Sensoren oder

⁸Anzahl einschließlich der Fragmente aus dem virtuellen Screening

einem Verlust an immobilisiertem p97 zurückgeführt werden kann. Um die Signale der Messungen besser vergleichen zu können wurden diese, wie in Kapitel 5.5.4.2 beschrieben, normiert. In Abbildung 6.38 sind exemplarisch die Zwischenschritte der Normierung für einen Sensor gezeigt. Der so erhaltene R_{norm} -Wert gibt somit die relative Verschiebung eines Fragments im Verhältnis zur Verschiebung der ADP-Positivprobe wider.

Fragmente, die nach dieser Normierung auffällig negative oder positive Verschiebungen zeigten, wurden anhand ihrer Sensorgramme visuell kontrolliert. Fragmente, deren Kurvenverläufe nicht den für BLI-Messungen typischen Verlauf zeigten, wurden entfernt. Abbildung 6.39 zeigt die Verschiebungen aller Fragmente nach der Normierung und visuellen Bereinigung. Die Auswertung dieser Daten ergab eine mittlere relative Verschiebung $\mu_{R_{norm}}$ von 0.050 mit einer Standardabweichung $\sigma_{R_{norm}}$ von 0.0786. Der eingesetzte Grenzwert für die weitere Selektion wurde auf $cutoff = \mu_{R_{norm}} + \sigma_{R_{norm}} = 0.1287$ gesetzt. Von 679⁹ Fragmenten blieben 82 (11.91 %) Fragmente für die weitere Auswertung übrig.

Die Sensorgramme der selektierten 82 Fragmente wurden anschließend nach den in Kapitel 5.5.4.2 genannten Kriterien visuell kontrolliert, wobei 48 (6.97 %) Fragmente für weitere Messungen übrig blieben.

Während der kurzen Aufenthaltszeit von 2.5 Wochen, die zur Etablierung und Durchführung des Screenings im Oktober 2019 an der UiB zur Verfügung standen, konnten von diesen 48 Fragmenten 18¹⁰ mittels einer 1:1-Konzentrationsreihe mit fünf Konzentrationen ausgehend von 1000 μM , auf ihre Dosisabhängigkeit geprüft werden¹¹.

⁹Anzahl Fragmente nach visueller Kontrolle.

¹⁰Hierbei handelte es sich um eine zufällige Auswahl, die anhand der Platzierung auf den Screeningplatten getroffen wurde.

¹¹Aufgrund der Corona-Pandemie 2020 waren weitere Forschungsaufenthalte an der UiB nicht mehr möglich, sodass die restlichen 30 Fragmente nicht mehr vermessen werden konnten.

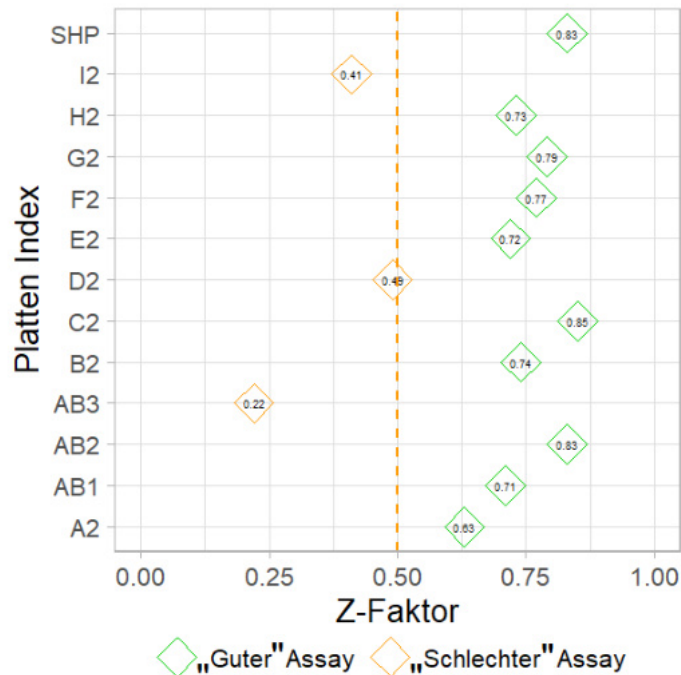


Abbildung 6.36: Z-Faktoren

Übersicht der Z-Faktoren der einzelnen Screeningplatten. 10 der 13 Platten zeigen Z-Faktoren größer 0.5 und spiegeln damit qualitativ sehr gute Assays wider. Zwei Screeningplatten (D2, I2) liegen im Grenzbereich zu einem Assay mit schwacher Aussagekraft. Screeningplatte AB3 weist mit einem Z-Faktor von 0.22 den niedrigsten Z-Faktor auf. Keine der Platten zeigt einen negativen Z-Faktor, der auf einen nicht funktionalen Assay deuten würde.

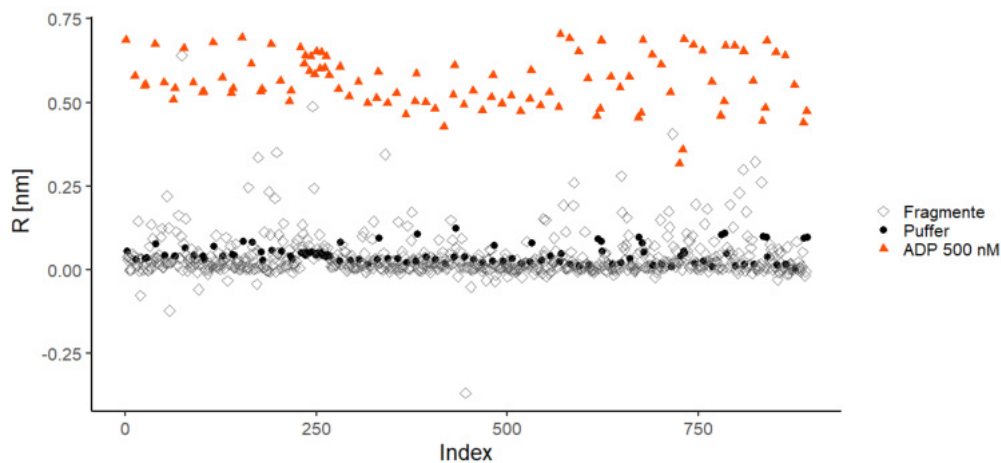


Abbildung 6.37: Gemessene Verschiebungen der Fragmente des initialen-Screening mit p97-ND1

Übersicht aller doppelreferenzierter Verschiebungen der eingesetzten Fragmente im initialen p97-ND1-Screening (grau), einschließlich aller Positiv- (rot) und Negativkontrollen (schwarz). Aufgrund der hier mit eingezeichneten Kontrollmessungen liegt die Indexanzahl höher als 688.

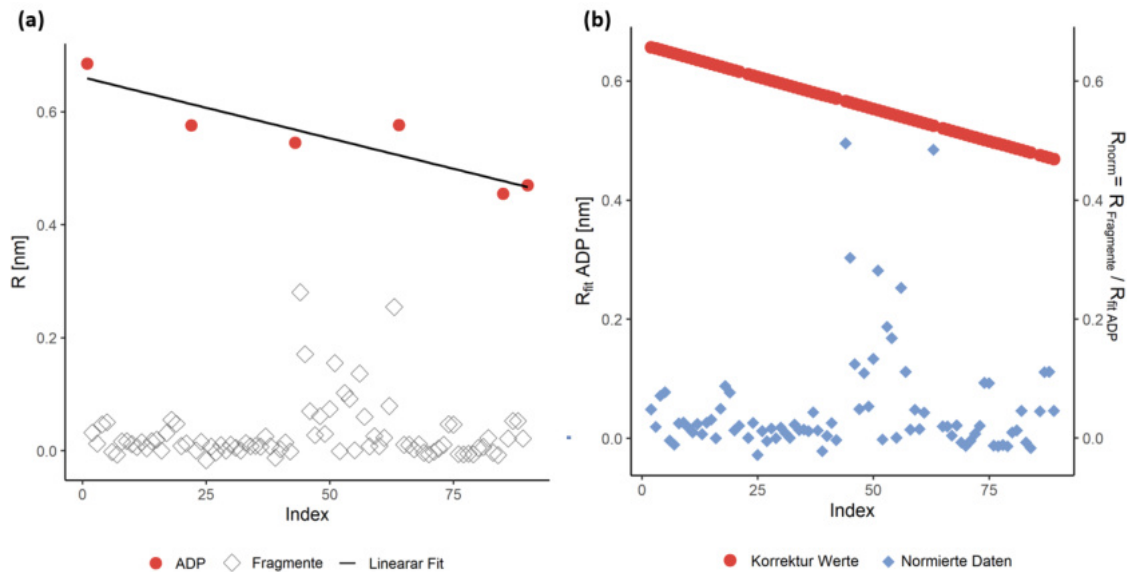


Abbildung 6.38: Normierung der Signale

Zu sehen sind die Schritte der Korrektur am Beispiel eines Sensors des Screenings. Teilabbildung (a) zeigt die doppelreferenzierten Verschiebungen eines Sensors, der für fünf Platten eingesetzt wurde. In Rot sind die vor und nach jeder Platte gemessenen Positivkontrollen eingezeichnet. Anhand dieser wurde durch lineare Regression die Gerade (schwarz) ermittelt.

(b) Anhand der Regressionsgeraden wurde die theoretische Verschiebung durch ADP zum Zeitpunkt der Messung eines jeden Fragments berechnet (rote Kreise). Jede Verschiebung eines Fragments wurde durch die dazugehörige theoretische Verschiebung von ADP geteilt und somit normiert. Die so erhaltenen Werte sind in Blau gezeigt. Durchgeführt wurden die Korrekturen in dafür geschriebenen Python- und R-Skripten.

6.2.3.2 Messung der Konzentrationsabhängigkeit ausgewählter Fragmente des initialen p97-ND1-Screenings

Von den eben genannten 18 Fragmenten zeigten sieben ein konzentrationsabhängiges Signal, welches zur Abschätzung der Affinität mit einem 1:1-Langmuir-Modell ausgewertet wurde. Die Langmuir-Modelle der sieben Fragmente sind in Anlage E zu finden. Die Strukturen und erhaltenen Affinitäten aller sieben Fragmente werden in der zusammenfassenden Übersichtstabelle 6.17 der Fragmentenscreenings mit p97 gezeigt. Im Folgenden soll hier nur auf die in Abbildung 6.40 gezeigten Sensorgramme von vier ausgewählten Fragmenten näher eingegangen werden.

Die Fragmente VIK20¹² und VIK50 repräsentieren mit ihren Sensorgrammen den Großteil der vorgefundenen Kurvenverläufe. Auch wenn sich die zu erwartende Verschiebung (wie oben bei den Messungen mit ADP gezeigt) nicht in der

¹²Abkürzung VIK steht für VIKING.

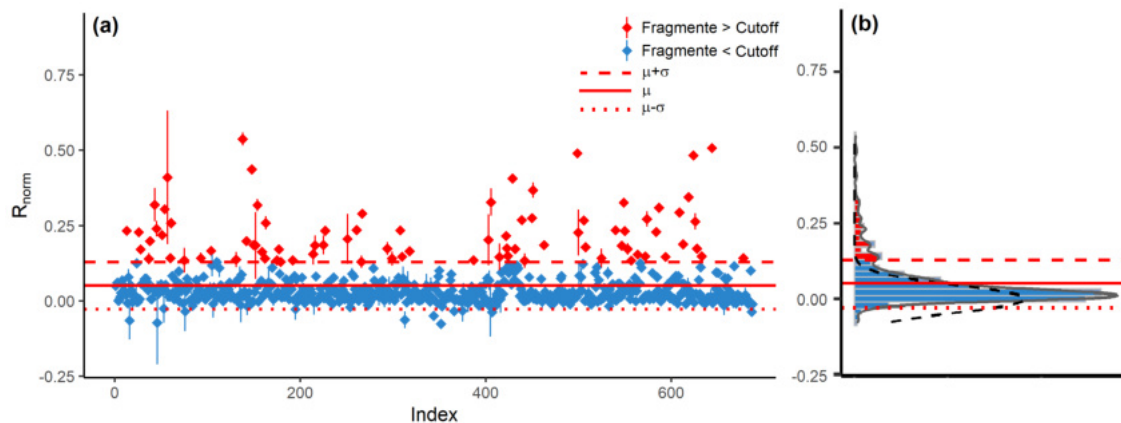


Abbildung 6.39: Übersicht des Fragmentscreenings mit p97-ND1

(a) Zu sehen sind die normierten Verschiebungen der Fragmente. In Rot sind diejenigen Fragmente gezeigt, die oberhalb des gesetzten Grenzwerts von $\mu + \sigma = 0.1287$ lagen. Alle in Blau gezeigten Fragmente wurden nicht weiter verfolgt.

(b) Häufigkeitsverteilung der Verschiebungen der Fragmente überlagert mit einer auf Basis der Daten berechneten Normalverteilung in Schwarz. Deutlich zu sehen ist ein leichtes Abweichen (rot) der Daten von der berechneten Normalverteilung. Diese Abweichung spricht für eine Interaktion mit p97.

Weise berechnen lässt, wie es bei SPR-Experimenten möglich ist, so liegen sie im Vergleich zu ADP in einem realistischen Größenbereich. Deutlich zu sehen sind biphasische Kurvenverläufe, auf deren Ursache in Kapitel 6.2.4.1 im Detail eingegangen werden soll. Die Signale der Zweifachmessungen (rot: 1. Messung, blau: 2. Messung) der jeweiligen Konzentrationen verlaufen annähernd gleich, zeigen Plateauphasen, sowie eine Konzentrationsabhängigkeit. In den meisten Fällen lagen die Signale der zweiten Messung höher als die der ersten Messung. Dies wird darauf zurückgeführt, dass durch die recht lange Assayzeit von fast vier Stunden eine Verdunstung des Wassers aus den Nöpfchen (200 μL Volumen) stattfand, die zu einer Konzentrationserhöhung der Fragmente führte. In den neusten BLI-Geräten kann durch einen Evaporationsschutz dieser Effekt vermindert werden. Das Gerät der UiB besaß diese Vorrichtung jedoch noch nicht¹³.

Das Sensorgramm von VIK90 besaß zwar eine Konzentrationsabhängigkeit, zeigte aber neben einem starken biphasischen Verlauf vor allem bei hohen Konzentrationen keine Plateauphase mehr. Sensorgramme mit einem solchen Erscheinungsbild lassen eine eher unspezifische Bindung des Moleküls vermuten [211]. Die für kleine Konzentrationen erhaltenen Plateauphasen könnten dennoch eine zunächst spezifische Bindung anzeigen, die bei höheren Konzentrationen von

¹³Messungen des tatsächlichen Verlustes an Wasser (z.B. über den Gewichtsverlust der Platten) wurden nicht durchgeführt, da dieser Effekt erst bei der Datenauswertung der beiden Screenings auffiel.

einer unspezifischen überlagert wird. Grundlage des 1:1-Langmuir-Modell ist allerdings das Erreichen eines chemischen Gleichgewichts, wodurch die hier abgeschätzte Affinität extrem fehlerbehaftet ist.

Das Fragment VIK1 zeigte eine dreimal so starke Verschiebung wie die restlichen Fragmente. Zunächst wurde VIK1 aufgrund seiner hohen Affinität zu den bestbindenden Molekülen eingestuft. Spätere Messungen mit dem bei Otava Chemicals nachbestellten Fragment zeigten allerdings nur noch eine schwache Verschiebung und eine deutlich geringere Affinität ($>1000 \mu\text{M}$). Grund könnte hier ein in VIK1 enthaltenes reaktives Michael-System sein. Das Fragment war zum Zeitpunkt der Messungen bereits über zwei Jahre gelagert, und es kann daher nicht ausgeschlossen werden, dass hier Abbau- und/oder Kondensationsprodukte in der Stammlösung vorlagen. Größere Kondensationsprodukte mit einem höheren Molekulargewicht wären eine mögliche Erklärung für die deutlich stärkeren Verschiebungen. Die Ursache des hier zu sehenden Signals sind Gegenstand aktueller analytischer Untersuchungen seitens der UiB. Abschließende Ergebnisse lagen zum Zeitpunkt des Schreibens dieser Arbeit noch nicht vor. Aufgrund dieser Befunde wird das Fragment VIK1 zunächst nicht weiter betrachtet. Das Beispiel zeigt, dass eine analytische Charakterisierung der gefundenen Treffer für den weiteren Verlauf des Projektes sicherlich sinnvoll ist, um solche falsch-positiven-Fragmente zu identifizieren.

6.2.3.3 Zusammenfassung des initialen-Screenings mit p97-ND1

Zusammenfassend lässt sich festhalten, dass das Ziel der Etablierung eines Fragmentscreenings mit p97-ND1 erreicht wurde. Es konnte gezeigt werden, dass das p97-ND1-Konstrukt in den eingesetzten Screeningbedingungen sehr stabil ist. Darüber hinaus konnte die Qualität des Screenings anhand der Z-Faktoren bestätigt werden. Von insgesamt 48 initial gefundenen Treffern konnten zwar leider nur 18 Fragmente näher auf ihre Konzentrationsabhängigkeit untersucht werden. Sieben Fragmente zeigten eine Abhängigkeit der Verschiebung von der Konzentration. Die in diesem Screening identifizierten Fragmente wurden aufgrund des Fokus auf die N-Domäne nicht weiter verfolgt und sind daher nur zur Demonstration der Funktionalität des Assays zu sehen. Die gewonnenen Erkenntnisse flossen jedoch in das initiale-Screening mit der isolierten N-Domäne mit ein.

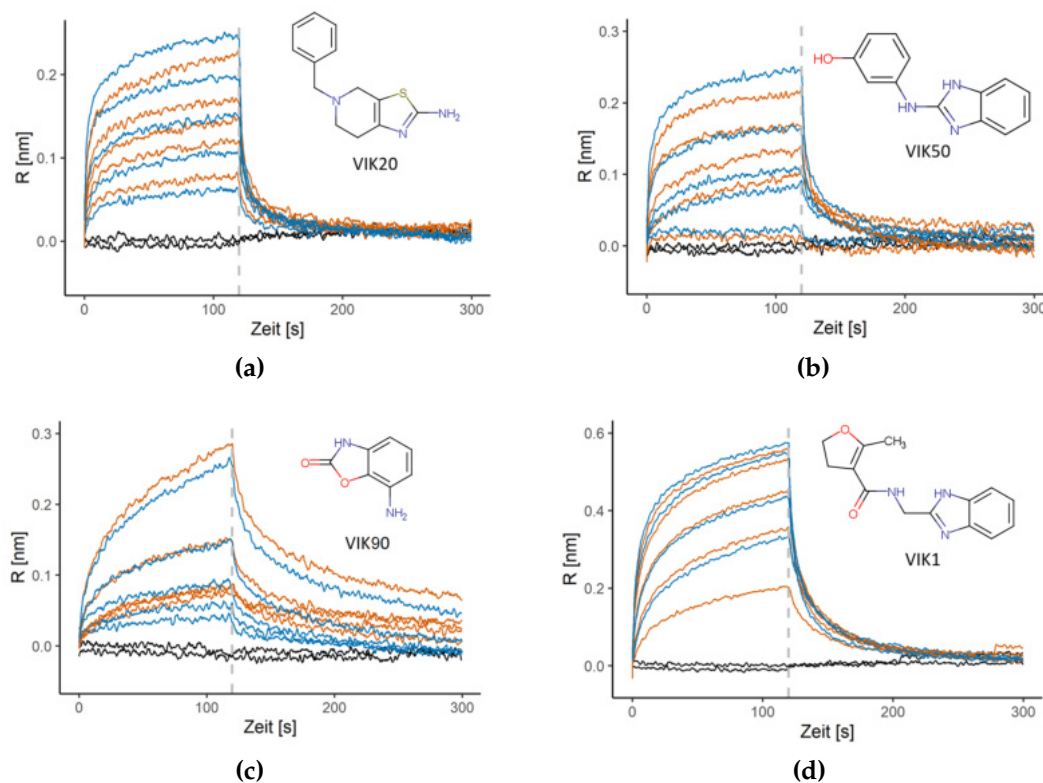


Abbildung 6.40: Vier Sensorgramme von Fragmenten die eine Konzentrationsabhängigkeit der Verschiebungen im p97-ND1 Screening zeigten

6.2.3.4 Initiales-Screening mit der isolierten N-Domäne von p97

Zum Zeitpunkt der Durchführung des initialen-Screenings mit p97-N lag, wie bereits erwähnt, keine valide Positivprobe für die isolierte N-Domäne vor. XN wäre hierfür sicherlich ein geeigneter Kandidat gewesen, allerdings wurden die XN BLI-Messungen erst nach den Fragmentscreenings durchgeführt. Auf Grundlage der Messungen mit der ND1-Domäne wurden die gleichen Assaybedingungen für das initiale-Screening mit der N-Domäne übernommen. Die Normalisierung der doppelreferenzierten Daten wurde nur über die Verschiebung der Beladung der Sensoren durchgeführt (siehe Kapitel 5.5.4.2). Zeitabhängige Effekte, wie Verlust oder Denaturierung von p97-N an der Sensoroberfläche, konnten hiermit allerdings nicht korrigiert werden. Bevor die Normalisierung durchgeführt wurde, wurden besonders auffällig hohe bzw. negative Signale anhand ihrer Sensorgramme visuell kontrolliert. Dabei stellten sich nahezu alle diese Fragmente als Artefakte heraus, die aus dem Datensatz entfernt wurden. Von 651 blieben 638 Fragmente für die weitere Auswertung übrig. Abbildung 6.41 zeigt die Übersicht aller doppelreferenzierten Verschiebungen des p97-N Screenings, vor (a) und nach (b) dieser visuellen Bereinigung.

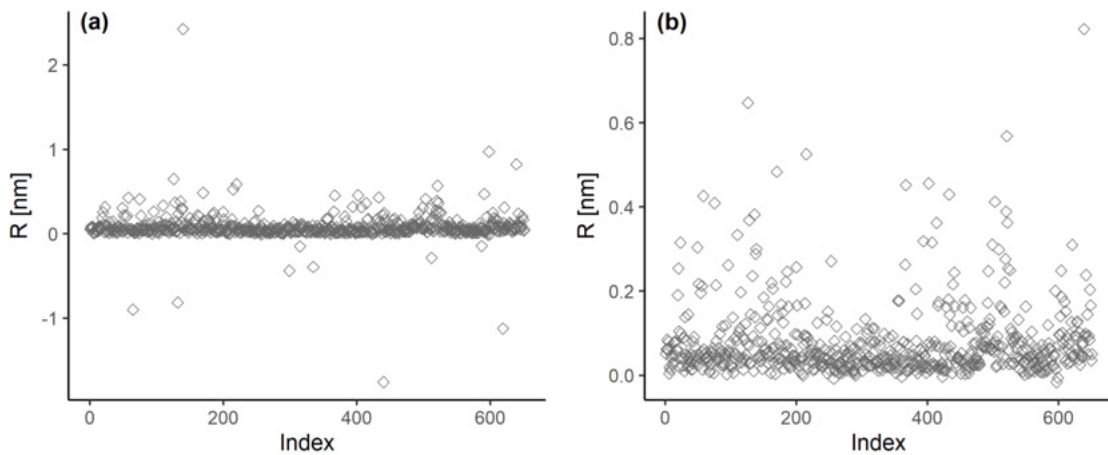


Abbildung 6.41: Doppelreferenzierte Daten aller Fragmente des initialen-Screenings mit p97-N

(a) Zu sehen sind die Daten vor der visuellen Kontrolle mit Fragmenten die auffällig hohen und negativen Verschiebungen zeigten. (b) Die Daten nach der Bereinigung der Artefakte.

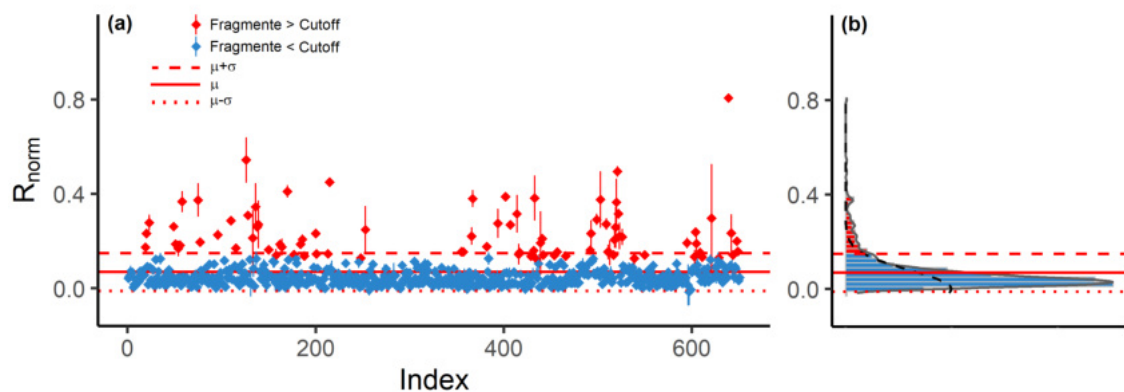


Abbildung 6.42: Übersicht des Fragmentscreenings mit p97-N

(a) Zu sehen sind die normierten Verschiebungen der Fragmente. In Rot sind diejenigen Fragmente gezeigt die oberhalb des gesetzten Grenzwerts von $\mu + \sigma = 0.1491$ lagen. Alle in Blau gezeigten Fragmente wurden nicht weiter verfolgt.

(b) Häufigkeitsverteilung der Verschiebungen der Fragmente überlagert mit einer auf Basis der Daten berechneten Normalverteilung in Schwarz. Deutlich zu sehen ist ein leichtes Abweichen (rot) der Daten von der berechneten Normalverteilung. Diese Abweichung spricht für eine Interaktion mit p97.

Nach Korrektur der Daten lag die mittlere normierte Verschiebung $\mu_{R_{norm}}$ bei 0.0690, mit einer Standardabweichung $\sigma_{R_{norm}}$ von 0.0802. Der eingesetzte Grenzwert für die weitere Selektion wurde somit auf $cutoff = \mu_{R_{norm}} + \sigma_{R_{norm}} = 0.1492$ gesetzt. Eine Fraktion von 66 (10.14 %) Fragmenten lag hier oberhalb des Grenzwerts. Innerhalb dieser Fraktion befanden sich die im initialen-Screening mit

p97-ND1 identifizierten Fragmente VIK20, VIK40 und VIK50, sowie das im virtuellen Screening gefundene Fragment SHP9, das ebenfalls unter gleichen Bedingungen untersucht wurde. Da die Bindung von SHP9 durch STD-NMR-Messungen an der isolierten N-Domäne bereits bestätigt war, wurde das Wiederfinden dieses Fragments als Bestätigung der Funktionalität des Assays angesehen.

Nach visueller Inspektion der Sensorgramme dieser 66 Fragmente wurden 36 ausgewählt. 6 weitere Fragmente wurden aufgrund ihrer Sensorgramme zusätzlich mit in Beracht gezogen, obwohl sie knapp unterhalb des gesetzten Grenzwertes lagen. In Summe blieben 42 (6.58 %) Fragmente für die folgenden Konzentrationsabhängigkeitsmessungen übrig.

6.2.3.5 Messung der Konzentrationsabhängigkeit ausgewählter Fragmente des initialen p97-N-Screenings

Die ausgewählten 42 Fragmente wurden auf die Konzentrationsabhängigkeit ihrer Verschiebungen geprüft und die erhaltenen Daten anschließend mit Hilfe eines 1:1-Langmuir-Modells ausgewertet. Aufgrund der so erhaltenen Affinitäten wurden sie in die in 5.5.4.4 genannten Kategorien eingeteilt. Hierbei fielen 22 Fragmente in die Kategorie **Binder**, 7 in die Kategorie **schwache Binder**, 7 in die Kategorie **unspezifische Binder** und 6 in die Kategorie **keine Binder**. Abbildung 6.44 zeigt jeweils zwei Beispielkurven für jede Kategorie aus den vermessenen Fragmenten¹⁴. Deutlich erkennbar sind hier sigmoidale Kurvenverläufe (Binder bzw. schwache Binder), lineare Verläufe (unspezifische Bindung) und solche, die keine sichtbare Abhängigkeit besitzen (keine Binder). Die Strukturen und Affinitäten sind in der zusammenfassenden Übersichtstabelle 6.17 der Fragmentscreenings mit p97 zu finden. In den Sensorgrammen fanden sich dabei die bereits im p97-ND1-Screening beobachteten Kurvenverläufe (Abbildung 6.43). Die meisten der Fragmente zeigten hier den bei TROLL6 (**a**) und TROLL7 (**b**) dargestellten Verlauf mit einem biphasischen Erscheinungsbild und annähernden Plateauphasen.

¹⁴Sämtliche Regressionskurven mit den dazugehörigen Sensogrammen der Kategorie **Binder** sind in Anhang F zu finden.

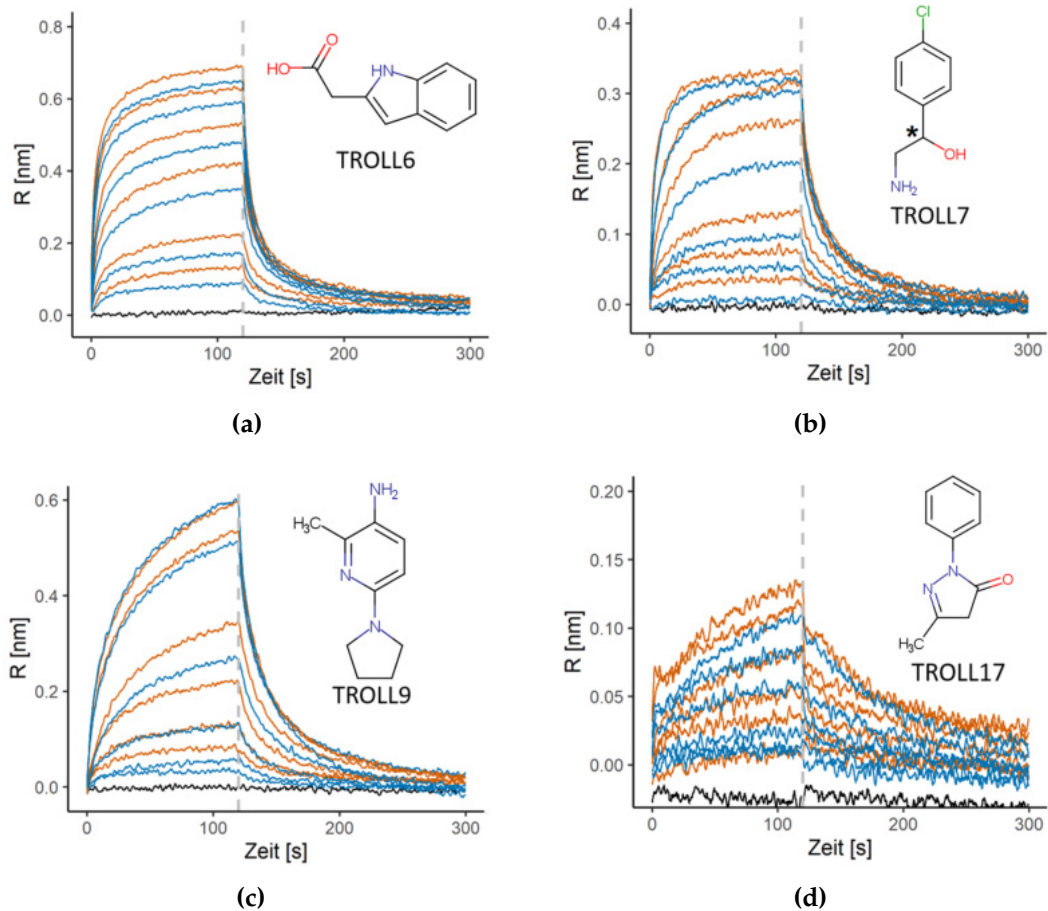
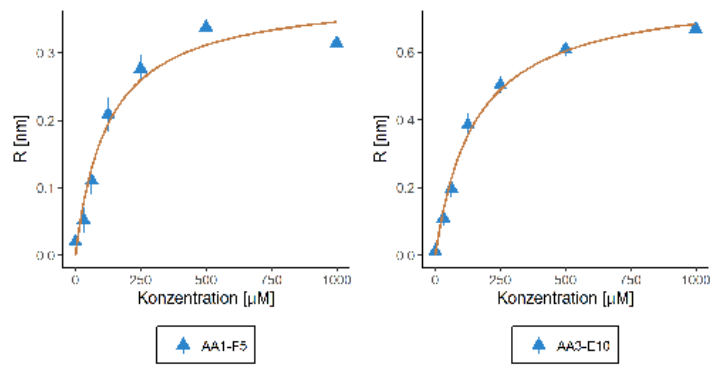
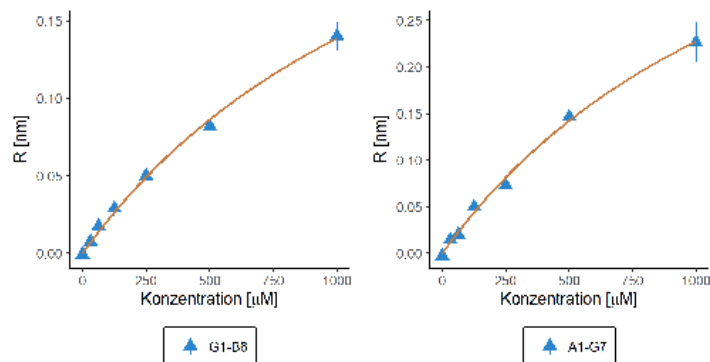


Abbildung 6.43: Vier Sensorgramme von Fragmenten die eine Konzentrationsabhängigkeit, der Verschiebungen im p97-N Screening zeigten

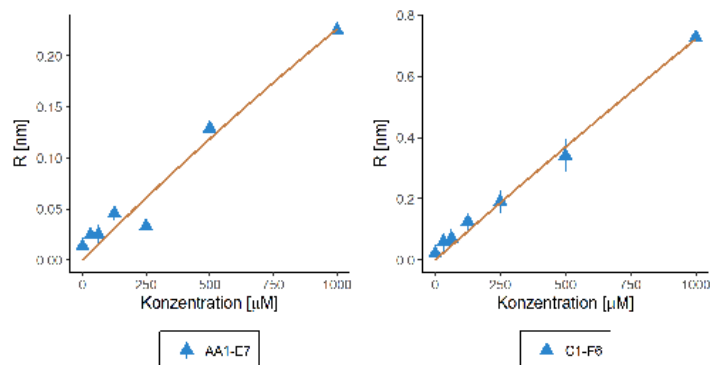
Die Höhe der Verschiebungen waren hier, aufgrund der kleineren Masse der N-Domäne und der höheren Beladung der Sensoren, wie zu erwarten höher als bei dem p97-ND1-Screening und beliefen sich zwischen 0.3 und 0.6 nm für die maximale Konzentration. Im Fall von besonders hohen Verschiebungen, wie bei dem hier gezeigten Fragment TROLL6, sind diese vermutlich auf die Bindung an mehrere Bindestellen von p97-N zurückzuführen, was bei dem hier gezeigten kleinen Fragment sicherlich nicht untypisch ist. Daneben fanden sich ebenfalls Sensorgramme, wie hier am Beispiel von TROLL9 (c) zu sehen, die auf eine unspezifische Bindung des Moleküls durch nicht-Erreichen einer Plateauphase hindeuteten. TROLL17 (d) fiel zusätzlich mit einer geringen Verschiebung auf. Alle Fragmenten, die laut Langmuir-Modell eine höhere Affinität als 700 μM aufwiesen, wurden erneut mit der ND1-Domäne vermessen.



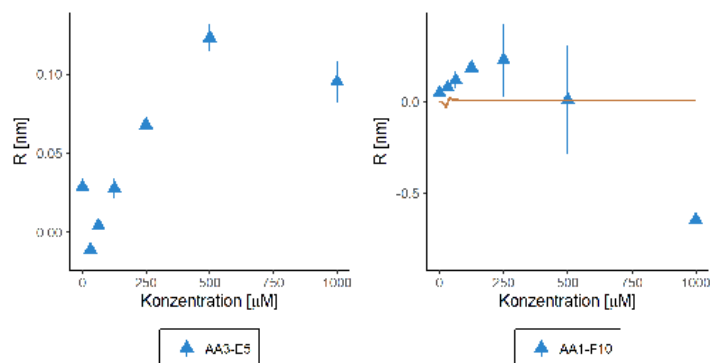
(a)



(b)



(c)



(d)

Abbildung 6.44: Langmuir-Modelle für die vier definierten Kategorien von Fragmenten

(a) Binder, (b) schwache Binder, (c) unspezifische Binder und (d) keine Binder.

6.2.3.6 Vermessung gefundener Treffer des p97-N-Screenings mit p97-ND1

19 Fragmente zeigten im Langmuir-Modell Affinitäten mit K_D -Werten unter 700 μM für die N-Domäne und wurden noch einmal mit der ND1-Domäne vermessen. Dies galt der Beantwortung der zwei folgenden Fragen:

- War die gefundene Bindung das Resultat der Immobilisierung von p97-N an den Sensor (Hintergrund: siehe XN Messungen 6.2.2.4)?
- Liegt die mögliche Bindestelle an einer biologisch irrelevanten Stelle, da sie dort sterisch durch die D1-Domäne maskiert ist?

Die Messungen hierzu wurden an einem K2 Octet Leihgerät der Firma Sartorius an der Universität Würzburg durchgeführt¹⁵. Alle ausgewählten Fragmente (mit Ausnahme von VIK20 und VIK40, für die bereits Messungen mit der ND1-Domäne aus Norwegen vorlagen) wurden mit der ND1-Domäne vermessen und die erhaltenen Daten mit Hilfe des Langmuir-Modells ausgewertet. Ein Blick auf die im Anhang G gezeigten Sensorgramme zeigt die zu erwartenden, deutlich geringeren Verschiebungen (im Mittel kleiner als 0.1 nm) als bei der isolierten N-Domäne. Die Qualität der erhaltenen Kurvenverläufe ist deutlich schlechter als bei den Messungen mit der N-Domäne. Sichtbar sind die genannten Probleme beispielsweise in den nicht sauber verlaufenden Kurven der Negativproben im Assaypuffer (schwarz). Acht Fragmente ließen eine Konzentrationsabhängigkeit in den Langmuir-Modellen erkennen. Drei Fragmente (TROLL1, TROLL13 und TROLL15) zeigten mit der ND1-Domäne kein Signal mehr und weitere sechs Fragmente keine auswertbaren Signale. Die bestimmten Affinitäten sind aufgrund der Messproblematik hier unter Vorbehalt zu sehen.

¹⁵Aufgrund der Corona Pandemie 2020 waren erneute Reisen nach Norwegen nicht mehr möglich. Um die Messungen dennoch weiter durchführen zu können wurde von Sartorius dankenswerterweise ein K2 Octet zur Verfügung gestellt. Das K2 Octet besitzt, im Gegensatz zu dem größeren Octet 96REDe, nur zwei Spektrometer. Neben dem Problem des viel geringeren Durchsatzes (ein Fragment gegenüber sechs Fragmenten pro Messungen), zeigten sich hier Probleme bei der Referenzierung der Messungen. Am größeren acht-Kanal Octet 96REDe Gerät wurden stets zwei Kanäle für Pufferkorrekturen verwendet. Dies hatte den Vorteil, dass bei ungewöhnlichen Kurvenverläufen einer Messung auf die andere Messung zurückgegriffen werden konnte. Am K2 war dies nicht möglich, sodass eine komplett neue Messung hätte durchgeführt werden müssen, was aufgrund der begrenzten Messzeit von knapp 3 Wochen nicht immer möglich war. Generell zeigten sich in den aufgenommenen Daten für die Pufferkorrektur deutlich stärkere Schwankungen und Abweichungen, als bei den UiB Messungen. Ursachen waren hier vermutlich neben den berichteten Problemen mit den SSA-Sensoren (Kapitel 5.5.2.1), die Tatsache, dass das K2 Octet nur die Funktion des Heizens und nicht, wie das 96REDe, die Funktion des Kühlens und Heizens der Probenplatte besitzt. Temperaturschwankungen besitzen einen starken Einfluss auf die Signalqualität, die bei den hier durchgeführten Messungen nahe der Detektionsgrenze besonders ins Gewicht fallen können. Aufgrund dieser nicht optimalen Messbedingungen wurden von einigen Fragmenten zwar Signale erhalten, diese konnten aber nicht ausgewertet werden.

In Abbildung 6.45 ist eine Gegenüberstellung der erhaltenen Affinitäten (gelb) mit der ND1-Domäne und der N-Domäne gezeigt (die absoluten Werte finden sich wieder in der Übersichtstabelle 6.17). Die eingezeichnete Diagonale entspricht einer gleichen Affinität für die N- und ND1-Domäne. Fragmente, die auf oder in der Nähe der Diagonalen liegen, lassen auf eine D1-Domänen-unabhängige Bindung schließen und sich dahingehend interpretieren, dass ihre Bindestelle innerhalb der N-Domäne zu finden sein müsste. Das bereits vorgestellte XN ist auf dieser Diagonalen zu finden. Fragmente rechts der Diagonalen zeigen eine höhere Affinität in Gegenwart der D1-Domäne. Dies kann durch einen positiven Einfluss der D1-Domäne (z.B. Bindung im Kontaktbereich beider Domänen) oder durch eine bessere Bindestelle in der D1-Domäne interpretiert werden. In Blau gezeigte Fragmente lieferten kein oder kein auswertbares Signal mit der ND1-Domäne. Durch die beschriebenen Probleme bei den Messungen kann hier nicht mit Sicherheit gesagt werden, ob dies das Resultat der Maskierung einer Bindestelle durch die D1-Domäne darstellt. Fragmente links der Diagonalen würden einen negativen Einfluss der D1-Domäne auf die Bindung zeigen. In Rot gezeigte Fragmente wurden nur mit der ND1-Domäne gemessen. Die in der gelben Ellipse befindlichen Fragmente liegen nahe der Diagonalen und lassen sich daher mit einer Bindung an die N-Domäne, ohne Einfluss der D1-Domäne, interpretieren. Dies würde auf die favorisierten Bindestellen in der SHP-Binderegion (SHP-I/SHP-II) zutreffen.

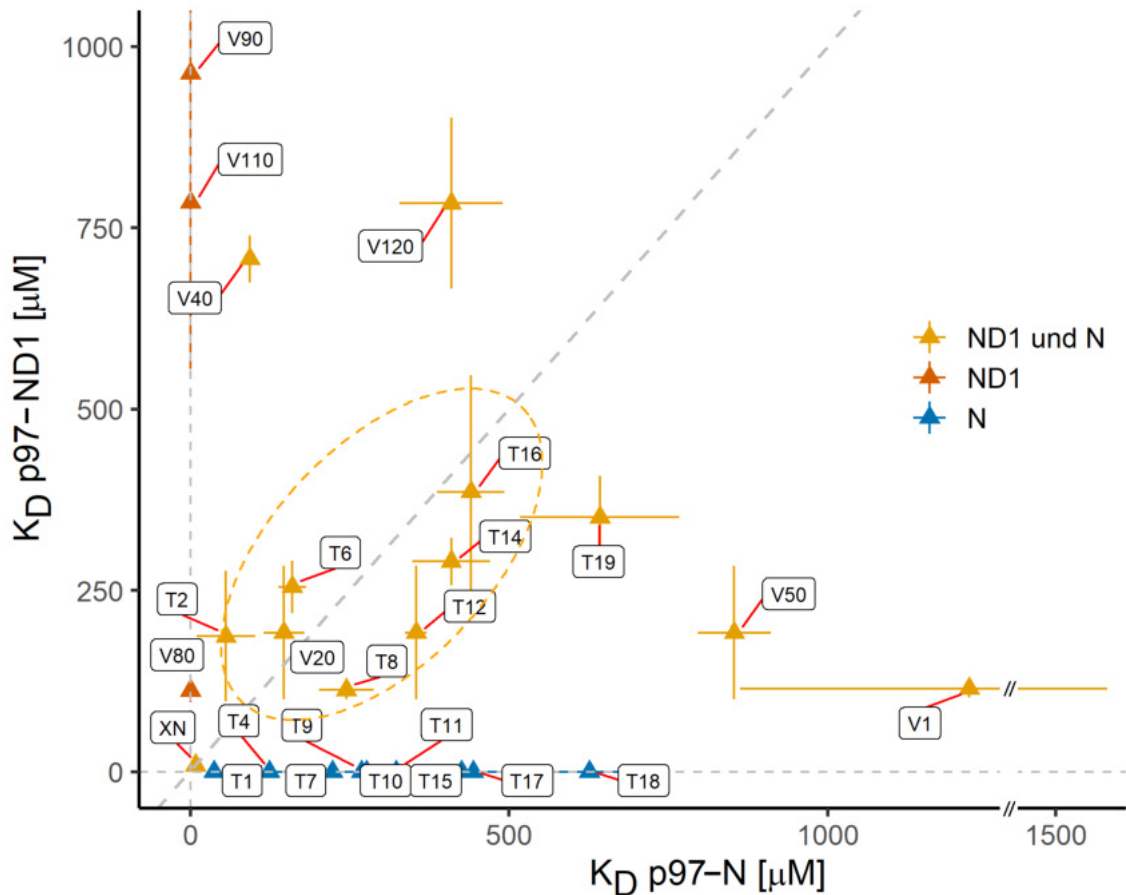


Abbildung 6.45: Gegenüberstellung der erhaltenen Affinitäten mit p97-N und p97-ND1

Gelbe Datenpunkte zeigt Fragmente, für die mit beiden Konstrukten eine Affinität bestimmt werden konnte. Für blaue (p97-N) bzw. rote Datenpunkte (p97-ND1) konnte nur eine Affinität für das jeweilige Konstrukt erhalten werden. Die Fehlerbalken zeigen die Fehler der Regression an. Die durch den Ursprung laufende Gerade zeigt den Bereich, für den beide Affinitäten gleich sind. In der gelben Ellipse sind diejenigen Fragmente markiert, die in der Nähe dieser Gerade zu finden sind und somit für beide Konstrukte eine fast gleiche Affinität zeigen. Anhand der gleichen Affinitäten für beide Konstrukte kann für diese Fragmente die Hypothese einer Bindung an die N-Domäne aufgestellt werden.

6.2.4 Analyse der Sensorgramme des Screenings mit der N-Domäne

6.2.4.1 Kurvenverläufe in den Sensorgrammen

Wie bereits erwähnt, ergibt eine Betrachtung der in Anlage F gezeigten Sensorgramme des p97-N Screenings, dass mehr oder weniger alle Kurven einen augenscheinlich biphasischen Verlauf zeigen. Um diesen gezielt zu analysieren und beurteilen zu können, wurden die jeweiligen Assoziationsphasen der unterschiedlichen gemessenen Konzentrationen der Fragmente einer exponentiellen Regression mit Hilfe einer einfachen Exponentialfunktion (siehe Gleichung 2.50, 1:1-Modell) bzw. einer heterogenen Exponentialfunktion (Gleichung 2.57, 2:1-Modell) unterzogen. Abbildung 6.47 zeigt dies beispielhaft an den Sensorgrammen von Fragment TROLL9 für zwei unterschiedliche Konzentrationen (1:1-Modell in Rot, 2:1-Modell in Blau). Für die Beurteilung der jeweiligen Regressionen wurde das jeweilige Bestimmtheitsmaß R^2 herangezogen. Für die geringe Konzentration von 31.25 μM liefern beide Modelle eine gute Vorhersage der experimentellen Daten, während die Daten für die Konzentration von 1000 μM eine deutliche Abweichung von einem 1:1-Modell zeigen und sich besser mit einem heterogenen Modell beschreiben lassen. Diesen Trend zeigten mehr oder weniger alle Fragmente in den Sensorgrammen.

Um die Abhängigkeiten von den zwei oben genannten Regressionen besser vergleichen zu können, wurde für jedes Fragment ein mittleres Bestimmtheitsmaß \overline{R}^2 für das jeweilige Modell bestimmt. Hierbei wurde der Mittelwert über alle Bestimmtheitsmaße aller Konzentrationen gebildet.

In Abbildung 6.50 sind die mittleren Bestimmtheitsmaße \overline{R}^2 für ein 1:1-Modell in Gelb und für ein 2:1-Modell in Blau für jedes Fragment, sowie ADP und XN zu sehen. Alle Fragmente zeigen hier eine bessere Übereinstimmung mit dem Kurvenverlauf eines 2:1-Modells als mit dem eines 1:1-Modells. Im Gegensatz hierzu zeigen die ADP und Xanthohumol Kurvenverläufe eine leicht bessere Übereinstimmung mit einem 1:1-Modell.

Bevor hier auf die möglichen Ursachen der vorgefundenen Heterogenität der Messkurven eingegangen wird, soll der Einfluss von möglichen Massentransport-Effekten auf die Messungen analysiert werden. Eine rein visuelle Inspektion der Kurven zeigte bei keinem Fragment das typische Profil, dass auf Massentransport-Effekte¹⁶ schließen lassen würde. Zur Bestätigung wurden die Daten mittels eines globalen 1:1-Modells mit Massentransport-Effekt in der Octet Analyse Software simuliert. Die erhaltenen Kurven zeigten deutlich, dass die Messungen nicht von diesem Effekt beeinflusst wurden (siehe Abbildung 6.46 am Beispiel von TROLL2).

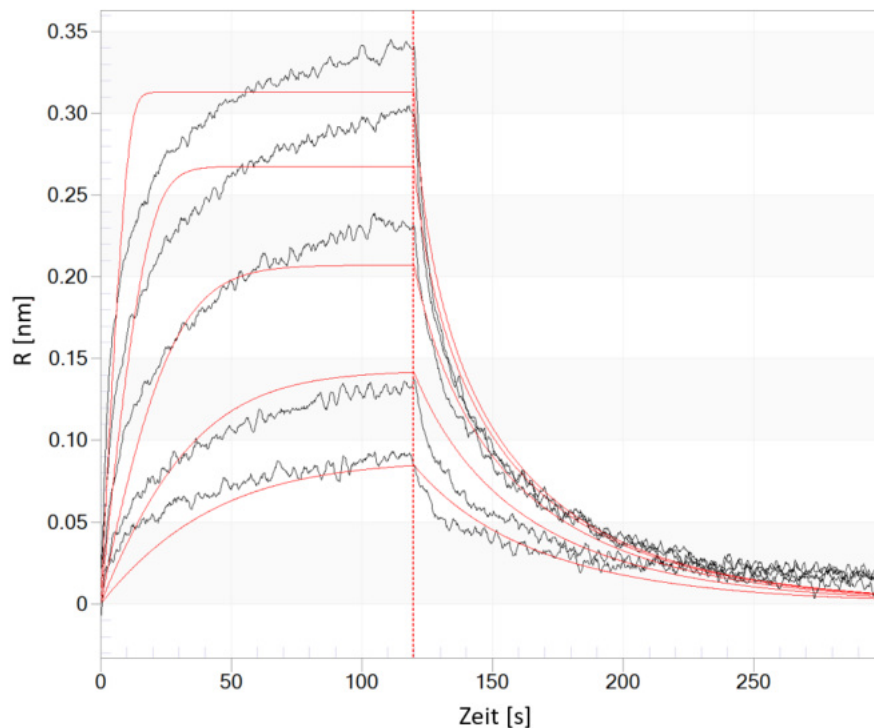


Abbildung 6.46: Globales 1:1-Modell mit Massentransport-Effekt anhand der Daten von TROLL2

In Schwarz sind die Verschiebungen für TROLL2 zu sehen. Die in Rot sind das Modell für Annahme eines Massentransport-Effektes zu sehen. Die Abbildung zeigt deutlich, dass die Messungen nicht von Massentransport-Effekten beeinflusst waren.

¹⁶Da die Bindung eines Liganden bei Vorliegen von Massentransport-Effekten nur noch von der Massentransfer-Rate k_m abhängig ist (siehe Kapitel 2.4.3.4), zeigen solche Sensorgramme einen fast linearen Anstieg.

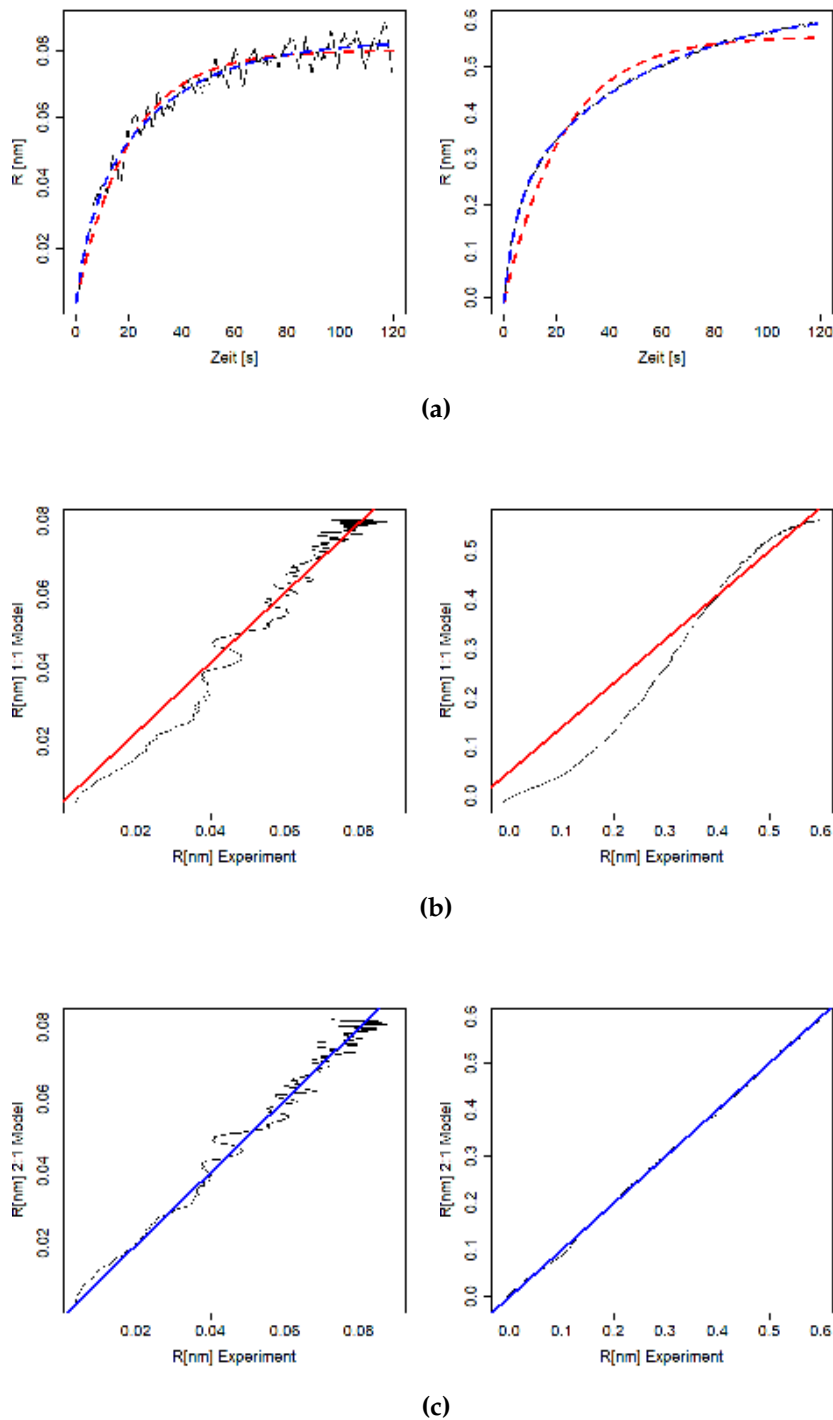


Abbildung 6.47: Abhängigkeit der Kurvenverläufe von einem 1:1- bzw. 2:1-Bindungsmodell am Beispiel von TROLL9

(a) Links: Die Modelle für ein 1:1-(rot) bzw. 2:1-(blau) Bindemodell überlagert mit den Daten von TROLL9 für die Konzentration von 31.25 μM . Rechts: das selbige für eine Konzentration von 1000 μM . (b) Vorhersage der experimentellen Daten anhand eines 1:1-Modells. Während der Kurvenverlauf der geringen Konzentration dem eines 1:1-Modells folgt, weicht er für die höhere Konzentration deutlich von diesem ab. Die Bestimmtheitsmaße wurden anhand der gezeigten Geraden berechnet. (c) Vorhersage der experimentellen Daten anhand eines 2:1-Modells. Beide Kurvenverläufe lassen sich über ein 2:1-Modell beschreiben.

Welche Ursachen können den vorgefundenen heterogenen Kurvenverläufen zugrunde liegen? Biphasische Verläufe können durch unterschiedliche Effekte [214, 299, 300] hervorgerufen werden, die in der Bindung an sich liegen, wie:

- die Bindung an mehrere Bindestellen;
- unspezifische Bindung;
- oder Konformationsänderungen während der Bindung;

oder auf Effekten des Assays beruhen, wie:

- Einflüsse der Immobilisierung;
- nicht optimale Assaybedingungen in Form des verwendeten Puffers, Einsatz von Detergenzien oder nicht optimaler pH-Bereich.

Eine einfache Schlussfolgerung ist hier also nicht möglich. Je mehr Terme ein Regressionsmodell enthält, desto besser lassen sich solche komplexen Kurven modellieren. Diese zusätzlichen Parameter stellen jedoch noch keinen realen Zusammenhang zwischen wirklichen biologisch ablaufenden Prozessen dar. Die Vorgehensweise in der Literatur [214, 299, 300] sagt hierzu, dass bei gefundener Heterogenität der erste Schritt eine Optimierung der Assaybedingungen darstellt (wie Beladungsdichten der Sensoren, Einsatz von Detergenzien, Pufferbedingungen, eingesetzte Konzentrationen etc.), mit dem Ziel, Kurven, die einem 1:1-Modell folgen, zu erhalten. Zeigen die Kurven dennoch einen biphasischen Verlauf, so muss die vermutete Ursache stets mit einer unabhängigen Messmethode bestätigt werden.

Im Kontext der hier eingesetzten Fragmente zeigen die Kurven meistens ein biphasisches Erscheinungsbild. Durch die sehr niedrigen Affinitäten und die daher hohen eingesetzten Konzentrationen, werden die Signale hier schnell von unspezifischen Bindungen beeinflusst. Zudem können die noch recht unspezifischen Fragmente unterschiedliche Bindeposen in mehreren Bindetaschen oder innerhalb einer Binderegion einnehmen, wobei jede der Posen ihre eigene Affinität aufweist, die in Summe nicht mit dem Kurvenverlauf eines 1:1-Modell erklärt werden können. Dennoch bietet sich das einfach handhabbare 1:1-Modell zur ersten groben Abschätzung der Affinitäten an. Die hier in einem Screening erhaltenen ersten Daten der Fragmente stellen sicherlich nur erste Anhaltspunkte dar und sollten im weiteren Verlauf des Projektes noch einmal gezielt für vielversprechende Fragmente individuell gemessen und optimiert werden. Eine solche Optimierung ist auch vor dem Hintergrund einer kinetischen Auswertung von einzelnen Fragmenten unabdingbar.

6.2.4.2 Abhängigkeiten von einem 1:1-Bindemodell

Der eben analysierte reine Kurvenverlauf der einzelnen Konzentrationen bestätigt alleine noch kein 1:1-Bindungsmodell. Hierzu bedarf es gemäß Gleichung 2.52 der Abhängigkeit der empirischen Geschwindigkeitskonstanten von der eingesetzten Konzentration des Fragments $k_{obs} \propto [L]$. Daher wurden die aus der exponentiellen Regression erhaltenen Werte für k_{obs} auf ihre Abhängigkeit von der eingesetzten Konzentration durch eine lineare Regression mit dem Programm Origin geprüft. Die hier durchgeführte Analyse über den linearen Zusammenhang zwischen k_{obs} und der Konzentration ist sehr anfällig für komplexere Bindeverhältnisse [301]. Aufgrund der Heterogenität der Bindekurven ist eine solche Analyse mit Vorbehalt zu sehen und dient nur einer zusätzlichen Einschätzung der erhaltenen Daten des Screenings.

In Anlage H sind alle erhaltenen Regressionsgeraden der Fragmente zu finden. Diese lassen nur bei wenigen einen linearen Zusammenhang erkennen. Hierzu zählen die Fragmente TROLL2, TROLL13 und VIK20. TROLL6, TROLL7, TROLL8, TROLL11, TROLL12, TROLL14, TROLL18 und VIK40 die für einen Teil der Konzentrationen lineare Zusammenhänge zeigen. Hierbei wurden Werte für die höchsten Konzentrationen oder kleinsten Konzentrationen, die augenscheinlich Ausreißer darstellen, nicht in die Regressionen mit einbezogen. Über die erhaltenen Geraden konnten die kinetischen Geschwindigkeitskonstanten k_a und k_d erhalten und dadurch die Affinität K_D auf Basis der Kinetik berechnet werden. In Abbildung 6.48 sind diese den Affinitäten aus dem Langmuir-Modell gegenübergestellt. Hier zeigt sich, dass ein Teil der Fragmente (innerhalb der gelben Ellipse: TROLL2, 6, 7, 8, 11, 12 und VIK20) eine ähnliche Größenordnung für die Affinitäten zeigt, wie sie auf Basis des Langmuir-Modells berechnet wurde.

Für die genannten Fragmente kann dies als eine weitere Bestätigung der Bindung gesehen werden. Alle anderen Fragmente zeigen in ihren Affinitäten teils erhebliche Fehlerbereiche und sind daher als nicht aussagekräftig einzustufen. Interessanterweise finden sich hier fast die gleichen Fragmente wieder, die bereits bei der Gegenüberstellung der Affinitäten für die N- und ND1-Domäne identifiziert wurden. Die erhaltenen Bestimmtheitsmaße R^2 der linearen Regression aller Fragmente, sowie von ADP und XN sind in Abbildung 6.50 gezeigt.

Es ist hier zu beachten, dass die aus dem Screening erhaltenen Messwerte nur mittels Zweipunktmessungen mit einem Sensor gewonnen wurden. Wie die analog durchgeführte Analyse der XN-Daten zeigte, sind die Messungen einer recht großen Streuung unterworfen, weswegen eine größere Anzahl an Messungen mit unterschiedlichen Sensoren die Aussagekraft dieser kinetischen Analyse deutlich verbessern würde.

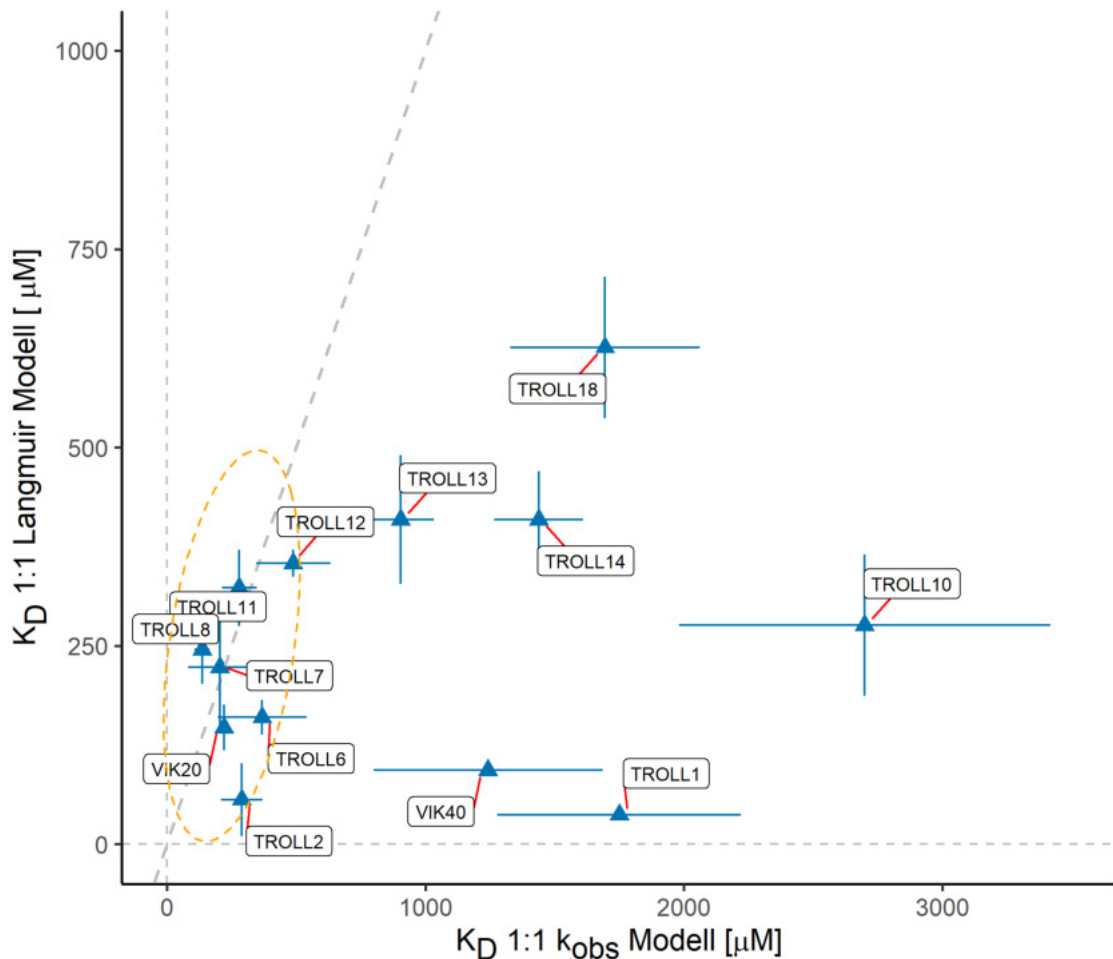


Abbildung 6.48: Gegenüberstellung der kinetischen Affinitäten (basierend auf k_{obs}) und der im Langmuir-Modell erhaltenen Affinitäten des p97-N Screenings
Die Diagonale zeigt den gleichen Wert für beide Affinitäten an. Für die Fragmente innerhalb der gelben Ellipse liegen die Affinitäten, abgeschätzt über das chemische Gleichgewicht, in der gleichen Größenordnung, wie die über die Kinetik bestimmten Werte.

6.2.4.3 Analyse der Plateauphasen der Sensorgramme

Die oben durchgeführte Auswertung durch ein 1:1-Langmuir-Modell setzt die Bedingung des Erreichens eines chemischen Gleichgewichts der Bindung zwischen einem Liganden und dessen Zielprotein voraus. Im Sensorgramm wird dies durch Ausbildung einer Plateauphase sichtbar. Ein nicht Ausbilden des chemischen Gleichgewichts führt zwangsläufig zu stark fehlerbehafteten Affinitäten auf Basis des Langmuir-Modells. Um das Erreichen der Plateauphase genauer abschätzen zu können und somit eine Aussage über die Belastbarkeit der Affinitätswerte zu erhalten, wurde die im Methodenteil 5.5.4.5 genannte Vorgehensweise auf die Sensorgramme angewendet.

ADP zeigte eine Steigung $m_{plateau}$ von $-1.84 \cdot 10^{-5} \text{ nm s}^{-1}$ und lag somit nahe dem Idealwert von Null. Abbildung 6.49 zeigt die Ergebnisse der Auswertung für XN, das in den Sensorgrammen eine eindeutige Plateauphase ausbildete, und für TROLL9, dessen Sensorgramme gerade bei hohen Konzentrationen keine Plateaus mehr erreichten. XN zeigte eine Steigung $m_{plateau}$ von $-2.38 \cdot 10^{-5} \text{ nm s}^{-1}$ und lag somit im gleichen Bereich wie ADP. TROLL9 zeigte dagegen eine Steigung $m_{plateau}$ von $3.51 \cdot 10^{-4} \text{ nm s}^{-1}$, und lag vom Betrag der Werte um eine Zehnerpotenz höher als ADP. Anhand der Steigungen $m_{plateau}$ lässt sich somit qualitativ (im Vergleich zu ADP) das Erreichen einer Plateauphase abschätzen. Alle errechneten Steigungen $m_{plateau}$ der Fragmente sind ebenfalls in Abbildung 6.50 gezeigt.

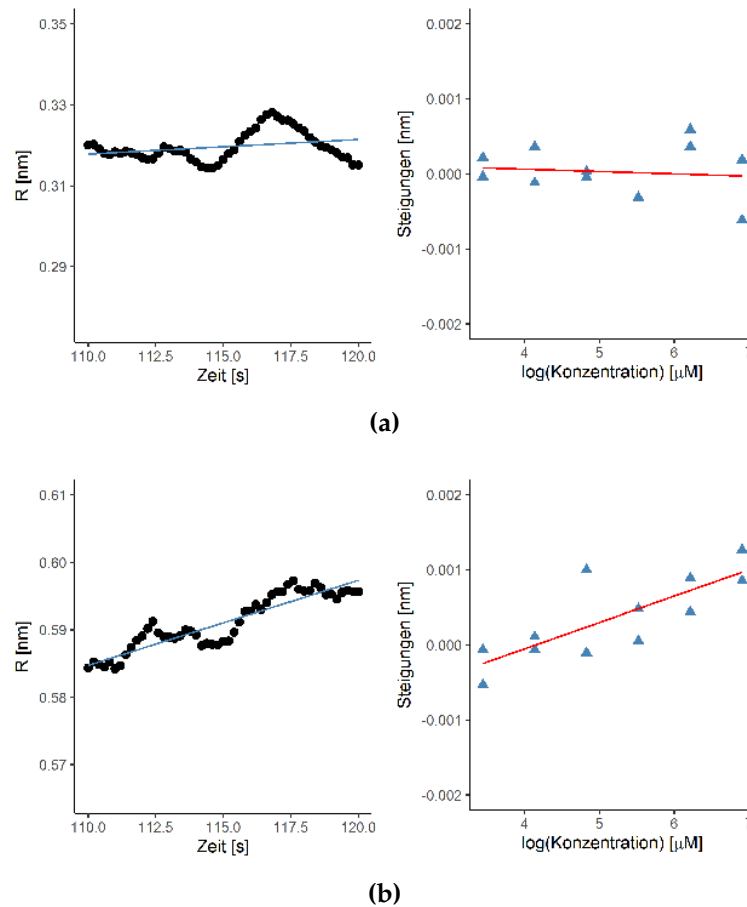


Abbildung 6.49: Auswertung der Plateauphasen

(a) zeigt die lineare Regression für eine Konzentration von XN. Hierbei fließen die Datenpunkte der letzten zehn Sekunden der Assoziationsphase mit ein. Die Gerade verläuft nahezu horizontal und zeigt somit das Ausbilden eines chemischen Gleichgewichts an. Daneben sind die aus den einzelnen Assoziationsphasen erhaltenen Steigungen $m_{Konz.}$ gegen die Konzentration aufgetragen. Wieder zeigt die erhaltene Gerade einen fast horizontalen Verlauf. Die Steigung dieser Geraden wird in dieser Arbeit als $m_{plateau}$ bezeichnet.

(b) zeigt den gleichen Sachverhalt für TROLL9. Die Gerade der Einzelkonzentration zeigt noch einen deutlichen Anstieg der Verschiebung. Die so erhaltene Gerade für alle Konzentrationen zeigt ebenfalls einen deutlichen Anstieg und somit größeren $m_{plateau}$ -Wert. Da die Steigungen $m_{Konz.}$ gerade für die hohen Konzentrationen zunehmen, kann hier von einer unspezifischen Bindung des Fragments ausgegangen werden.

6.2.4.4 Zusammenfassung der Analysen der Sensorgramme des p97-N-Screenings

Abbildung 6.50 fasst noch einmal alle eben genannten Werte für alle Fragmente zusammen. Aufgrund der geschilderten Problematik bei den Messungen der ND1-Domäne am K2 Octet Gerät wurde auf die eben beschriebene Auswertung

der Sensorgramme verzichtet. Um eine abschließende Priorisierung der Fragmente zu erstellen, wurde ein Qualität-Score auf Basis dieser Daten aufgestellt. Der $Score_{BLI}$ wurde hierbei wie folgt berechnet:

$$Score_{BLI} = \frac{R_{1:1}^2}{R_{2:1}^2} \cdot R_{kobs}^2 - 100 \cdot |m_{plateau}| \quad (6.2)$$

$\frac{R_{1:1}^2}{R_{2:1}^2}$ stellt das Verhältnis der Bestimmtheitsmaße zwischen einem 1:1- und einem 2:1-Kurvenverlauf dar. Folgt die Kurve einem 1:1-Verlauf ist der entsprechende Term größer-gleich 1.0 und bei einem 2:1-Verlauf unter 1. Dieser Term wird multipliziert mit dem Bestimmtheitsmaß R_{kobs}^2 der linearen Auswertung auf Basis der empirischen Geschwindigkeitskonstanten. Liegt ein 1:1-Bindungsmodell vor, liegt er im Idealfall ebenfalls bei 1. Der Term $-100 \cdot |m_{plateau}|$ stellt eine Art „Strafterm“ für das nicht-Erreichen der Plateauphase dar. Um ihn in den gleichen Skalenbereich der Bestimmtheitsmaße zu bekommen, wird der Betrag von $m_{plateau}$ mit 100 multipliziert. Je weiter ein Wert unter 1.0 liegt, desto schlechter wird ein 1:1-Modell erfüllt. Tabelle 6.14 zeigt die erhaltenen Scorewerte für die einzelnen Fragmente. Die Werte zeigen eine auffallende Grenze zwischen dem Fragment VIK40 und TROLL15, weshalb alle Fragmente mit einem Score schlechter als VIK40 nicht weiter priorisiert wurden. Wieder liegen die bereits in den Analysen identifizierten Fragmente TROLL2, 6, 7, 8, 12 und VIK20 oberhalb der Grenze zwischen VIK40 und TROLL15.

Tabelle 6.14: Priorisierung der Fragmente anhand des $Score_{BLI}$

ID	$Score_{BLI}$	ID	$Score_{BLI}$
ADP	1.02	TROLL10	0.79
VIK20	0.94	TROLL6	0.75
TROLL14	0.93	TROLL1	0.73
TROLL13	0.92	VIK40	0.71
TROLL7	0.91	TROLL15	0.33
TROLL8	0.91	TROLL16	0.31
TROLL11	0.90	TROLL4	0.19
XN	0.89	TROLL17	0.17
TROLL18	0.89	TROLL19	-0.12
TROLL12	0.84	TROLL9	-0.15
TROLL2	0.83		

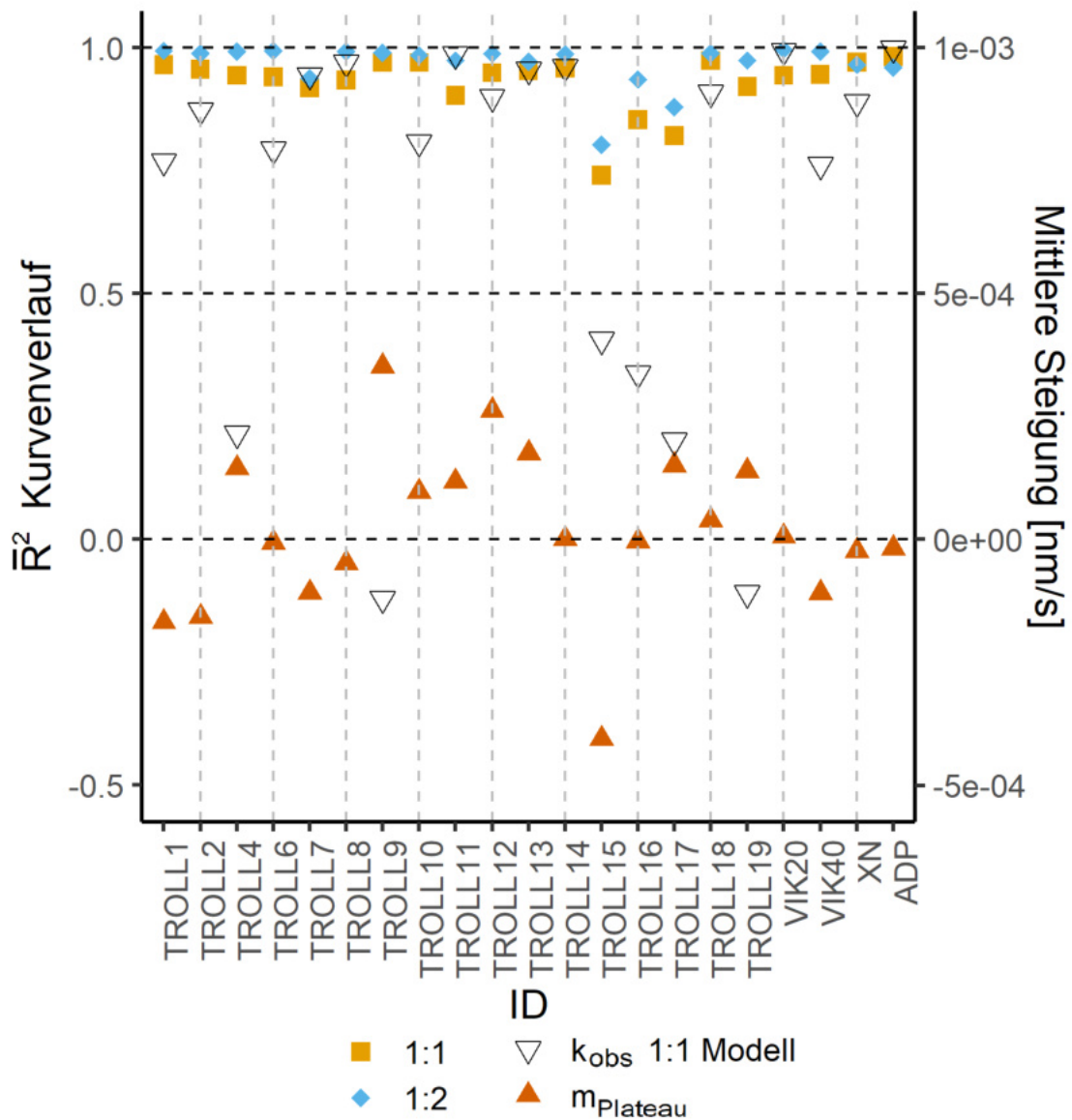


Abbildung 6.50: Zusammenfassung der Auswertung der Sensorgramme

Die Abbildung zeigt alle erhaltenen Werte der verschiedenen Analysen für jedes Fragment. ADP fungiert hier mit seinen eindeutigen Sensorgrammen als Positivkontrolle. Zu sehen sind die Bestimmtheitsmaße für den Kurvenverlauf eines 1:1-Modells (gelb), eines 2:1-Modells (blau), und für eine angenommene 1:1-Bindung (weiße Dreiecke), sowie die $m_{plateau}$ Werte (rot). Auf Grundlage dieser Werte wurde der Qualitäts-Score $Score_{BLI}$ berechnet.

6.2.5 Ergebnisse der orthogonalen STD-NMR-Messungen

Ziel des orthogonalen STD-NMR-Assays war die Identifizierung von möglichen falsch-positiven Treffern im BLI-Assay. Tabelle 6.15 und Abbildung 6.51 zeigen zusammenfassend die Ergebnisse der Messungen. Alle Fragmente der TROLL-Serie sowie einzelne Fragmente der VIK-Serie (VIK20, 40 und 50) die mit der N-Domäne ein Signal im BLI-Assay zeigten, wurden mit der ND1-Domäne vermessen. Die Auswertung der Messungen wurde, wie in Kapitel 5.6 beschrieben durchgeführt und die Werte für std_{mean}^{arom} und std_{max}^{arom} für jedes Fragment berechnet.

Wie auch bei STD-NMR-Messungen mit ID5 und SHP9 (siehe Kapitel 6.3), wurden teilweise hohe Signale bis über 70% erhalten. Im Mittel lagen die std_{mean}^{arom} -Signale bei 36.3 % und die std_{max}^{arom} -Signale bei 40.7 %. Für die Fragmente TROLL10, TROLL16 und TROLL19 konnten keine signifikanten STD-Signale erhalten werden. Die Fragmente TROLL9 und TROLL17 zeigten im Vergleich zum allgemeinen Mittel geringe Signale, gefolgt von den Fragmenten TROLL1, TROLL4 und TROLL15, die etwas höhere Signale zeigten. Für VIK20 konnten keine Signale für die Protonen des hydrierten Rings gemessen werden, da diese allesamt in den Bereich der Wasserunterdrückung fielen. Ein Hauptteil des Moleküls konnte also messtechnisch nicht erfasst werden, weswegen das scheinbar geringe STD-Signal hier mit Vorsicht zu betrachten ist und durchaus höher liegen könnte. Interessanterweise zeigten nun auch diejenigen Fragmente eindeutige Signale mit der ND1-Domäne, die bei den BLI-Messungen keine auswertbaren Signale ergaben. Hierzu gehören TROLL1, TROLL4, TROLL7, TROLL15 und TROLL18.

Auf Grundlage dieser Ergebnisse können die Fragmente TROLL9, TROLL10, TROLL16, TROLL17 und TROLL19 als mögliche Artefakte des BLI-Screenings angesehen werden. Die in Tabelle 6.14 genannten $Score_{BLI}$ -Werte dieser Fragmente zeigen für alle außer TROLL10 ebenfalls schlechte Werte. TROLL4 und TROLL15, die ebenfalls im Vergleich zu anderen Fragmenten schwächere STD-Signale zeigten, wurden auch über die $Score_{BLI}$ -Werte schlecht bewertet. Die STD-Messungen bestätigen demnach zum großen Teil die dort gemachten Beobachtungen.

Trotzdem bedeutet ein nicht vorhandenes Signal nicht zwangsläufig, dass das entsprechende Molekül nicht bindet. In einer bei Schiebel *et al.*[302] beschriebenen Fallstudie zeigten 71 Treffer eines kristallographischen Screenings, die mit sechs unterschiedlichen biophysikalischen Methoden vermessen wurden, keine

eindeutige Übereinstimmung. 44 % der Treffer des kristallographischen Screenings konnten sogar mit keiner biophysikalischen Methode detektiert werden, während die restlichen Fragmente nur eine schlechte Überschneidung der benutzen Methoden zeigten. Vor diesen Hintergrund sollten Fragmente, welche in einem einzelnen Assay kein Signal zeigen, nicht sofort verworfen werden. Daher wurde das Fragment TROLL10, das in den anderen Auswertungen vernünftige Ergebnisse zeigte, in die Betrachtungen in Kapitel 6.5 mit aufgenommen.

Tabelle 6.15: STD-NMR-Messungen aller Fragmente mit p97-ND1

ID	std_{mean}^{arom} -Signal [%]	std_{max}^{arom} -Signal [%]
TROLL1	18.0	22.5
TROLL2	36.0	37.7
TROLL4	20.5	24.2
TROLL6	66.5	72.5
TROLL7	34.5	37.3
TROLL8	33.1	41.0
TROLL9	10.5	10.8
TROLL10	-	-
TROLL11	37.8	43.7
TROLL12	44.5	47.9
TROLL13	33.0	38.7
TROLL14	55.4	67.4
TROLL15	25.2	32.1
TROLL16	-	-
TROLL17	11.3	12.5
TROLL18	54.3	57.4
TROLL19	-	-
VIK20	18.3	18.3
VIK40	43.6	51.4
VIK50	73.9	78.0
Mittelwert	36.3	40.7

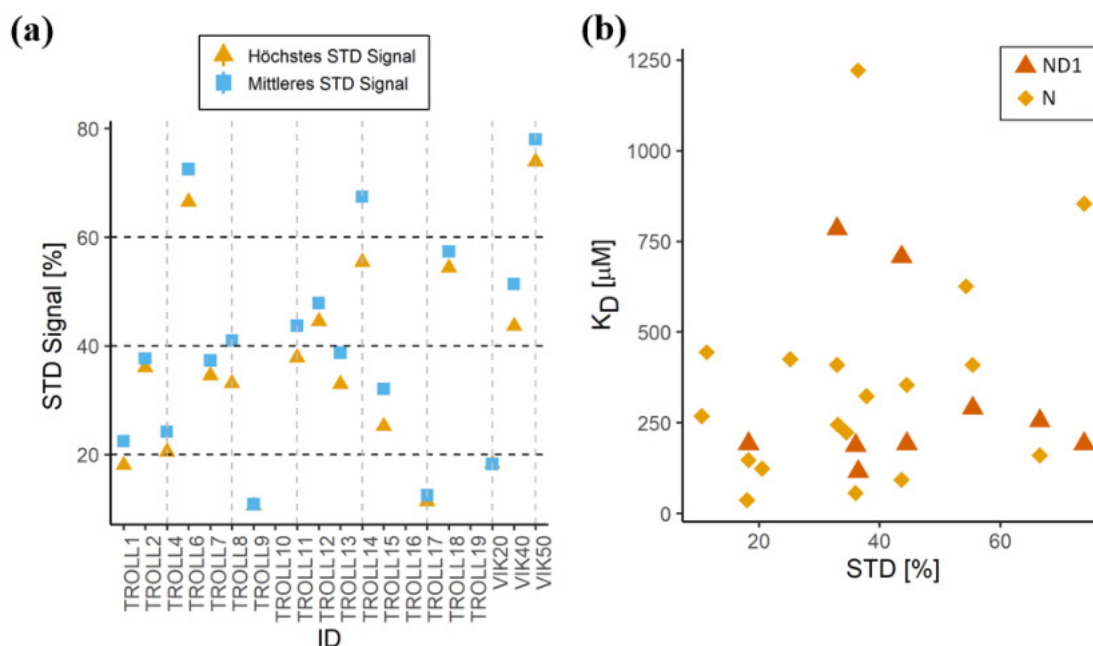


Abbildung 6.51: STD-NMR Messungen aller Treffer des N-Domänen-Screenings

(a) In Blau sind jeweils die std_{max}^{arom} -Signale und in Orange die std_{mean}^{arom} -Signale gezeigt. Auffallend sind die geringen Werte für TROLL9 und TROLL17. TROLL10, -16 und -19 zeigten keine signifikanten Signale.

(b) Gezeigt sind die std_{mean}^{arom} -Signale aufgetragen gegen die gemessenen Affinitäten (N-Domäne in Orange und ND1-Domäne in Rot). Es konnte keine Abhängigkeit der STD-Signalstärke von der Affinität festgestellt werden.

6.2.6 Zusammenfassung des Fragmentscreenings an p97

Nach der Vorstellung der Ergebnisse und erster Auswertungen des Fragment-screenings mit p97 sollen die Ergebnisse kurz zusammengefasst und ein Überblick über alle gefundenen Strukturen und Affinitäten gegeben werden. Generell zeigte das durchgeführte BLI-Fragmentscreening ähnliche Resultate bezüglich der beobachteten Signale und Streuung der Daten, wie bei in der Literatur berichtete BLI-Screenings [202, 273]. Tabelle 6.16 fasst die Treffer der initialen-Screenings, sowie die Konzentrationsabhängigkeitsmessungen zusammen. Sie zeigt, dass 22 Fragmente in beiden initialen-Screenings vertreten waren und für 13 Fragmente mit beiden Konstrukten eine Konzentrationsabhängigkeit gemessen werden konnte. Tabelle 6.17 gibt eine Übersicht aller im Fragmentscreening gefundenen Strukturen, die an p97-N oder p97-ND1 binden und für mindestens ein Konstrukt eine Affinität besser als 1000 µM aufweisen. Zusätzlich sind die Ergebnisse der STD-NMR-Messungen, sowie der $Score_{BLI}$ mit angegeben. Die IDs von Fragmenten, die für weitere Betrachtungen ausgewählt wurden, sind **fett** gedruckt.

Tabelle 6.16: Übersicht Fragmentscreening

„x“: Treffer;

„b.“: bestätigter Treffer;

„-“: kein Treffer bzw. nicht bestätigt;

„n.v.“: nicht vermessen

ID Screen	Interne ID	Treffer initial-Screening		Konzentrationsabhängigkeit		Treffer STD p97-ND1
		p97-ND1	p97-N	p97-ND1	p97-N	
1-C08	TROLL21	-	x	-	b.	n.v.
1-C09	TROLL25	x	x	n.v.	b.	n.v.
1-D02	VIK20	x	x	b.	b.	b.
1-G07	-	x	x	-	-	n.v.
1-G10	VIK80	x	-	b.	-	n.v.
1-H05	TROLL16	-	x	b.	b.	-
2-A02	TROLL15	x	x	-	b.	b.
2-B04	VIK90	x	-	b.	-	n.v.
2-H08	TROLL7	-	x	-	b.	b.
2-H09	TROLL18	x	x	-	b.	b.
2-H10	TROLL27	-	x	-	b.	n.v.
3-A02	TROLL4	-	x	-	b.	b.
3-A05	TROLL2	-	x	b.	b.	b.
3-B07	-	x	x	-	-	-
3-F09	-	x	x	-	-	-
3-F11	TROLL23	x	x	-	b.	n.v.
4-B05	TROLL20	-	x	-	b.	n.v.
4-D10	TROLL17	-	x	-	b.	b.
4-E09	VIK100	x	-	b.	-	-
5-C02	TROLL24	x	x	-	b.	n.v.
6-G06	TROLL11	-	x	-	b.	b.
7-B05	VIK40	x	x	b.	b.	b.
7-B06	VIK1	x	x	b.	b.	n.v.
7-B06	-	x	x	-	-	n.v.
7-E03	-	x	x	-	-	n.v.
7-F06	TROLL19	x	x	b.	b.	-
7-F11	VIK50	x	x	b.	b.	b.
7-F11	-	x	x	-	-	n.v.
7-G03	TROLL12	-	x	b.	b.	b.
7-G05	TROLL14	-	x	b.	b.	b.
8-C02	TROLL8	x	x	b.	b.	b.
8-D10	TROLL9	x	x	b.	b.	b.
8-D11	TROLL1	x	x	-	b.	b.
8-E03	TROLL13	x	x	b.	b.	b.
8-E06	TROLL10	x	x	-	b.	-
8-H10	TROLL6	-	x	b.	b.	b.
In beiden enthalten:		22		13		

Tabelle 6.17: Übersicht aller an p97 bindenden Fragmente

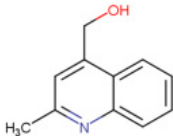
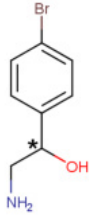
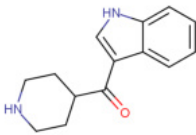
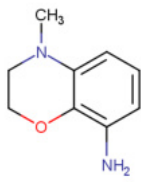
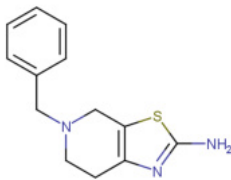
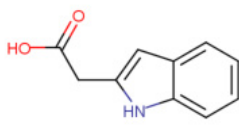
Gezeigt sind alle Fragmente, die mindestens für ein Konstrukt eine Affinität besser als 1000 μM zeigten. K_D Werte sind in μM angegeben. Fehler sind Standardfehler des Regressionsmodells ($n=2$).

„k.B.“: keine Bindung detektiert;

„n.a.“: Signale waren nicht mit Hilfe des 1:1-Langmuir-Modells auswertbar;

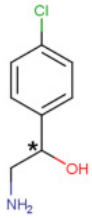
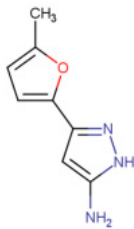
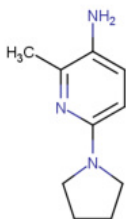
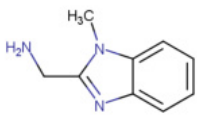
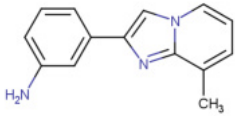
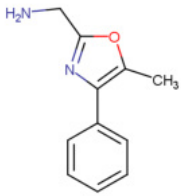
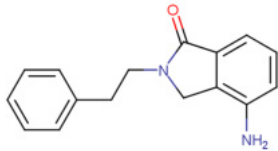
„n.v.“: nicht bekannt, da nicht mit p97-N vermessen;

fett gedruckte interne ID Nummern zeigen Fragmente an, die für die weitere Analyse ausgewählt wurden.

Struktur	ID	$K_D \pm \Delta_{Reg}$		std_{mean}^{arom}	Score _{BLI}
		p97-N	p97-ND1		
	ID _{Screen} : H1-D11 ID _{intern} : TROLL1	37 ± 3	n.a.	18.0	0.73
	ID _{Screen} : AA1-F5 ID _{intern} : TROLL2	56 ± 46	187 ± 90	36.0	0.83
	ID _{Screen} : G1-B5 ID _{intern} : VIK40	93 ± 4	707 ± 32	43.6	0.71
	ID _{Screen} : AA1-F2 ID _{intern} : TROLL4	124 ± 29	n.a.	20.5	0.19
	ID _{Screen} : A1-D2 ID _{intern} : VIK20	147 ± 32	192 ± 92	18.3	0.94
	ID _{Screen} : AA3-E10 ID _{intern} : TROLL6	160 ± 22	255 ± 36	66.5	0.75

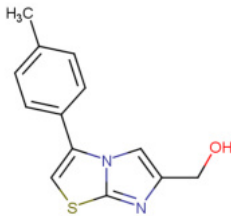
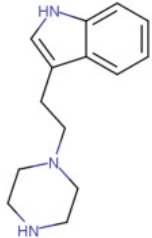
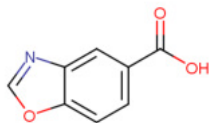
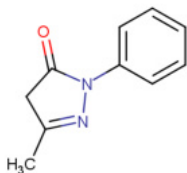
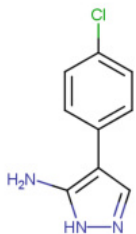
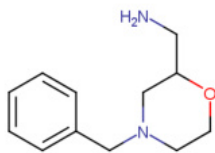
Fortsetzung auf nächster Seite

Tabelle 6.17 Fortsetzung von letzter Seite

Struktur	ID	p97-N	p97-ND1	std_{mean}^{atom}	Score _{BLI}
		$K_D \pm \Delta_{Reg}$			
	ID _{Screen} : AA1-E8 ID _{intern} : TROLL7	223 ± 83	n.a.	34.5	0.91
	ID _{Screen} : H1-C2 ID _{intern} : TROLL8	245 ± 43	113 ± 13	33.1	0.91
	ID _{Screen} : H1-D10 ID _{intern} : TROLL9	268 ± 65	n.a.	10.5	-0.15
	ID _{Screen} : H1-E6 ID _{intern} : TROLL10	276 ± 89	n.a.	k.B.	0.79
	ID _{Screen} : F1-G6 ID _{intern} : TROLL11	323 ± 48	726 ± 72	37.8	0.90
	ID _{Screen} : G1-G3 ID _{intern} : TROLL12	354 ± 17	192 ± 92	44.5	0.84
	ID _{Screen} : H1-E3 ID _{intern} : TROLL13	409 ± 81	784 ± 118	33.0	0.92

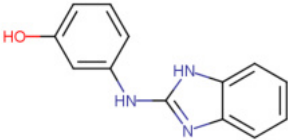
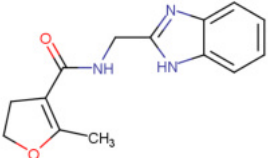
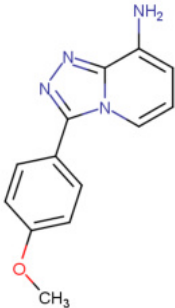
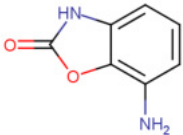

Fortsetzung auf nächster Seite

Tabelle 6.17 Fortsetzung von letzter Seite

Struktur	ID	p97-N	p97-ND1	std_{mean}^{arom}	Score _{BLI}
		$K_D \pm \Delta_{Reg}$			
	ID _{Screen} : G1-G5 ID _{intern} : TROLL14	409 ± 61	290 ± 33	55.4	0.93
	ID _{Screen} : AA1-D2 ID _{intern} : TROLL15	425 ± 31	n.a.	25.2	0.33
	ID _{Screen} : AA1-C5 ID _{intern} : TROLL16	440 ± 53	386 ± 161	k.B.	0.31
	ID _{Screen} : D1-D10 ID _{intern} : TROLL17	444 ± 56	k.B.	11.3	0.31
	ID _{Screen} : AA1-E9 ID _{intern} : TROLL18	626 ± 89	n.a.	54.3	0.89
	ID _{Screen} : G1-F6 ID _{intern} : TROLL19	642 ± 125	351 ± 57	k.B.	-0.15

Fortsetzung auf nächster Seite

Tabelle 6.17 Fortsetzung von letzter Seite

Struktur	ID	p97-N	p97-ND1	std_{mean}^{atom}	Score _{BLI}
		$K_D \pm \Delta_{Reg}$			
	ID _{Screen} : G1-F11 ID _{intern} : VIK50	853 ± 57	192 ± 92	73.9	-
	ID _{Screen} : G1-B6 ID _{intern} : VIK1	1221 ± 359	115 ± 13	n.v.	-
	ID _{Screen} : A1-G10 ID _{intern} : VIK80	n.v.	112 ± 16	n.v.	-
	ID _{Screen} : B1-B4 ID _{intern} : VIK90	n.v.	963 ± 408	n.v.	-
	ID _{Screen} : E1-B7 ID _{intern} : VIK100	n.v.	785 ± 149	n.v.	-

6.3 Biophysikalische Untersuchungen ausgewählter Fragmente aus dem virtuellen Screening

Wie in Abschnitt 6.1.3 beschrieben, wurden insgesamt zehn Fragmente aus dem virtuellen Screening für experimentelle Untersuchungen bei dem Anbieter Enamine bestellt. Als primärer Assay wurde das STD-NMR-Verfahren zur Detektion einer möglichen Bindung an die p97 N-Domäne verwendet. Als orthogonale Messmethode wurde ein BLI-Assay eingesetzt.

6.3.1 Etablierung des STD-NMR-Assays

Für die Etablierung des Assays wurde auf ein bei *Chimenti et al.* [61] berichtetes Fragment (ID5) zurückgegriffen, das in dieser Studie mit p97-ND1 ein STD-NMR-Signal zeigte und dessen Bindung an die isolierte N-Domäne mit einer Affinität von $241 \pm 4 \mu\text{M}$ unter Anwendung der SPR-Methode bestimmt wurde.

Für die Auswertung (auch aller späteren STD-NMR-Experimente) wurde der aromatische Bereich des Spektrums herangezogen (siehe Methodenteil 5.6). Abbildung 6.52 zeigt die Spektren des aromatischen Bereiches von ID5. Die Zuordnung der Protonen wurde unter Zuhilfenahme des *1H prediction Servers* auf www.nmrdb.org [303–305] durchgeführt. Für das Molekül konnte das Dublett des Protons 1 bei 7.295/7.285 ppm, sowie das Dublett des Protons 2 bei 6.989/6.979 ppm dem 1,3-Thiazolring zugeordnet werden. Der Phenylring erscheint als nicht näher aufgelöstes Multiplett zwischen 7.300 und 7.400 ppm. Die Methylengruppe sollte theoretisch im Bereich von 4 ppm zu finden sein, ist aber aufgrund der Wasserunterdrückung nicht zu sehen. Das Differenzspektrum zeigt ein deutliches Signal (std_{max}^{arom} -Signal mit S/N von 21.5 %) für alle aromatischen Protonen des Fragments. Das größte std_{max}^{arom} -Signal liegt bei 22.8 % und ist dem Proton 1 des 1,3-Thiazolringes bei 6.98 ppm zuzuordnen. Das std_{mean}^{arom} -Signal liegt bei 18.4 %.

Eine Negativprobe ohne p97-N gab Aufschluss darüber, ob mögliche beobachtete STD-Effekte tatsächlich auf der Bindung von ID5 beruhen. Das Differenzspektrum in Abbildung 6.52 (c) zeigt kein signifikantes Signal für ID5. Ein minimaler Ausschlag im Bereich des Phenylrings liegt mit einem S/N von 3.26 % überhalb des in 5.6 genannten Grenzwerts. Aufgrund der großen Differenz zwischen der

Probe (mit p97-N) und Negativkontrolle, wurde das Signal dennoch als nicht aussagekräftig gewertet. Das negative STD-Differenzsignal der Negativprobe bestätigte somit, dass das beobachtete STD-Signal ausschließlich von der Bindung des Fragmentes an p97-N herrührt.

Neben der Messung an der isolierten N-Domäne, wurden zusätzlich Messungen mit der ND1-Domäne durchgeführt. Die dabei erhaltenen STD-Signale lagen um das 3.4-fache (bezogen auf das std_{max}^{arom} -Signal) höher als bei den Messungen mit der N-Domäne. Das höhere STD-Signal lässt sich zum einen auf den besseren NOE-Effekt des größeren p97-ND1-Monomers (51.2 kDa), zum anderen auf mögliche weitere adressierte Bindestellen zurückführen. Das std_{max}^{arom} -Signal lag nun bei 77.1 % und war wieder Proton 1 des 1,3-Thiazolrings zuzuordnen. Das std_{mean}^{arom} -Signal lag bei 70.4 %. Das std_{max}^{arom} -Signal lag nahe dem bei [61] berichteten höchsten gemessenen STD-Signal von 87 %.

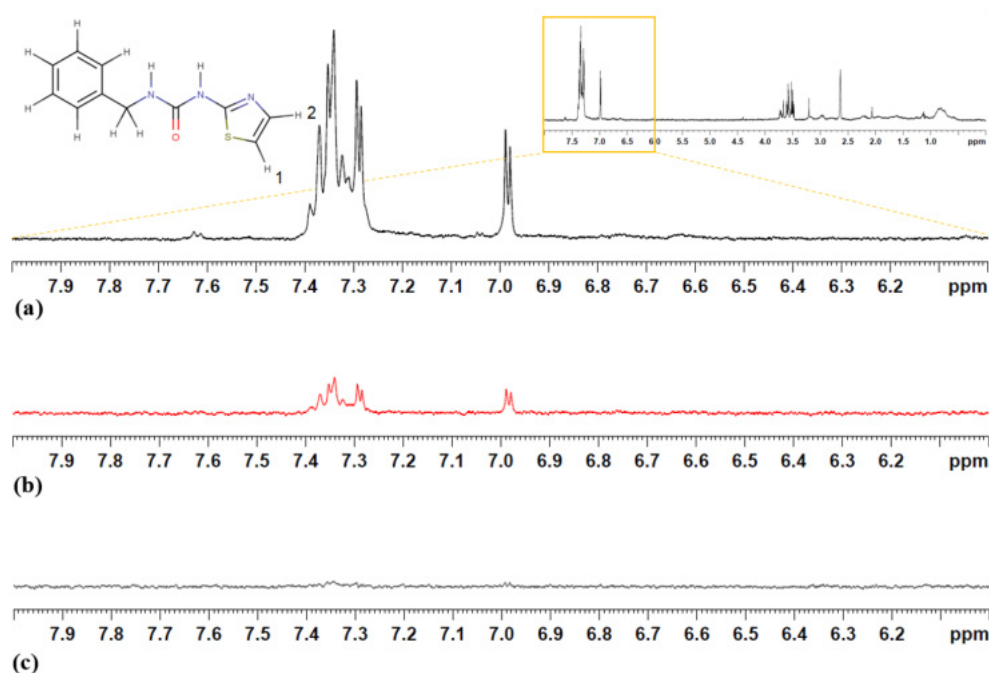


Abbildung 6.52: STD-NMR Messungen an ID5

Gezeigt ist der aromatische Bereich des Spektrums zwischen 6.0 und 8.0 ppm, der für die Auswertung der STD-Messungen herangezogen wurde. (a) zeigt das off-Resonanz-Spektrum, der p97-N (25 μ M) und ID5 (500 μ M) enthaltenden Probe. (b) zeigt das Differenzspektrum mit einem deutlichen Signal für die aromatischen Protonen von ID5 (c) zeigt das Differenzspektrum der Negativkontrolle. Diese enthielt nur das Fragment ID5 (500 μ M) unter sonst gleichen Bedingungen. In dem Spektrum ist kein signifikantes Signal sichtbar. Für weitere Erläuterungen siehe Text.

Um die Frage zu klären, ob es sich bei dem beobachteten Signal um eine unspezifische oder spezifische Bindung handelt, wurden die std_n^{rel} -Signale¹⁷ berechnet und den Strukturelementen des Fragments zugeordnet. Bei einer unspezifischen Interaktion sollten die std_n^{rel} -Signale in etwa gleiche Stärken zeigen, da kein bevorzugter Bindemodus des Moleküls vorliegt und im Mittel jedes Proton in etwa den gleichen Betrag an Magnetisierung über den NOE-Effekt bekommt. Bei einer spezifischen Bindung mit einem bevorzugten Bindemodus können die Protonen eines Moleküls durch ihre definierte Lage unterschiedliche Beträge an Magnetisierung erhalten und die std_n^{rel} -Signale können sich hierdurch voneinander unterscheiden.

Abbildung 6.53 zeigt die std_n^{rel} -Signale für die beiden unterschiedlichen Konstrukte p97-N und p97-ND1. Protonen 1 und 2 des 1,3-Thiazolringes zeigen vor allem für p97-N einen deutlich sichtbaren Unterschied in den std_n^{rel} -Signalen. Folglich bekommt Proton 2 im Durchschnitt mehr Magnetisierung über den NOE-Effekt als Proton 1. Dies kann mit einer definierten Orientierung des Thiazolringes an seiner Bindestelle erklärt werden und würde für eine spezifische Bindung sprechen. Ein Vergleich des std_n^{rel} -Signals des Phenylrings gegenüber dem Thiazolring zeigt mit beiden Konstrukten im Schnitt eine geringere Signal-Stärke. Auch dies kann für eine definierte Bindepose sprechen und wäre ein Indiz für eine spezifische Bindung von ID5 an p97. Interessanterweise fallen die Unterschiede bei den Messungen mit p97-ND1 geringer aus. Dies könnte durch zusätzliche Bindestellen an p97 hervorgerufen werden, wodurch die erhaltenen Signale die Summe der jeweiligen Bindungen darstellen. Zu beachten ist bei diesen Betrachtungen, dass die Unterschiede in den std_n^{rel} -Signalen auch auf unterschiedlichen Relaxationszeiten der Protonen beruhen können, die hier nicht im Detail untersucht wurden. Abschließend lässt sich sagen, dass anhand der hier durchgeführten STD-NMR-Messungen zwar die Bindung eines Moleküls detektiert, aber keine definitive Aussage gegeben werden kann, ob es sich bei der Bindung um eine spezifische oder unspezifische handelt. Es können hier nur Indizien beobachtet werden, die aber weiterer experimenteller Untersuchungen bedürfen. Die Etablierung des STD-NMR-Assays mit Hilfe des Fragments ID5 kann als erfolgreich eingestuft werden, mit ähnlichen STD-Effekten mit der ND1-Domäne, wie sie bei *Chimenti et al.* berichtet werden.

¹⁷Das relative STD-Signal eines Protons std_n^{rel} wurde über $std_n^{rel} = std_n^{arom} / std_{max}^{arom}$ mit dem STD-Signal std_n^{arom} des Protons und dem höchsten gefundenen STD-Signal std_{max}^{arom} berechnet.

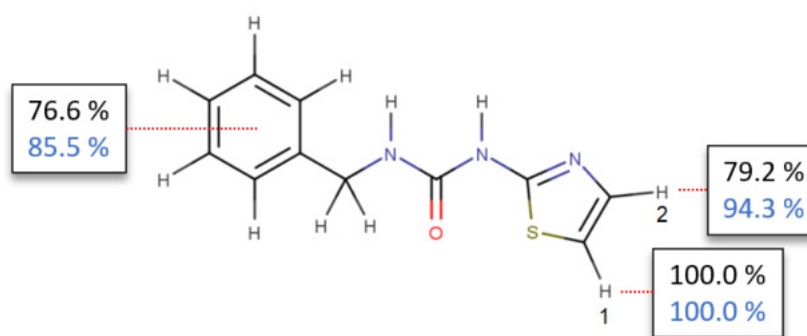


Abbildung 6.53: Relative std_n^{arom} -Signale des Fragments ID5

In Schwarz sind die std_n^{rel} -Signale, die mit p97-N erhalten wurden, gezeigt. Sie zeigen deutlich unterschiedliche Stärken für die beiden Protonen des Thiazolrings, sowie zwischen den Protonen des Phenyl- und des Thiazolrings. In Blau sind die std_n^{rel} -Signale, die mit p97-ND1 gemessen wurden, zu sehen. Der Effekt fällt hier deutlich geringer aus.

6.3.2 STD-NMR-Messungen ausgewählter Fragmente aus dem virtuellen Screening

Die zehn ausgewählten Fragmente aus dem virtuellen Screening wurden anschließend ebenfalls mit der N-Domäne vermessen. Von zehn Fragmenten zeigten vier ein signifikantes Signal im Differenzspektrum, wobei SHP9 das stärkste Signal aufwies (siehe Tabelle 6.18). Für die vier identifizierten Fragmente wurde ebenfalls eine Negativkontrolle ohne p97-N durchgeführt. SHP3 zeigte hier auch in der Negativkontrolle ein std_{mean}^{arom} -Signal von 7 %, dessen Ursprung nicht abschließend geklärt werden konnte. In der ersten Messung zeigte SHP1 bei 500 μM nur einen schwachen, nicht signifikanten STD-Effekt. Erst bei einer Konzentration von 1000 μM konnte ein signifikantes Signal erhalten werden. SHP6 zeigte ebenfalls nur einen schwachen STD-Effekt, der nur knapp oberhalb der Signifikanzgrenze lag (std_{max}^{arom} -Signal mit S/N von 3.71 %). SHP9 wurde daher als der beste „Treffer“ der STD-NMR-Messungen eingestuft.

Zur Bestätigung wurde SHP9 daraufhin erneut mit p97-ND1 vermessen. Ähnlich dem Fragment ID5 stiegen die STD-Signale ebenfalls um das 3.5 fache (bezogen auf das std_{max}^{arom} -Signal) an. Dies bestätigte, dass es sich bei dem massiven Anstieg des STD-Effekts vor allem um einen Konstrukt-abhängigen Effekt handelte. Das std_{max}^{arom} -Signal lag bei 67.1 %, das std_{mean}^{arom} -Signal bei 62.7 %.

Eine Betrachtung der std_n^{rel} -Signale der einzelnen Protonen zeigt vor allem für die Messungen mit der N-Domäne deutliche Unterschiede in den Signal-Stärken. Der 2-Chloro-Phenylring zeigt einen deutlichen Anstieg der Signale der Protonen von 5 über 4 nach 3 an (siehe Abbildung 6.54). Position 4 und 6 sind hierbei magnetisch äquivalent und können daher nicht unterschieden werden. Der Pyrimidinring zeigt ebenfalls unterschiedlich starke Signale für die Protonen 4' bzw. 3' und 5'. Diese Unterschiede in den Signal-Stärken der Protonen können ähnlich ID5 mit einer spezifischen Bindung von SHP9 an p97-N interpretiert werden. Interessanterweise verlieren sich diese Unterschiede in den Signal-Stärken bei der Verwendung des p97-ND1-Konstrukts fast vollständig. Analog zu ID5 kann dies ebenfalls mit zusätzlichen Bindungen an andere Regionen des größeren Konstruktes erklärt werden, die in Summe die Unterschiede in den std_n^{rel} -Signalen „verwischen“.

Tabelle 6.18: Ergebnisse der STD-NMR-Messungen ausgewählter Fragmente (500 μM) des virtuellen Screenings, die ein signifikantes Signal mit der N-Domäne (40 μM) von p97 zeigten. Zusätzlich sind die Ergebnisse der Messung mit der ND1-Domäne (30 μM) für das Fragment SHP9 (428 μM) gezeigt.

ID	std_{mean}^{arom} -Signal [%]	std_{max}^{arom} -Signal [%]
SHP1 ^a	2.2	7.5
SHP3	10.7	13.3
SHP6	7.1	7.8
SHP9	15.2 (p97-N)	19.3 (p97-N)
	62.7 (p97-ND1)	67.1 (p97-ND1)

^a Signal-Stärke bei einer Konzentration von 1000 μM .

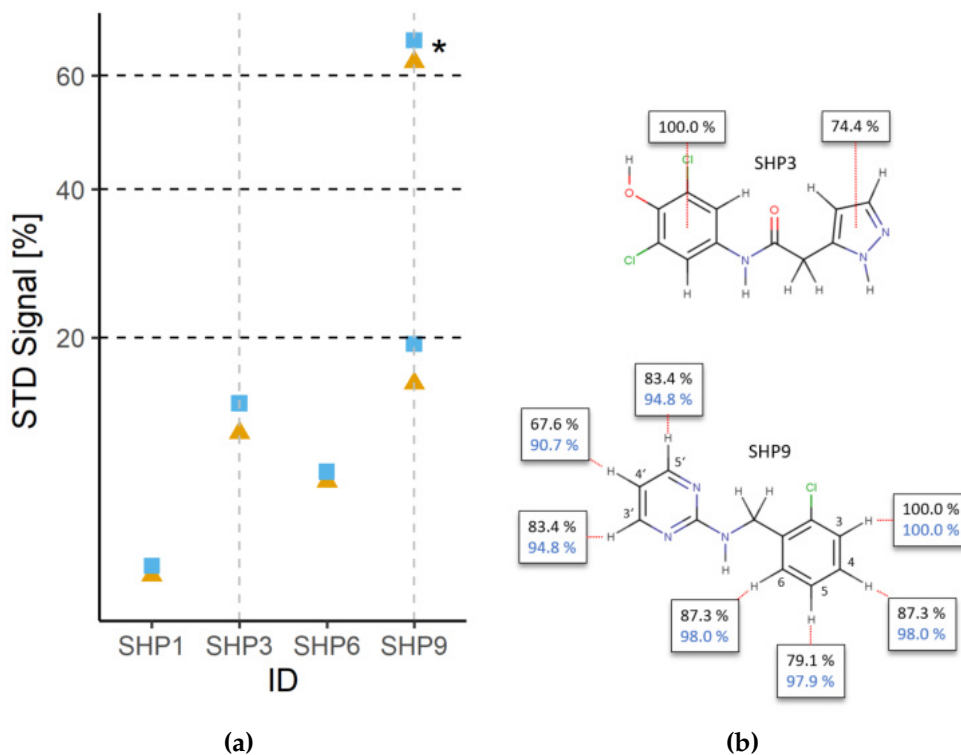


Abbildung 6.54: STD-NMR Messungen der Fragmente des virtuellen Screenings
(a) Ergebnisse der STD-NMR Messungen mit p97-N. In Blau sind die jeweils std_{max}^{arom} -Signale und in Orange die std_{mean}^{arom} -Signale gezeigt (* Messung mit p97-ND1).
(b) std_n^{rel} -Signale. In Schwarz sind die std_n^{rel} -Signale der Protonen mit der N-Domäne und in Blau mit der ND1-Domäne gezeigt.

6.3.3 BLI-Messungen ausgewählter Fragmente aus dem virtuellen Screening

Als orthogonale Messmethode wurde hier der BLI-Assay verwendet. Die zehn Fragmente wurden während des Fragmentscreenings an p97-ND1 analog den Molekülen aus der Fragmentbibliothek behandelt (siehe 6.2.3.1). Abbildung 6.55 zeigt die Ergebnisse der initialen Messung bei einer Konzentration von 500 μM . In Rot ist der in 6.2.3.1 genannte Grenzwert des initialen-Screenings mit der ND1-Domäne eingezeichnet. Wie auch bei den STD-Messungen lagen ID5, SHP9 wie auch SHP3 über dem Grenzwert, wobei letzteres Fragment mit einem Signal von 0.8 einen extrem hohen Wert zeigt. Die in den STD-Messungen nicht eindeutigen Fragmente SHP1 und SHP6 zeigen hier keine nennenswerten Verschiebungen. SHP2 lag hier über dem Grenzwert und wurde daher als „Treffer“ eingestuft.

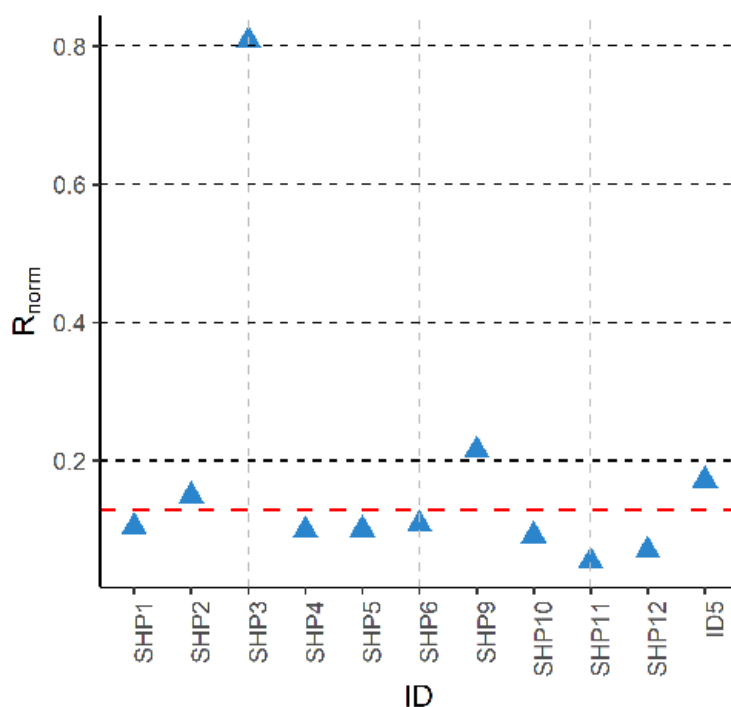


Abbildung 6.55: BLI Einpunktmessungen der Fragmente des virtuellen Screenings Gezeigt sind die Ergebnisse der BLI-Messungen mit p97-ND1 und den Fragmenten des virtuellen Screenings. Die Messungen wurden mit einer Ligand-Konzentration von 500 μM durchgeführt. In Rot ist der in 6.2.3.1 genannte Grenzwert für die Auswahl von Treffern eingezeichnet.

Die im Anschluss durchgeführten Messungen der Verdünnungsreihen ergaben keine Abhängigkeit für SHP2, und für SHP9 bzw. ID5 keine auswertbaren Signale. SHP3 zeigte als einzige Verbindung eine Konzentrationsabhängigkeit. Das Langmuir-Modell ergab für SHP3 eine Affinität von $790 \pm 304 \mu\text{M}$. SHP9 und ID5 wurden darauf hin erneut mit der N-Domäne vermessen. ID5 zeigte auch diesmal keine konzentrationsabhängigen Signale und konnte demnach im BLI-Assay nicht als Binder an p97-N bestätigt werden. SHP9 zeigte dagegen eine Konzentrationsabhängigkeit mit einem K_D -Wert von $208 \pm 48 \mu\text{M}$ im Langmuir-Modell. Die maximale Verschiebung war mit rund 0.1 nm auffallend schwächer als bei anderen bindenden Fragmenten mit vergleichbarem Molekulargewicht. Die Auswertung über die empirische Geschwindigkeitskonstante k_{obs} zeigte für das Fragment einen annähernd linearen Zusammenhang mit der eingesetzten Konzentration von SHP9, sodass hier eine 1:1-Bindung des Fragments angenommen werden kann. Der so über die Kinetik ermittelte K_D -Wert liegt um eine Potenz schlechter und zeigt einen großen Fehlerbereich (Abbildung 6.58). ID5 lieferte wieder nicht auswertbare Sensorgramme mit nicht erklärbaren Kurvenverläufen und konnte somit nicht im BLI-Assay als Ligand bestätigt werden.

Zusammenfassend lässt sich festhalten, dass anhand der STD-NMR-, sowie der BLI-Messungen, SHP9 als bindendes Fragment aus den virtuellen Screenings bestätigt werden konnte. SHP3 zeigt ebenfalls ein konzentrationsabhängiges Signal im BLI-Assay, lässt anhand der untypisch starken Verschiebungen in den Sensorgrammen aber keine endgültige Aussage zu, ob es sich bei der beobachteten Bindung doch eher um eine unspezifische handelt.

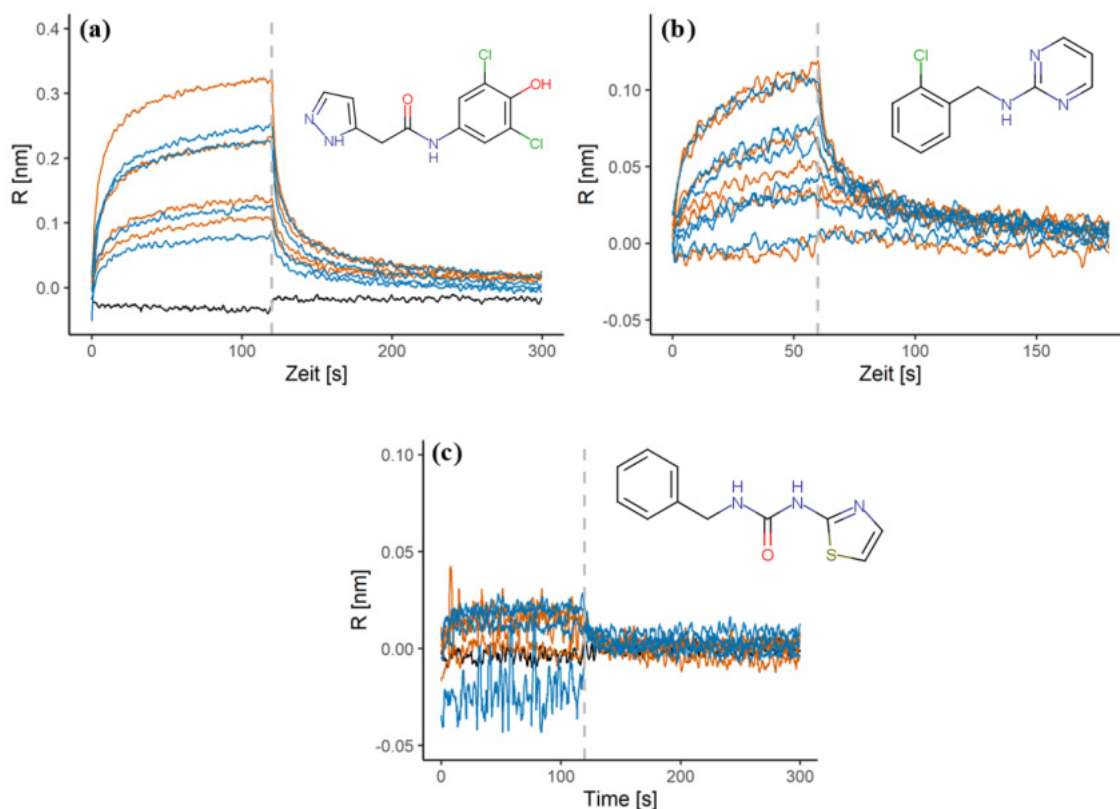
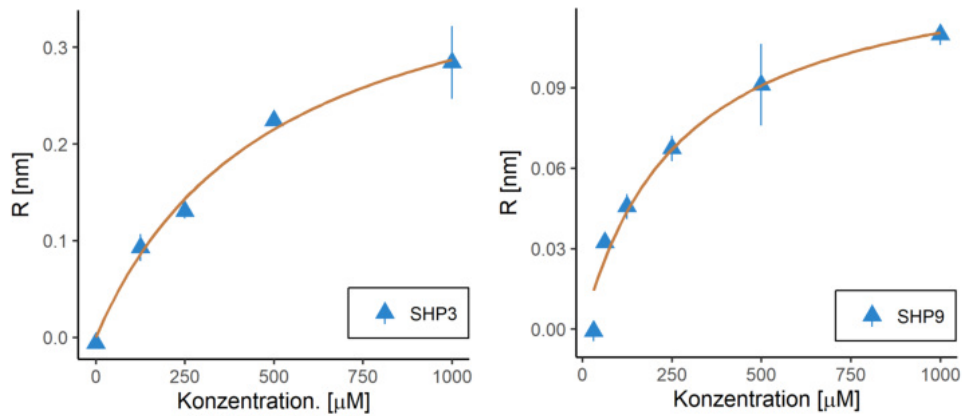


Abbildung 6.56: Sensorgramme der Fragmente des virtuellen Screenings

(a) Sensorgramm von SHP3 (p97-ND1) mit auffallend hohen Verschiebungen

(b) Sensorgramm von SHP9 (p97-N)

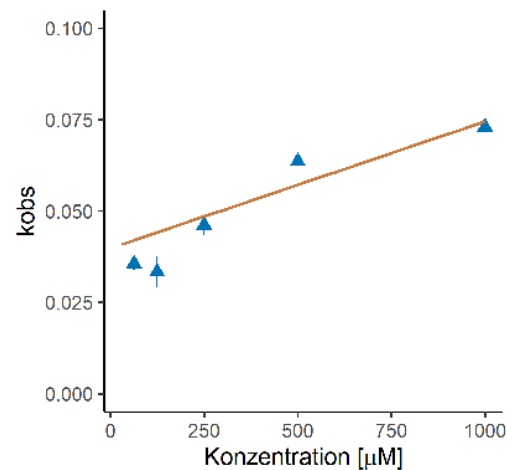
(c) Sensorgramm von ID5 (p97-ND1) mit negativen Verschiebungen und auffälligem Rauschen bei höheren Konzentrationen. Die Ursachen der negativen Signale sind bisher unbekannt.



ID	K_D [μM]	$\Delta_{Reg.}$ [μM]	R_{max}	R^2
SHP3 (mit p97-ND1)	790.3	304.4	0.578	0.979
SHP9 (mit p97-N)	208.6	47.9	0.132	0.933

Abbildung 6.57: Gemessene Affinitäten der Fragmente SHP3 und SHP9

Auswertung der Daten von SHP3 (p97-ND1) und SHP9 (p97-N) mit Hilfe eines 1:1-Langmuir Modells



k_d	$\Delta_{Reg.}$	k_a	$\Delta_{Reg.}$	K_D	$\Delta_{Reg.}$
	[s^{-1}]	[$L \mu mol s^{-1}$]		[$\mu mol L^{-1}$]	
0.0398	0.0058	$3.47 \cdot 10^{-5}$	$7.44 \cdot 10^{-6}$	1147	414

Abbildung 6.58: Kinetische Auswertung der Daten von SHP9 (p97-N) mittels k_{obs}

6.4 Identifizierung möglicher Bindestellen der gefundenen Fragmente

6.4.1 *mixed-solvent* Simulationen ausgewählter Fragmente

Auf Basis der oben durchgeführten Analysen (6.2.4, 6.2.5), sowie der in 6.5.2 genannten ligandbasierten Pharmakophormodelle, wurden die Fragmente TROLL2, -6, -7, -8, -11, -12, -13, -14, -18, VIK20 und SHP9, für *mixed-solvent* Simulationen ausgewählt. Für TROLL2 und TROLL7 wurden beide Enantiomere simuliert. Mit Hilfe dieser Simulationen sollten mögliche Bindestellen an die N-Domäne postuliert werden. Die Simulationen wurden hierfür wie in 5.1.6.7 berichtet aufgesetzt und ergaben insgesamt eine totalen Simulationszeit von fast 2 μ s. Abbildung 6.59 zeigt eine Übersicht der erhaltenen Ergebnisse, anhand der Auswertung der Dichten der virtuellen Dummy-Atome. In die Auswertung gingen jeweils die zwei höchsten gefundenen Dichten eines jeden Fragments ein, sodass die Gesamtzahl der identifizierten Bindestellen bei n=26 lag. Wie schon in den Ergebnissen der FTMap-Analyse, lag der Hauptteil der vorgefundenen Interaktionsstellen auf der Nc-Subdomäne von p97-N, und nur an wenigen Fällen waren Aminosäuren der Nn-Subdomäne beteiligt. Der Hauptanteil aller Interaktionen fiel auf den SHP Bindebereich mit zusammengekommen über 50% der gefundenen Bindestellen. Hierbei waren die SHPI-II und SHP-III Bindestelle mit 23.1% gleich stark vertreten, die SHP-I Bindestelle enthielt 11.5% aller Interaktionen. Die restlichen gefundenen Interaktionsregionen entfielen vor allem auf den Linker-Bereich zwischen den Nn- und Nc-Subdomäne, sowie auf die der D1-Domäne zugewandten Bereiche. Der als UBXL benannte zentrale Bereich der Spalte zeigte nur 3.8 % der Interaktionen. Des Weiteren fallen die vorgefundenen Regionen größtenteils mit den anhand der Kristallstrukturen identifizierten *hot-spot* Bereichen zusammen.

Vergleicht man die Ergebnisse mit denen der FTMap-Analyse, so stimmen die Aussagen bezüglich der Fokussierung der Bindestellen auf die Nc-Subdomäne überein. Im Detail unterscheiden sie sich jedoch im Hinblick auf das nicht-Detektieren der SHP-II Bindestelle und den höheren Anteil an Interaktionen im UBXL-Bereich. Eine Erklärung ist die in der Simulation nicht mehr starre Proteinstruktur, sowie das explizite Wasser. Gerade die von Seitenketten gebildete SHP-II Bindestelle ist deutlich lösungsmittlexponiert und somit stark von Interaktionen mit Wasser betroffen, die die Konformation der SHP-II Bindestelle beeinflussen. Diese möglichen unterschiedlichen Konformationen der Bindestelle, sowie ein möglicher *induced fit* bei Bindung eines Fragments fehlen in

der Betrachtung der FTMap-Analyse vollständig. Dies spiegelt sich auch in den Ergebnissen für die UBXL-Bindestelle wieder, die in den *mixed-solvent* Simulationen deutlich weniger präsent sind, als in den FTMap-Analysen. Die starke Solvensexponiertheit dieser Region erschweren hier möglicherweise eine stabile Bindung von Fragmenten. Wassermoleküle werden in den FTMap-Analysen allerdings nicht berücksichtigt, wodurch die UBXL-Bindestelle hier höher vertreten ist. Die *mixed-solvent* Simulationen bieten, also ein realistischeres Bild des Systems, was allerdings mit einem enormen Mehraufwand an Rechenzeit bezahlt wird¹⁸. Die Wahl der SHP-Bindestelle als primäres Ziel zum Design von PPI-Inhibitoren wurde somit von den computerbasierten Ergebnissen bestätigt, und es kann die Schlussfolgerung gezogen werden, dass es sich bei der Nc-Subdomäne um den Teil der N-Domäne handelt, an dem am ehesten wirkstoffähnliche Moleküle binden können. Anhand der vorgefundenen Ergebnisse wurde die Frage gestellt, ob neben einer Lokalisation, auch die Vorhersage von Bindemodi der Fragmente möglich ist. Eine hohe Besetzungsdichte mit dem virtuellen Dummy-Atom sagt zwar aus, dass sich das Fragment in den vier Replika der Simulation an einer definierten Stelle bevorzugt aufhält, jedoch nicht, ob auch eine spezifische Bindung vorliegt. Hierzu wurden von den Fragmenten die Besetzungsdichten ausgewählter Schweratome erstellt und diese im Hinblick auf eine Orientierung an den vorgefundenen Bindestellen analysiert.

Hier konnten nur für TROLL2 (S), 7 (S), 8, 12 und SHP9 bevorzugte Orientierungen anhand der Besetzungsdichten gefunden werden. Die restlichen Fragmente zeigten keine eindeutigen Dichten. Im Folgenden soll kurz auf die einzelnen Fragmente eingegangen werden. Abbildung 6.60 zeigt die vorgefundenen Dichten. Überlagert sind diese mit Momentaufnahmen aus den Simulationen, die das Fragment und das Protein zur besseren Zuordnung der Dichten zeigen¹⁹. Der prozentuale Besetzungsgrad der jeweiligen Atome ist nicht einheitlich, sondern wurden zur besseren Visualisierung jeweils angepasst und variierte dabei zwischen 10% und 30%. Ein Zusammenhang zwischen einem höheren Besetzungsgrad und der Affinität konnte nicht beobachtet werden. Generell wurde in einer Trajektorie nur die Bindung eines Fragments in eine Binde-tasche, nicht aber das Verlassen dieser beobachtet. TROLL2 zeigte nur für das S-Enantiomer eine deutliche Orientierung in den Dichten. Eine lag innerhalb der SHP-II Bindestelle (Abbildung 6.59 (a)) und eine weitere in der SHP-I Bindestelle

¹⁸1 Stunde beim FTMap-Algorithmus im Vergleich zu mehreren Wochen Simulationen für die Fragmente

¹⁹Die Momentaufnahmen wurden nicht weiter bearbeitet und sind daher nicht in einem lokalen Minimum bezüglich der potentiellen Energie. Dies führt zu leichten Abweichungen der gezeigten Bindungen von idealen Winkeln und Abständen.

(Abbildung 6.59 **(b)**) (die prozentualen Besetzungsgrade sind für beide Abbildungen gleich gewählt worden). Das fast gleiche Fragment TROLL7, das sich nur hinsichtlich des Halogens unterscheidet, zeigte hier die gleichen Ergebnisse, (Abbildung 6.59 **(c)**) und (Abbildung 6.59 **(d)**). Der Besetzungsgrad beider Fragmente war jeweils für die Halogene und den lipophilen Aromaten am höchsten und fiel zu den deutlich solventexponierten polaren Gruppen hin ab. Für TROLL8 konnte nur eine definierte Bindepose in der SHP-III Region ausgemacht werden (Abbildung 6.59 **(e)**). Hierbei interagierte die 3-Aminopyrazol-Struktur mit dem Peptidrückgrat im Bereich der Aminosäuren 185 und 186, während sich der Methylfuran-Ring in die lipophile Region der Kavität orientierte. Anzumerken ist hier, dass nur das gezeigte Tautomer des Pyrazolrings simuliert wurde. Dichten für TROLL12 fanden sich ebenfalls innerhalb der SHP-II Region (Abbildung 6.59 **(f)**). Der hydrophobe Phenylring zeigte wieder in den lipophilen Bereich der Binderegion, während der polare Teil dem Solvent zugewandt war. Eine Salzbrücke konnte zwischen E167 und dem protoniertem Amin von TROLL12 beobachtet werden. Für SHP9 wurden zwei Posen gefunden. Die erste lag in der SHP-II Bindestelle (Abbildung 6.59 **(h)**), mit einer Lokalisation des Chloratoms ähnlich TROLL7 und einer Orientierung des Pyrimidinrings ähnlich TROLL8, wobei für diesen nur ansatzweise Dichten vorhanden waren. Die zweite Pose lag zwischen SHP-I und SHP-III in direkter Nachbarschaft zu F131 (Abbildung 6.59 **(g)**). Hier interagierte vor allem der Pyrimidinring mit I182 über eine Wasserstoffbrücke, ähnlich den Peptiden des SHP-Motivs. Der hydrophobe Teil zeigt in Richtung von C184. Die mit Hilfe von GOLD gefundene und anschließend in der MD-Simulation als stabil identifizierte Pose konnte in einer *mixed-solvent* Simulation ebenfalls vorgefunden werden. Zusätzlich gab es eine Dichte für den Pyrimidinring im polaren Teil der SHP-I Bindetasche in Richtung Y143. Diese Dichten waren allerdings schwächer als die in **(h)** gezeigten.

Zusammenfassend lässt sich festhalten, dass für die Fragmente TROLL2, 7, 8, 12, sowie SHP9 mögliche Bindeposen postuliert werden können. Eine Erhöhung der Simulationszeit wäre ein Schritt, um vielleicht noch deutlichere Ergebnisse zu erhalten. Die Simulationen wurden nicht auf deren Konvergenz hin untersucht. Hier könnte der von *Guvench & MacKerell* [306] beschriebene Ansatz durchgeführt werden, bei dem jeweils die Dichten aus der gleichen Anzahl an Replika zu zwei Summendichten addiert und diese dann voneinander subtrahiert werden

²⁰Beispiel: Bei 10 simulierten Replika werden die Dichten von jeweils 5 Replika zu einer Summendichte D_1 und D_2 summiert. Die Subtraktionen von $D_1 - D_2$ sollte eine Dichte ergeben, die nur mehr ein Rauschen um null zeigt. Dies würde zeigen, dass die *mixed-solvent* Simulation konvergiert ist

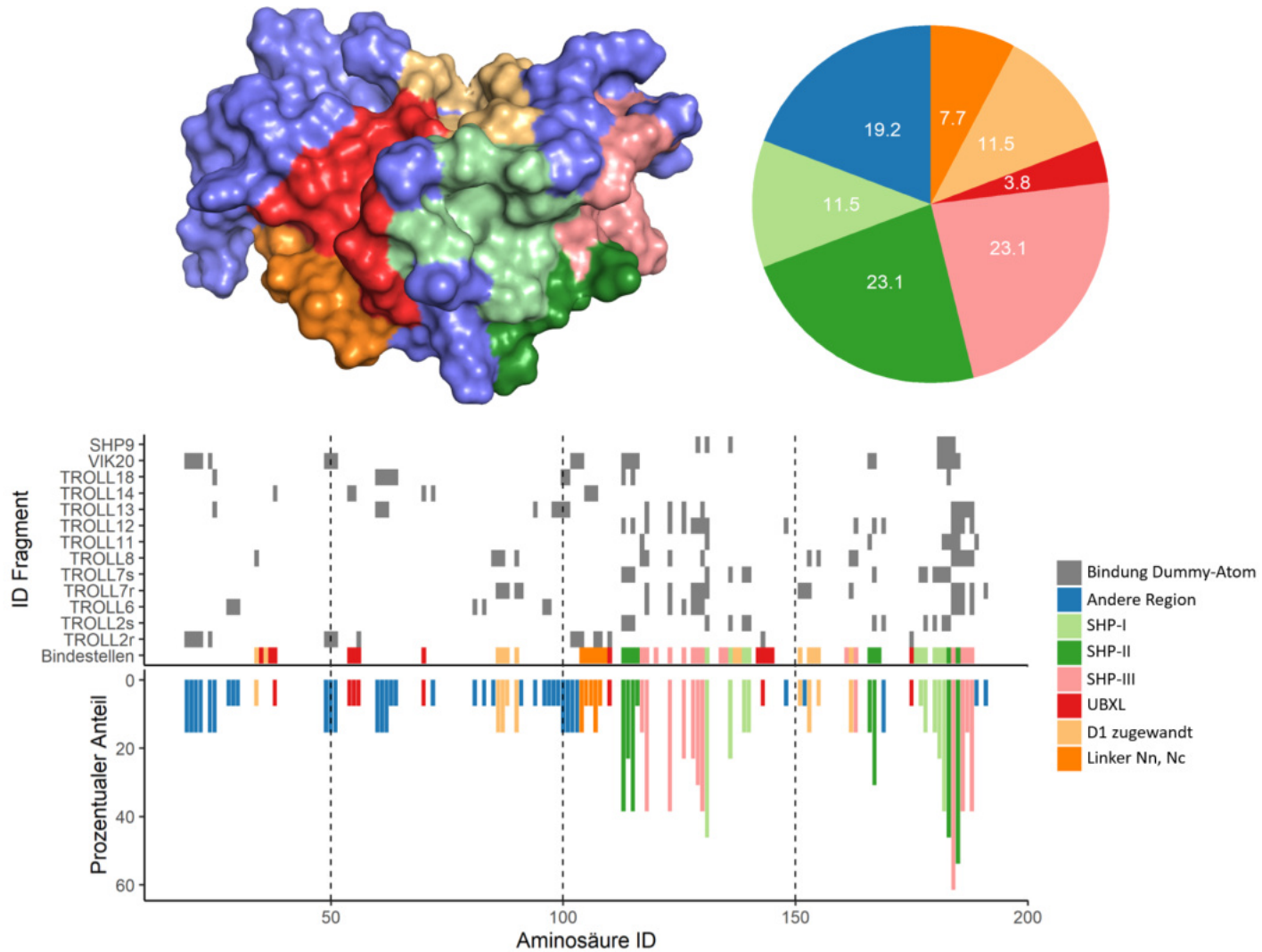


Abbildung 6.59: Übersicht der gefundenen Bindestellen anhand der Besetzungsdichten der Dummy-Atome.

Die meisten Interaktionsregionen lagen in der Nc-Subdomäne von p97-N, wobei der höchste Anteil in der SHP-Bindestelle lokalisiert war.

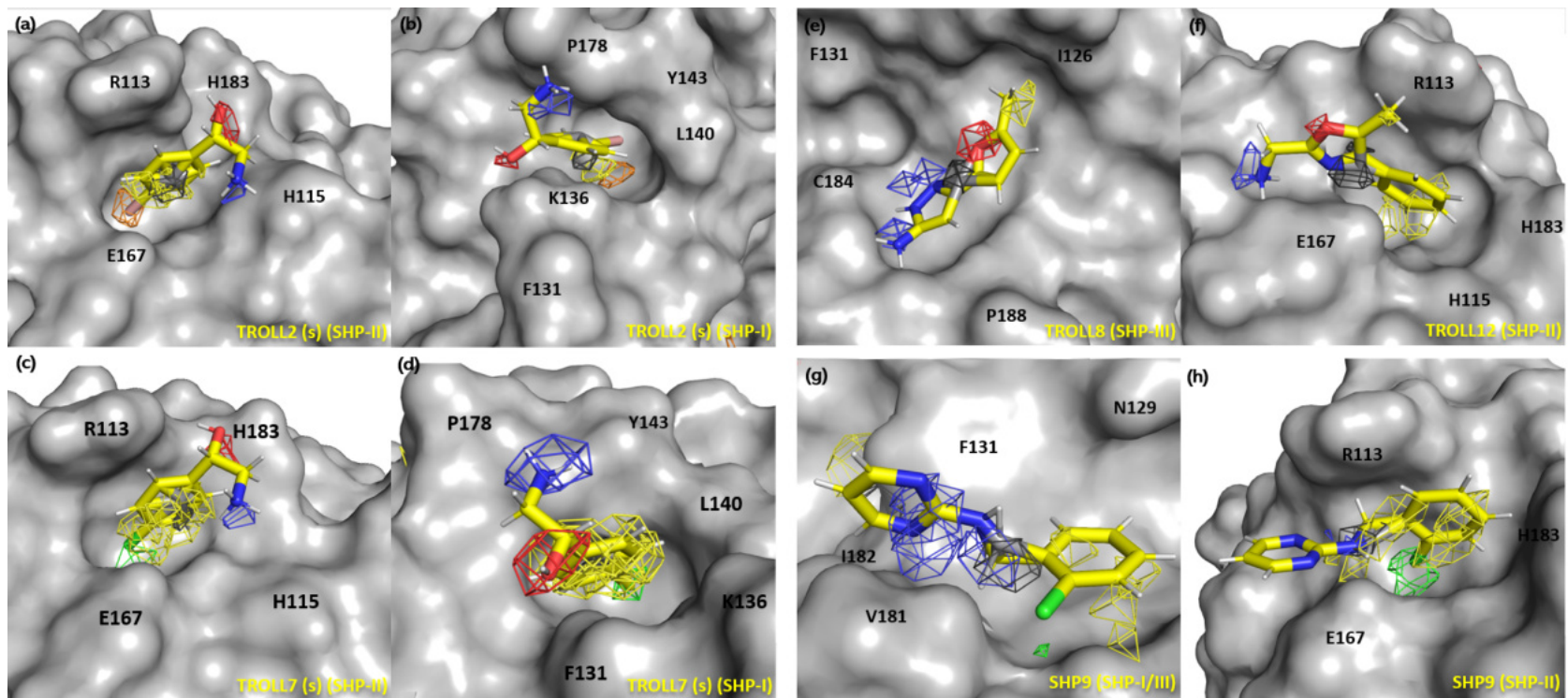


Abbildung 6.60: Besetzungsdichten für unterschiedliche Schweratome

Dichten für Kohlenstoffatome (gelb), Chloratome (grün), Stickstoffatome (blau), Sauerstoffatome (rot) und Dummy-Atome (grau) der Fragmente innerhalb ihrer jeweiligen Bindestelle. Es sind hier deutliche Orientierungen der Fragmente zu sehen. Überlagert sind die Dichten mit aus den Trajektorien ausgewählten Momentenaufnahmen, um die Bindemodi besser visualisieren zu können.

6.4.2 STD-NMR-Messungen an Punkt-Mutanten von p97-ND1

Neben den *mixed-solvent* Simulationen wurden STD-NMR-Messungen an zwei Punktmutanten von p97-ND1 durchgeführt. Bei den Mutanten handelte es sich jeweils um ein Konstrukt mit einer F131A-Mutation bzw. einer I182A-Mutation, mit der eine mögliche Bindung an die SHP-I Region detektiert werden sollte (Abbildung 6.62). Falls ein Fragment innerhalb dieser Region bindet, sollte sich dies auch in der Stärke der beobachteten STD-Effekte (positiv wie auch negativ) bemerkbar machen. So sollte eine F131A-Mutante einen deutlichen Einfluss auf die Geometrie der Bindestelle haben, und I182A neben geometrischen auch magnetische Änderungen zeigen, da die Seitenkette von I182A direkt gesättigt wurde. Voraussetzung hierfür sind allerdings die immer gleichen Ligand- und Proteinkonzentrationen bei den Messungen. Da für den Ligand immer die gleiche Stammlösung benutzt wurde, liegt die Hauptfehlerquelle in der Proteinkonzentration. Die Konzentrationen aller Proteinlösungen wurden nach dem Überführen in den NMR-Puffer daher immer mittels $n=3$ Messungen bestimmt.

Um den Einfluss der Proteinkonzentration auf die Stärke des STD-Signals besser einschätzen zu können, wurden drei Proteinlösungen von p97-ND1-WT mit $30 \mu\text{M} \pm 20\%$ hergestellt und mit TROLL2 vermessen. Abbildung 6.63 zeigt, dass eine Änderung von $\pm 20\%$ in der Proteinkonzentration eine Änderung um $\pm 3\%$ im mittleren STD-Signal hervorruft. Der Mittelwert der drei Messungen zeigt ein mittleres STD-Signal von 36.5% mit einem σ von 4.0% . Als Grenze für die weiteren Messungen wurde festgelegt, dass eine Änderung von mindestens 2σ in den STD-Signalen der unterschiedlichen Mutanten beobachtet werden musste, um sicherzugehen, dass es sich bei einem beobachteten Effekt nicht um bloße Variabilitäten der eingesetzten Proteinkonzentration handelte.

Für Messungen wurden basierend auf den *mixed-solvent* Simulationen und dem virtuellen Screening folgende Fragmente ausgewählt:

- SHP9 (SHP-I Bindestelle)
- TROLL2 (SHP-I Bindestelle)
- TROLL8 (SHP-III, in der Nähe zu I182)
- TROLL14, für das sich zwar keine eindeutige Orientierung in der *mixed-solvent* Simulation zeigte, das aber auf Basis der Pharmakophormodelle (siehe 6.5.2) mit einbezogen wurde.

Eine Analyse der Ergebnisse der Messungen ist in Abbildung 6.62 zu sehen. Für TROLL2 ist keine signifikante Änderung der STD-Signale zu sehen. Anhand dieser Ergebnisse kann keine Bindung an SHP-I für TROLL2 bestätigt werden.

TROLL8 zeigt für die F131A-Mutante das gleiche STD-Signal, wie für den Wildtyp, was mit den Ergebnissen der *mixed-solvent* Simulation in Übereinstimmung steht. Eine Änderung des Phenylalanins zu Alanin sollte keinen direkten Einfluss auf die Bindung in der SHP-III Region haben. Die I182A-Mutante zeigt dagegen ein extrem erhöhtes STD-Signal. Geht man weiterhin von der SHP-III Bindestelle aus, so könnte die Änderung von Isoleucin in ein Alanin einen nicht bekannten Einfluss auf die Konformation der SHP-III Bindestelle hervorrufen, mit dem Resultat einer verstärkten Wechselwirkung mit TROLL8. Alternativ könnte durch die kleinere Seitenkette des Alanins an Position 182 eine zusätzliche hydrophobe Tasche entstehen, die zusätzlich von TROLL8 adressiert wird. Beides sind nur Hypothesen, die weiterer Untersuchungen bedürfen (siehe unten). Auch für TROLL14 lässt sich ein ähnliches Verhalten beobachten, das so ohne Weiteres keine wirkliche Aussage über die Bindestelle zulässt.

SHP9 zeigte dagegen ein deutlich reduziertes Signal im Falle der F131A-Mutation. Geht man von einer Bindepose wie im virtuellen Screening oder auch in der *mixed-solvent* Simulation aus, so würde der Pyrimidinring mit F131 π - π Interaktionen eingehen. Ein Wegfall dieser Interaktion sollte die Bindung schwächen und somit eine negative Änderung im STD-Signal hervorrufen, wie dies auch zu beobachten war. Die Differenz zwischen dem höchsten beobachteten STD-Signal und dem mittleren STD-Signal, kann als Hinweis auf eine spezifische Bindung gesehen werden. Die F131A-Mutation zeigt im Vergleich zum WT eine deutliche Reduktion dieser Differenz, und kann mit einer mehr unspezifischen Bindung interpretiert werden. Interessanterweise zeigen alle anderen Fragmente keine solche Änderung. Die Differenzen sind hier für alle Mutanten nahezu gleich. Die Ergebnisse stützen eine Wechselwirkung von SHP9 mit der SHP-I Bindestelle, wie sie theoretisch vorhergesagt wurde. Die Erhöhung des Signals für I182A könnte, wie bereits bei TROLL8 erwähnt durch die Ausbildung einer zusätzlichen Kavität erklärt werden, die eine bessere Interaktion ermöglicht.

Zusammenfassend lässt sich hier sagen, dass die STD-NMR-Messungen nur für SHP9 erste Hinweise auf eine Bindestelle lieferten und nicht so aussagekräftig wie erhofft sind. Gerade die Mutante I182A zeigt, dass die Interpretation der STD-Signale sich als schwierig erweist. Unterschiedliche Effekte, wie ein veränderter NOE-Effekt und mögliche Konformationsänderungen des Proteins spielen hier vermutlich eine Rolle. Eine Messung der Affinitäten der Fragmente mit diesen Mutanten wäre sicherlich eine zusätzliche Hilfe bei der Interpretation der Ergebnisse.

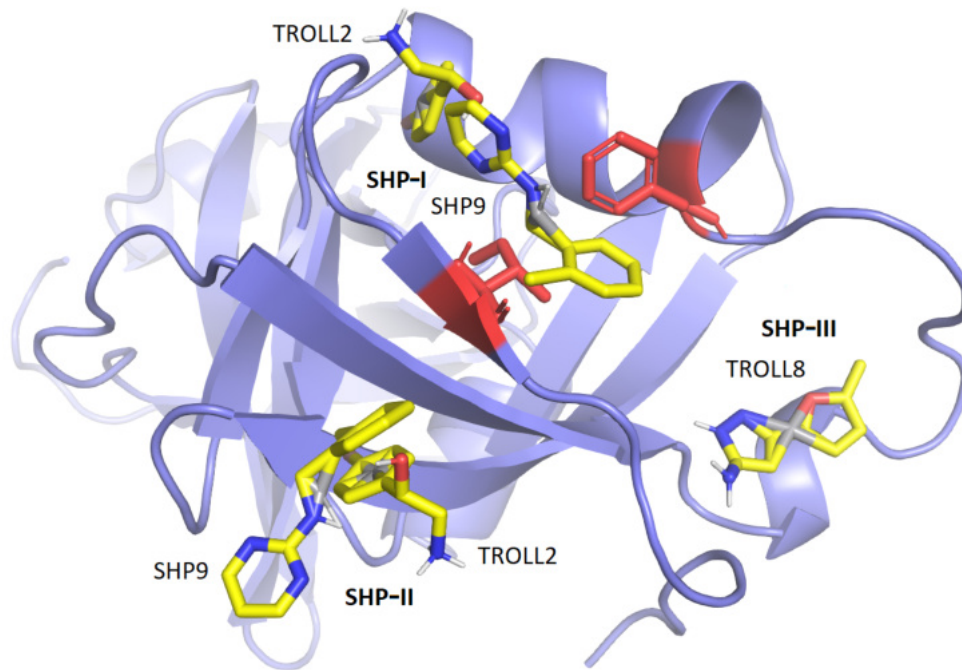


Abbildung 6.61: Lage der zwei Alanin-Mutationen (rot) in Relation zu den Ergebnissen der *mixed-solvent* Simulationen

Die rot markierten Aminosäuren F131 und I182 wurden jeweils durch Alanin ersetzt, um Einflüsse auf eine mögliche Bindung in die SHP-I Region mittels STD-NMR-Messungen zu erfassen. Die abgebildeten Fragmente zeigen die in den *mixed-solvent* Simulationen gefundenen Bindemodi und enthalten die in Grau gezeigten Dummy-Atome.

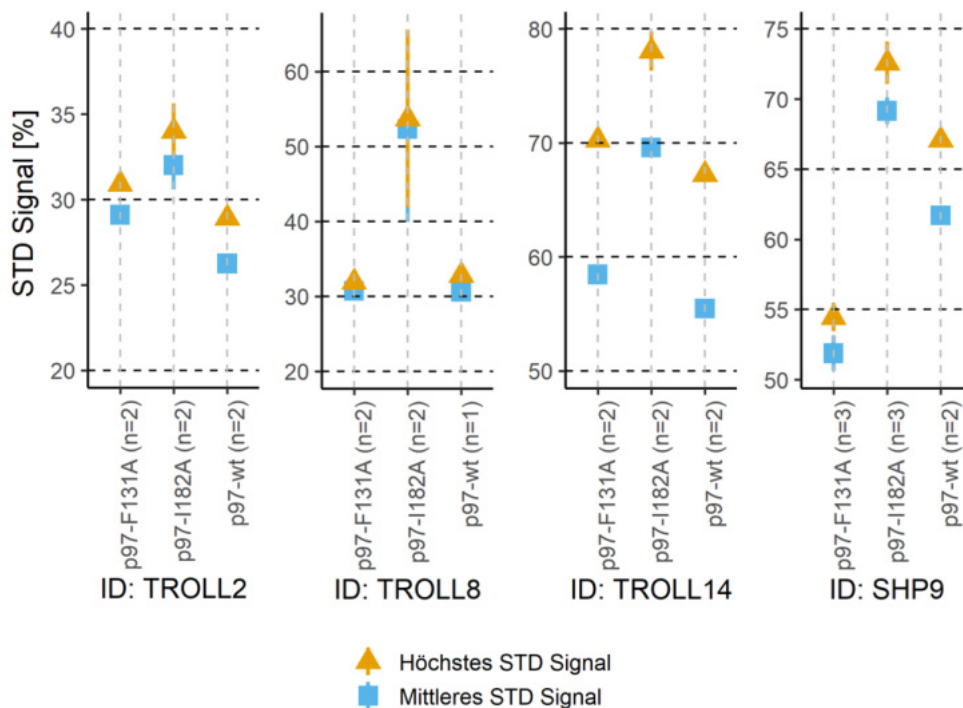
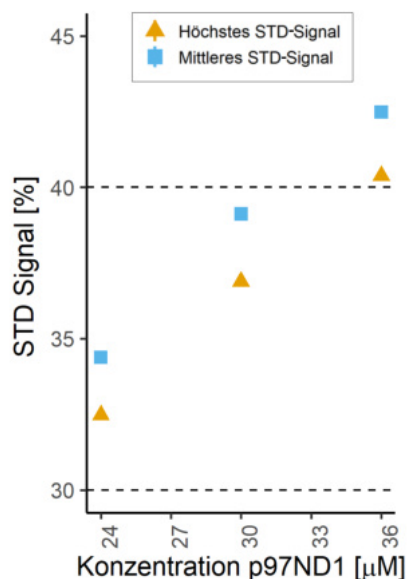


Abbildung 6.62: STD-NMR-Messungen an Alanin-Punkt-Mutanten innerhalb der SHP-I Bindestelle von p97-ND1



Konzentration [µM]	std_{mean}^{arom} -Signal [%]	std_{max}^{arom} -Signal [%]
24	32.4	34.3
30	37.0	39.1
36	40.3	42.4

Abbildung 6.63: Abhängigkeit des STD-Signals von der Proteinkonzentration

6.4.3 Kristallstruktur des TROLL2 p97-ND1 Komplexes

6.4.3.1 Kristallisation von Apo-p97-ND1^{L198W}

Die von *Tang & Xia* [261] berichtete Kristallisationsbedingung für das Konstrukt p97-ND1^{L198W} konnte nicht reproduziert werden und lieferte nur linsenförmige Sphärolite, die keine Streuung zeigten. Aus den daraufhin durchgeführten *de novo* Kristallisationsscreenings konnten zwar kleinere Proteinkristalle und Kristallnadeln gewonnen werden, jedoch ließ sich keine dieser Bedingungen weiter optimieren. Dementsprechend wurde ein individueller Natriumformiat-Screen durchgeführt, der auf der bei *Tang & Xia* [261] berichteten Bedingung aufbaute und zusätzliche ausgewählte Polyethylenglykole (PEG) enthielt (PEG 200, 600, 2000, 4000 und 6000). Der Zusatz von PEG600 führte hierbei zum Erfolg und ein Fein-Screening dieser Bedingung ergab die folgende Zusammensetzung der Mutterlauge, die für alle weiteren Kristallisationsexperimente benutzt wurde:

Tabelle 6.19: Kristallisationsbedingung für p97-ND1^{L198W}

Bestandteil	Konzentration
Natriumformiat (pH 6.0)	4.0 M
Glycerol (v/v)	10 %
PEG 600 (v/v)	5 %

Die Kristalle hatten ein hexagonales Erscheinungsbild in Form von „Rädern“ bzw. dickeren Stäben, die teilweise auf einer Seite hin spitz zuliefen (siehe Abbildung 6.64). Die finale Größe wurde innerhalb von 4 bis 6 Tagen bei 20 °C erreicht. Es konnten Kristalle ohne und mit ADP erhalten werden, wobei sich mit ADP deutlich größere Kristalle ergaben. Weitere Kristallisationen wurde daher stets in Gegenwart von 4 mM ADP durchgeführt.

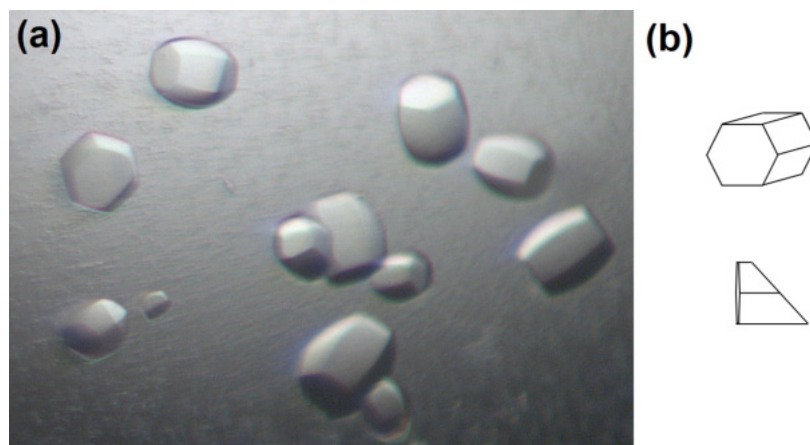


Abbildung 6.64: Proteinkristalle von p97-ND1^{L198W}.

(a) Aufnahme von Proteinkristallen der p97-ND1^{L198W} Mutante. Deutlich ist die hexagonale Form der Einzelkristalle zu sehen.

(b) Schematische Darstellung des hexagonalen Erscheinungsbilds der gefundenen Proteinkristalle. Etwa die Hälfte der Kristalle zeigte eine spitz zulaufende Seite.

6.4.3.2 Datensammlung und Strukturlösung von p97-ND1^{L198W} im Komplex mit Fragmenten

Erste kristallographische Modelle der untersuchten ND1^{L198W} Kristalle ohne Liganden entsprachen der bei *Tang & Xia* [261] berichteten Struktur. Die mit XDS bzw. mit POINTLESS bestimmte Raumgruppe der Kristalle war mit P622, sowie einem ND1-Monomer in der asymmetrischen Einheit identisch. Die statistisch noch sinnvollen Auflösungen lagen bei bis zu 1.9 Å und konnten somit gegenüber *Tang & Xia* leicht verbessert werden. Alle p97 Proteine lagen als biologisches Hexamer in der ADP-gebundenen Konformation vor. In allen D1-Domänen konnten

eindeutige Dichten für das gebundene ADP in der Nukleotidbindetasche gefunden werden.

In der vorgefundenen Kristallpackung (Abbildung 6.65) waren die Bindestellen für die UBXL-Domänen, die SHP-III Region, der Nn-Nc-Linkerbereich, sowie der Bereich zwischen N- und D1-Domäne frei von Kristallkontakten und somit zugänglich für Fragmente. Die zentrale Stelle der SHP-II Region ist ebenfalls frei von Kristallkontakten und zeigt nur in den Randregionen Kontakte mit benachbarten symmetrieverwandten N-Domänen. Im Gegensatz dazu ist die Bindestelle SHP-I durch einen Kristallkontakt mit dem N-Terminus einer symmetrieverwandten N-Domäne blockiert und somit nicht zugänglich.

Die ausgewählten Fragmente TROLL2,-6, -7, -8, -12, -14, VIK20 und SHP9 wurden daraufhin mit p97-ND1^{L198W} ko-kristallisiert. Für TROLL2, -7, -12 und VIK20 konnten Einzelkristalle mit ausreichender Größe erhalten werden. Auffällig war, dass die Fragmente sich teilweise günstig auf das Wachstum der Proteinkristalle auswirkten. So erreichten Kristalle mit TROLL2, -7 und VIK20 eine teilweise beachtliche Größe und zeigten schärfere Kanten als Kristalle ohne Fragmente. Alle Kristalle wurden vor dem Fischen zusätzlich in dem jeweiligen Fragment für unterschiedliche Zeiten getränkt und für einige Sekunden in die zwei Kryoprotektionslösungen überführt, bevor sie in flüssigen Stickstoff schockgekühlt wurden. Die Aufnahme der Datensätze fand an der European Synchrotron Radiation Facility (ESRF, Grenoble, Frankreich) an Beamline ID30B per Remoteverbindung durch Prof. Hermann Schindelin statt. Um ein mögliches anomales Signal zur Lokalisation des Fragments TROLL2 ausnutzen zu können, wurden die Datensätze nahe der Bromkante bei einer Wellenlänge von 0.9184 Å aufgenommen.

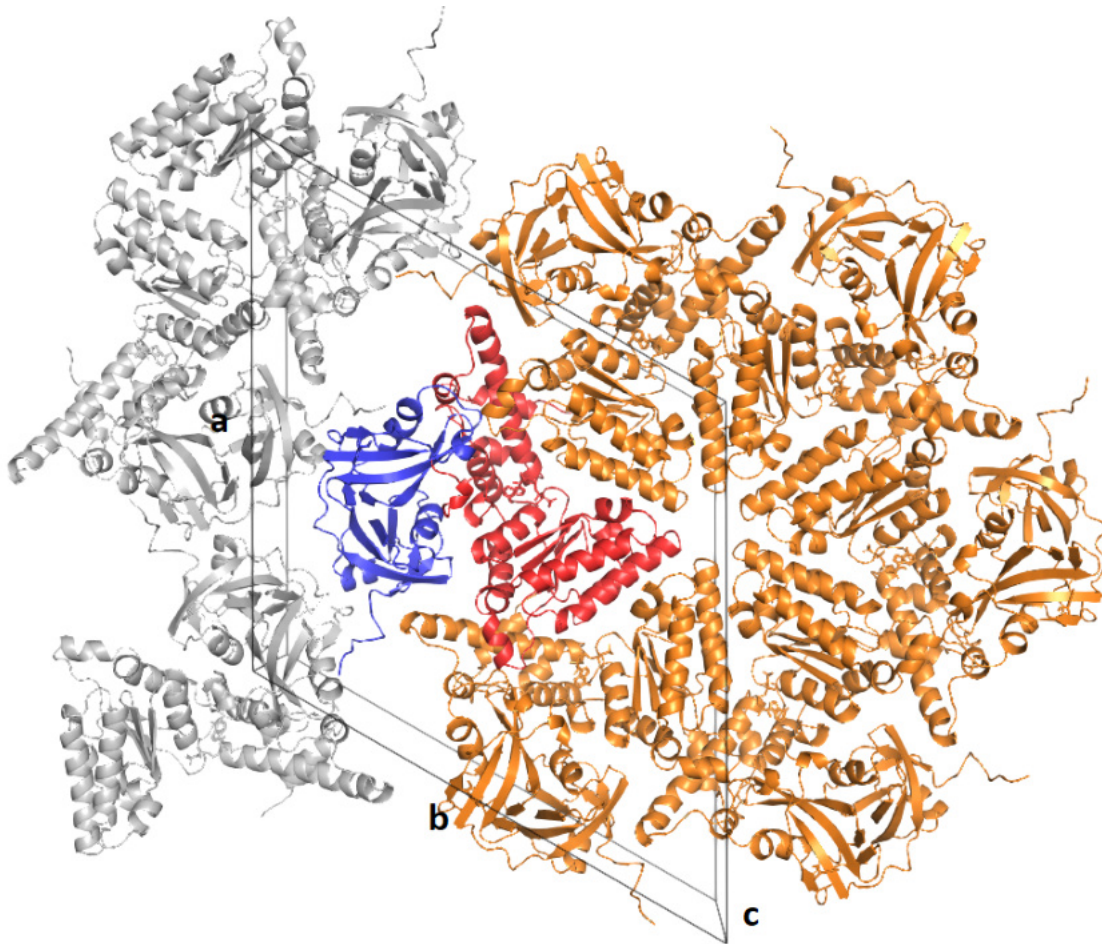


Abbildung 6.65: Kristallpackung der erhaltenen p97-ND1^{L198W} Kristalle in der P622-Raumgruppe.

Die hexagonale Einheitszelle ist in Schwarz eingezeichnet. In Rot ist die D1-Domäne und in Blau die N-Domäne eines Monomers gezeigt. In Orange sind die restlichen Monomere gezeigt, die zusammen ein p97-ND1 Hexamer bilden, sowie in Grau weitere Kristallkontakte von symmetrieverwandten p97-ND1-Hexameren. Die gezeigte Struktur basiert auf dem im Folgenden beschriebenen Datensatz mit dem Fragment TROLL2.

6.4.3.3 Struktur des Fragments TROLL2 im Komplex mit p97-ND1^{L198W}

TROLL2 konnte bis zum Abschluss dieser Arbeit als einziges der untersuchten Fragmente kristallographisch bestätigt werden. Der Nachweis gelang hierbei vor allem durch das eindeutige anomale Signal des in TROLL2 enthaltenen Bromatoms. Die Prozessierung, Integration und Skalierung der mit TROLL2 getränkten Datensätze erfolgte in XDS [238] und die weitere Auswertung des anomalen Signals mit Hilfe von SHELXC, SHELXD [235] und ANODE [217]. Die statistischen Ergebnisse der drei erhaltenen Datensätze zeigt Tabelle 6.20:

Tabelle 6.20: Statistiken der Datensammlungen von mit TROLL2 ko-kristallisierten und gesoakten p97-ND1^{L198W} Kristallen

Datensatz	1	2	3
Wellenlänge [Å]	0.918401	0.918401	0.918401
Raumgruppe	P622	P622	P622
Auflösung [Å]	19.72-2.0 (2.05-2.00) ^b	19.83-2.3 (2.38-2.3)	19.79-2.5 (2.6-2.5)
Mosaizität	0.09	0.1	0.24
Dimensionen der Einheitszelle [Å]	a/b=146.23, c=84.4	a/b=147.01, c=84.96	a/b=146.76, c=84.42
Anzahl gemessener Reflexe	1068655	948435	723244
Anzahl unabhängiger Reflexe	36339	24510	19018
Multiplizität	29.4 (29.1)	38.4 (31.0)	38 (32)
$\langle I/\sigma I \rangle$	19.3(3.0)	16.5 (2.6)	19.6 (2.4)
Vollständigkeit (%)	99.9 (100.0)	99.8 (100.0)	99.8 (99.9)
R _{merge}	0.122 (1.419)	0.171 (2.072)	0.162 (2.478)
R _{meas}	0.126 (1.469)	0.176 (2.124)	0.166 (2.539)
R _{pim}	0.032 (0.376)	0.039 (0.465)	0.037 (0.554)
Anzahl Brom Bindestellen (SHELXD)	1	1	0

Angaben in Klammern beziehen sich auf die Schale mit der höchsten Auflösung.

^b Die Datensätze wurden bei einer geringeren Auflösung von 20.0 Å beschnitten.

Alle Kristalle diffraktierten anisotrop mit einer höheren Auflösung in der h, k -Ebene und einer deutlich reduzierten entlang der l -Achse. Die Datensätze wurden anhand ihrer $\langle I/\sigma I \rangle$, R_{meas} auf die in Tabelle 6.20 genannten, als sinnvoll erachteten Auflösungen beschnitten. Datensatz 1 zeigte die höchste Auflösung von bis zu 2.0 Å, sowie ein eindeutiges anomales Signal für Brom mit einer maximalen Stärke von 11.25-fachen der RMSD^e (*root mean square deviation*)²¹, wie durch das Programm ANODE berechnet. Datensatz 2 zeigte eine um 0.3 Å schlechtere Auflösung mit einem nicht mehr so starken anomalen Signal von 5.2-fachen der RMSD^e (ANODE), ebenfalls in der Region der SHP-II Bindestelle. In Datensatz 3 konnte kein signifikantes anomales Signal für Brom gefunden werden. Datensatz 1 wurde daraufhin zur Generierung eines Modells mittels molekularem Ersatz

²¹RMSD-Werte im Kontext der Konturniveaus von Elektronendichten werden in dieser Arbeit mit RMSD^e bezeichnet, um diese von den genannten RMSD-Werten der computerbasierten Methoden besser abgrenzen zu können.

in PHASER (in CCP4i [219]) und zur weiteren Strukturverfeinerung in BUSTER [218] und Coot [220] ausgewählt. Tabelle 6.21 zeigt die Verfeinerungsstatistik für den ausgewählten Datensatz.

Tabelle 6.21: Statistiken der Verfeinerung von Datensatz 1

Auflösung [Å]	19.72-2.0
R _{free} [%]	24.09
R _{work} [%]	21.38
Mittlerer B-Faktor [Å ²]	52.58
Anzahl Nichtwasserstoffatome	
Protein	3512
Liganden	11 (TROLL2), 27 (ADP)
Puffermoleküle	52 (Formiat und Glycerin)
Ionen	1 (Na ⁺)
Wasser	227
RMS^aAbweichungen	
Bindungslängen [Å]	0.0091
Bindungswinkel [°]	1.10
Ramachandran-Diagramm	
Erlaubt [%]	96.4 %
Zusätzlich erlaubt [%]	3.15 %
Verboten [%]	0.45 %

^a *root mean square* (Wurzel aus dem quadratischen Mittel)

Die p97-ND1^{L198W}-TROLL2 Struktur zeigt die gleiche Tertiärstruktur wie die publizierte p97-ND1^{L198W}-Struktur (5DYG [261]). Eine Überlagerung beider Strukturen in Pymol ergab einen RMSD-Wert von nur 0.204 Å (Abbildung 6.66 (a)). Lediglich ein Bereich in der D1-Domäne zeigt in der erhaltenen Struktur eine ungeordnetere α -Helix, wobei die Struktur hier nicht genau aufgelöst ist (Abbildung 6.66 (b)). In diesen Bereich fallen auch die beiden Ramachandran-Ausreißer der Struktur. Das Bromatom zeigt ein eindeutiges anomales Signal in der Nc-Subdomäne von p97 (Abbildung 6.67). Bei dem gezeigten Konturniveau von 7.0 konnte kein weiteres Signal in der gesamten Einheitszelle beobachtet werden. Anhand dieses anomalen Signals konnte die Bindestelle des Fragments TROLL2 eindeutig in der SHP-II Bindestelle identifiziert werden. Im Kristall lagen an der

SHP-II Bindestelle zwei Kristallkontakte vor. R144* einer N-Domäne, sowie R64** einer weiteren N-Domäne (beide in Schwarz dargestellt) interagiert an dieser Stelle mit der SHP-II Region (Abbildung 6.68 (a)).

Abbildung 6.68 (b) zeigt die SigmaA-gewichtete 2Fo-Fc-Elektronendichte bei einem Konturniveau von 1.00, bei der nur eine eindeutige Dichte für das Bromatom und benachbarte Atome des Phenylrings zu sehen sind. Bei einem Konturniveau von 0.50 werden größere Teile des Phenylrings wie auch die Hydroxylgruppe sichtbar (Abbildung 6.68 (c)). Auch wenn diese Dichten nur sehr schwach (TROLL2 Besetzungsgrad 0.54) und fragmentarisch sind, können sie aufgrund des signifikanten anomalen Signals des Broms als eindeutig angesehen werden. In Phenix wurde daraufhin eine *Polder*-Elektronendichtekarte erstellt. Dabei handelt es sich um eine *omit*-Elektronendichtekarte, die das Signal um Einflüsse des Bulk-Solvens korrigiert und somit schwächere Signale verstärkt. Die so erhaltene Elektronendichtekarte zeigt eine deutliche Dichte für das Brom, sowie einen Teil des Phenylrings (Abbildung 6.68 (d)).

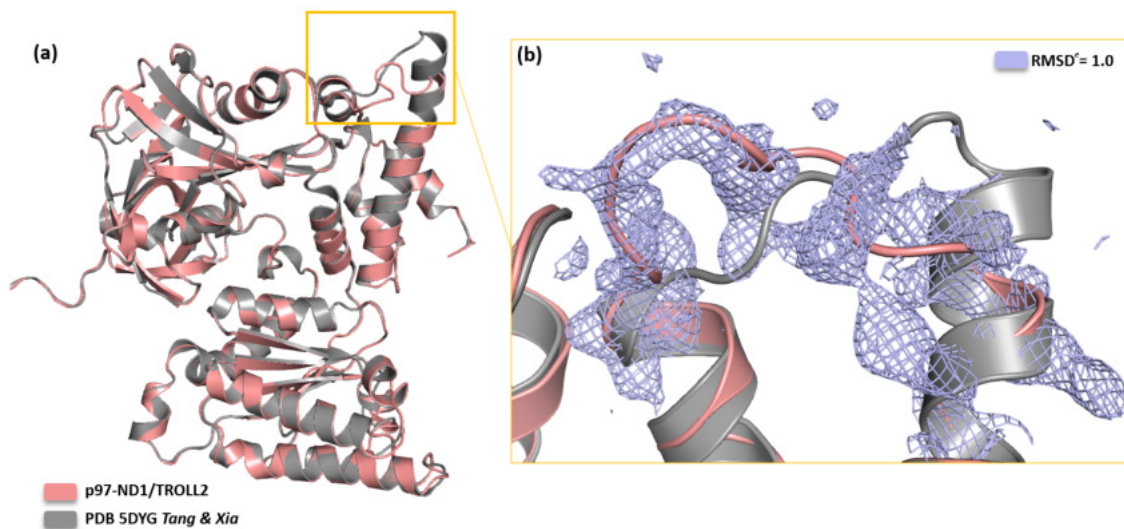


Abbildung 6.66: Vergleich der p97-ND1-TROLL2 Struktur mit PDB: 5DYG [261]

Eine Überlagerung der bekannten Struktur mit der hier erhaltenen zeigt keine signifikanten Unterschiede. Lediglich eine Stelle in der D1-Domäne, die jedoch in beiden Strukturen partiell ungeordnet ist, zeigte Unterschiede. Gezeigt ist für diesen Bereich die SigmaA-gewichtete 2Fo-Fc-Elektronendichte bei einem Konturniveau vom 1-fachen der $RMSD^e$.

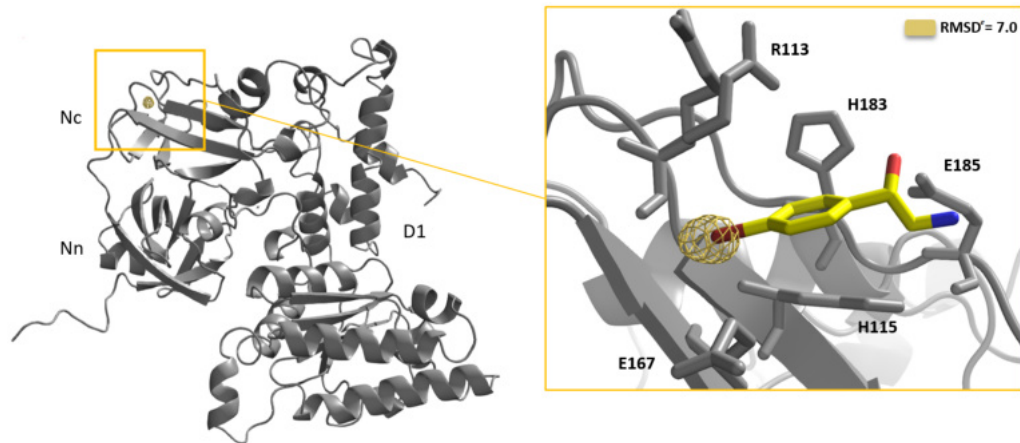


Abbildung 6.67: Anomales Signal des Bromatoms von TROLL2

Gezeigt ist die eindeutige Lokalisation des anomalen Signals in der SHP-II Bindestelle der p97-N Domäne, sowie eine detaillierte Darstellung der TROLL2 Bindetasche mit umgebenden Resten der N-Domäne. Die anomale Differenzdichte von Brom ist bei einer RMSD^e von 7.0 konturiert.

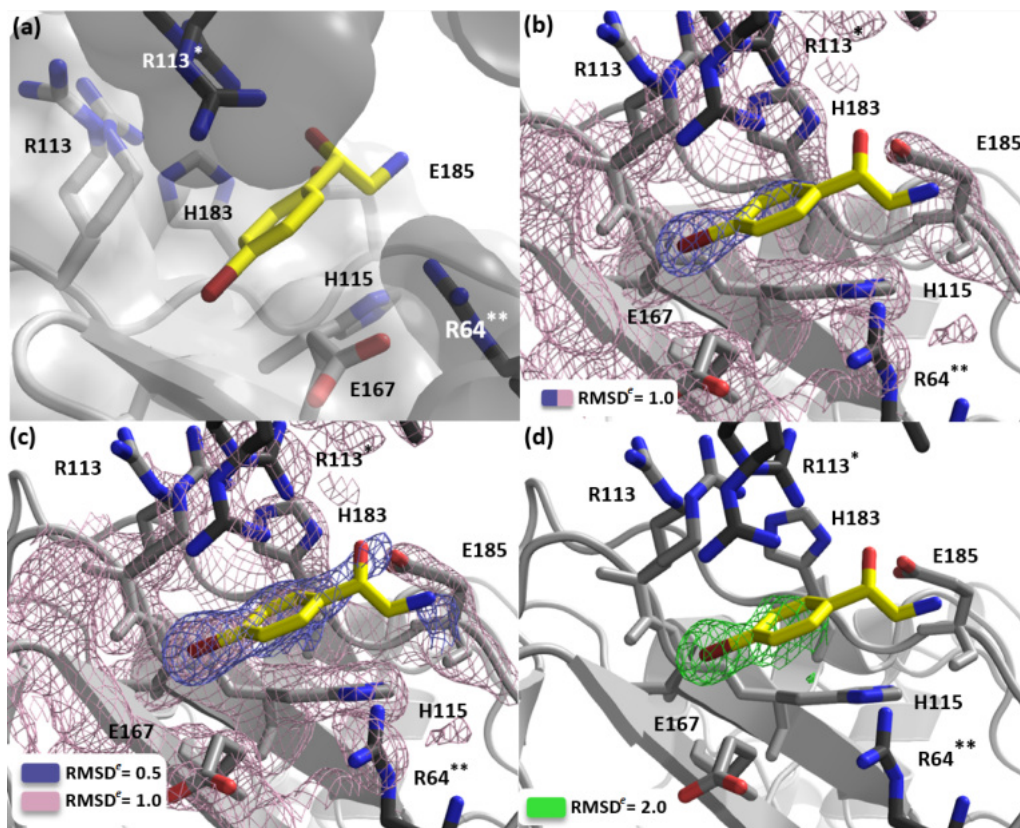


Abbildung 6.68: Bindetasche und Elektronendichten für TROLL2

(a) Struktur der Bindestelle im Kristall (b) Elektronendichte bei einer RMSD^e von 1.0 (c) Elektronendichte bei einer RMSD^e von 0.5 (d) Polder-Karte für TROLL2 bei einer RMSD^e von 2.0. Weitere Details siehe Text.

Das beste Modell konnte mit dem *S*-Enantiomer von TROLL2 erhalten werden. Hierbei liegt der bromierte Phenylring im hydrophoben Bereich der Bindestelle, der von den Seitenketten der Aminosäuren R113, H115, E167 und E169 gebildet wird. Für eine Interaktionsanalyse wurde die Kristallstruktur mit den benachbarten Kristallkontakten in MOE präpariert und anschließend eine Interaktionsanalyse durchgeführt. Die Ergebnisse hierfür sind in Tabelle 6.22 und Abbildung 6.69 gezeigt. Der Phenylring interagiert mit dem Imidazolring von H183 über eine T-förmige (*edge to face*) Wechselwirkung, sowie über eine CH- π -Wechselwirkung. Zusätzlich bildet der Phenylring über eine Kationen- π -Wechselwirkung eine Interaktion zu R113* des Kristallkontaktes aus. Inwiefern das direkt benachbarte R113 in Lösung diese Kationen- π -Interaktion oder eine π - π Wechselwirkung mit dem Guanidinsystem bilden kann, kann hier nur gemutmaßt werden. MD-Simulationen könnten hier eine eindeutige Antwort liefern.

Auch wenn die Elektronendichten im Bereich des sekundären Alkohols und des primären Amins sehr gering sind und somit mit einer entsprechenden Unsicherheit behaftet sind, sprechen die hier vorgefundenen Interaktionen für diesen Bindemodus. Das protonierte Amin interagiert über eine Salzbrücke mit der Carboxylgruppe von E185. Der Abstand von 3.35 Å (N·O) liegt hierbei in einem typischen Bereich für eine solche Interaktion [307]. Auch wenn die Wasserstoffbrücke zwischen dem sekundären Alkohol von TROLL2 und der Carboxylgruppe von E185 mit einem Abstand von 3.11 Å (O·O) an der oberen Grenze der ansonsten typisch vorgefundenen Abstände liegt [308], kann diese dennoch als eine, wenn auch schwache, Interaktion interpretiert werden. Eine bessere Elektronendichte und somit bessere Lokalisation des Sauerstoffes des Alkohols könnten hier eindeutiger Bindeverhältnisse liefern. Die für Kristallisation und Soaking eingesetzten Fragmentkonzentrationen lagen bei 1.5 mM. Eine höhere Konzentration könnte den Besetzungsgrad eventuell deutlich erhöhen und eindeutiger Elektronendichten zeigen. Das deutliche Abnehmen der Dichten zur Aminogruppe kann sicherlich mit der starken Solvensexponiertheit der Bindestelle erklärt werden. Das Fragment TROLL7, welches sich nur durch das Chloratom anstelle des Bromatoms unterscheidet, sollte ebenfalls an dieser Stelle binden.

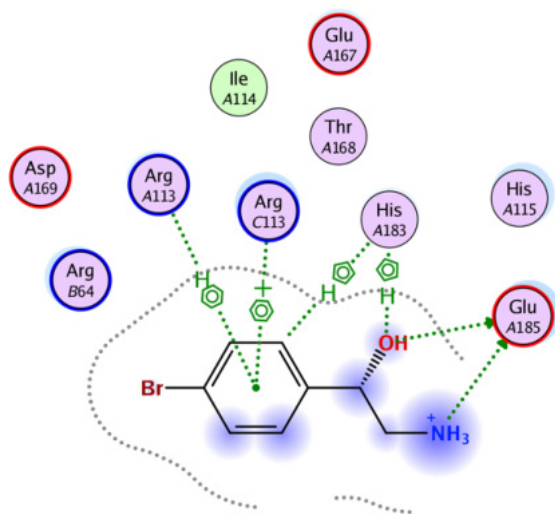


Abbildung 6.69: Interaktionen von TROLL2 mit der p97-N-Domäne

Dargestellt ist eine in MOE-generierte zweidimensionale Darstellung der Interaktionen von TROLL2 mit p97. Aminosäuren, die mit B und C bezeichnet sind, gehören zu symmetrieverwandten N-Domänen. Polare Aminosäuren sind violett, bei positiver Ladung mit einem blauem Kreis, bei negativer Ladung mit einem roten Kreis dargestellt. Lipophile Aminosäuren sind grün dargestellt. Die blauen Dichten um TROLL2 geben einen qualitativen Eindruck der Lösungsmittel-Exponiertheit wieder. Die in grün gezeigten Interaktionen sind in Tabelle 6.22 detailliert beschrieben.

Tabelle 6.22: Interaktionen von TROLL2 mit p97
Ergebnisse der Interaktionsanalyse mit MOE

TROLL2	p97-ND1	Interaktion	Distanz [\AA]
prim. Amin	E185 (A)	Salzbrücke	3.35 (N·O)
sek. Alkohol	E185 (A)	Wasserstoffbrücke	3.11 (O·O)
Phenylgruppe	H183 (A)	CH- π	3.78
Phenylgruppe	R113 (A)	π -CH	4.68
Phenylgruppe	R113 (C)	π -Kation	4.56

Abbildung 6.69 zeigt auch noch einmal, dass das Fragment relativ solvensexponiert ist (blaue Kreisflächen) und nur ein Teil des bromierten Phenylrings in einer Tasche vergraben liegt. Dies unterscheidet das Fragment von den meisten in Kristallstrukturen beobachteten Fragment-Protein-Komplexen, bei denen der Hauptteil des Fragmentes nur noch einen sehr kleinen Lösungsmittel exponierten Anteil besitzen [283]. Dies ist nicht zuletzt durch das Fehlen wirklich tiefer Bindetaschen in der N-Domäne begründet und stellt auch gleichzeitig die große Herausforderung für das Finden möglicher bindender Moleküle dar.

Vergleicht man die vorgefundene mit der in den *mixed-solvent* Simulationen erhaltenen Pose (Abbildung 6.60) für die SHP-II Stelle, so liefert die Simulation eine sehr gute Vorhersage für die im Kristall vorgefundene Struktur. Ob TROLL2 auch, wie in den *mixed-solvent* Simulationen gezeigt, an die SHP-I Region bindet, kann hier leider nicht gesagt werden, da diese durch einen Kristallkontakt größtenteils maskiert und unzugänglich ist. Um diese Problematik zu lösen und die SHP-I Bindestelle zugänglich zu machen sind weitere kristallographische Arbeiten mit einem p97-ND1-Konstrukt mit einem verkürztem N-Terminus geplant. Erste aufgenommene Daten zeigten hier Resultate, die für den Fortgang des Projektes hoffen lassen, die mögliche Bindung von TROLL2, wie auch anderer Fragmente an die SHP-I Bindestelle aufklären zu können.

Aufgrund der starken Solvens-Exponiertheit der möglichen Bindestellen in der N-Domäne und der damit schwer interpretierbaren partiellen Elektronendichten, wäre der bei [309] berichtete Ansatz (PanDDA-Algorithmus) eine weitere mögliche Lösung. Der PanDDA-Algorithmus benutzt mehrere Datensätze von unterschiedlichen Kristallen eines Fragmentes und subtrahiert ein so errechnetes Grundrauschen, um am Ende die partiellen schwachen Elektronendichten von Fragmenten besser sichtbar werden zu lassen. Eine weitere Möglichkeit wäre der gezielte Einsatz von mit Brom substituierten Fragmentanaloga, die wie bei TROLL2 die Lokalisation der Fragmente deutlich erleichtern würden.

6.5 Zusammenfassung und Diskussion des virtuellen Screenings und Fragmentscreenings an p97

Das langfristige Ziel dieses Projekts ist die Entwicklung eines PPI-Inhibitors mit der p97-N Domäne als Zielstruktur. Die vorgestellten Ergebnisse des virtuellen Screenings, des biophysikalischen Screenings, sowie die Ergebnisse der strukturellen Experimente sollen vor diesem Hintergrund diskutiert werden.

6.5.1 Vergleich der Struktur motive

Zunächst soll ein Überblick über die Strukturen der Treffer des p97-N Screenings gegeben werden. Die Bemis-Murcko-Gerüste²² und die vorkommenden aromatischen Systeme der in Tabelle 6.17 aufgeführten Fragmente, sowie SHP9 aus dem virtuellen Screening sind in Abbildung 6.70 gezeigt.

Insgesamt lassen sich die 11 ausgewählten Fragmente in 3 Gruppen aufteilen (Abbildung 6.70): in solche, die nur ein aromatisches System besitzen (1.); in solche, die zwei direkt über eine Bindung verbundene aromatische Systeme besitzen (2.); und in solche, die zwei über mehr als eine Bindung verbundene aromatische Systeme besitzen (3. und 4.). *Giordanetto et al.* [283] untersuchten und klassifizierten in „*Fragment Hits: What do they look like and how do they bind?*“ 489 Kristallstrukturen von Protein-Fragment-Komplexen anhand ihrer Struktur und physikochemischen Eigenschaften. Ein Vergleich mit der dort durchgeführten Bemis-Murcko-Analyse zeigt ähnliche Gerüste und eine ähnliche Verteilung der gefundenen Strukturen.

Die Analyse der vorkommenden aromatischen Systeme zeigt einen Phenylring an erster Stelle, gefolgt von Pyrazol enthaltenden Fragmenten. Daneben finden sich Imidazol und Analoga, sowie Indolsysteme und Analoga. Die bei *Giordanetto et al.* [283] berichteten Pyridin enthaltenden Fragmente fehlen hier gänzlich. Bis auf das Fehlen von Pyridinen folgen die Fragmente den in der Literatur bekannten Strukturen. Eine Auswertung mittels der Tanimoto-Distanz zeigte außer den eng verwandten Molekülen TROLL2 und TROLL7 keine auffallend ähnlichen Fragmente. Es ließen sich aber visuell wiederkehrende strukturelle Merkmale

²²In der Methode nach *Bemis & Murcko* [310] werden Moleküle wahlweise auf ihre Ringsysteme, Linkerstrukturen und Seitenketten reduziert. Hierdurch lässt sich deren Komplexität reduzieren und die Vergleichbarkeit von Molekülen vereinfachen. Für die Auswertung wurde die Stufe der Reduzierung auf ein Molekülgerüst gewählt, welches nur noch Ringsysteme und Linkerstrukturen anzeigt. Die Auswertung wurde mit dem in RDKit implementierten Murcko-Algorithmus durchgeführt [310].

erkennen, wie eine 3-Aminopyrazol-Substruktur, die in TROLL2 und TROLL7 vorkommende 2-Aminophenyl-1-Ethanol-Substruktur, sowie ein primäres Amin bzw. ein primärer Alkohol mit Bindung an ein aromatisches System.

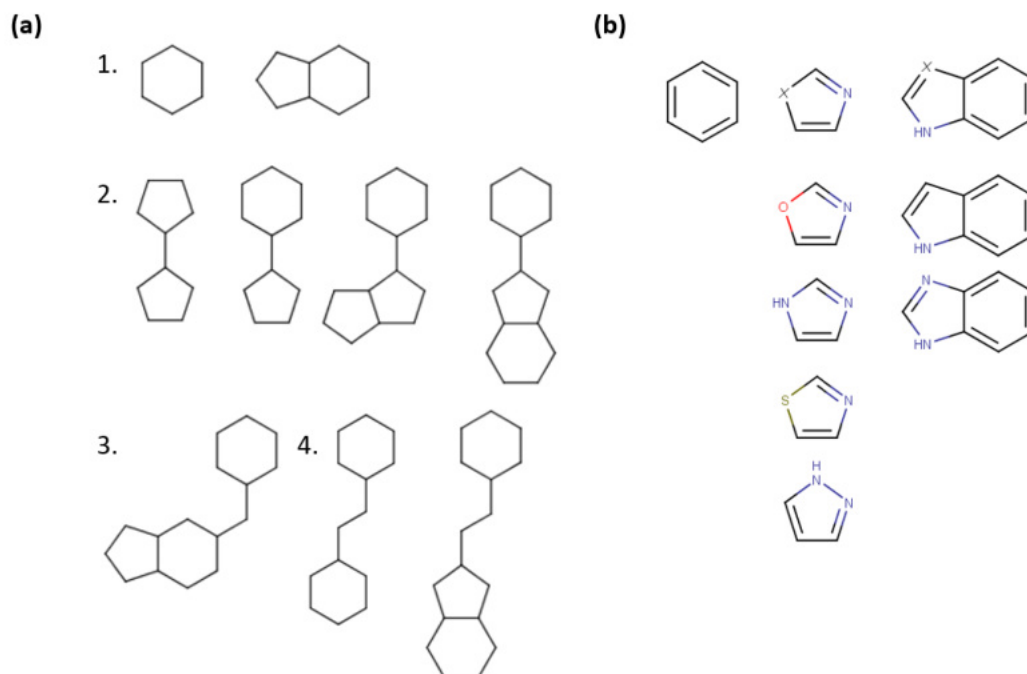


Abbildung 6.70: Strukturmotive der Treffer des p97-N-Screenings
(a) Bemis-Murcko-Analyse und (b) vorgefundene Ringsysteme in den Treffern des N-Domänen Screenings.

Die nächste Frage stellt die Position der Fragmente im chemischen Raum dar. Hierzu wurde die Lage der Treffer im chemischen Raum der UiB Fragmentbibliothek (Abbildung 6.71 (a)), sowie im chemischen Raum der virtuellen Fragmentbibliothek eingezeichnet (Abbildung 6.71 (b)). Zusätzlich wurden die Positionen der bei *Chimenti et al.* [61] berichteten Fragmente markiert.

Abbildung 6.71 (a) zeigt, dass sich die Fragmente nicht zufällig über den chemischen Raum der UiB Bibliothek verteilen, sondern sich fast alle im vierten Quadranten befinden. Die Lage der Fragmente wird hier hauptsächlich durch den $\log P$, sowie der damit korrelierenden Anzahl an aromatischen Atomen und Ringssystemen bestimmt. Der mit einer gelben Ellipse eingezeichnete Bereich enthält 9 der 12 Treffer der in dieser Arbeit durchgeführten Screenings, sowie 5 der 7 bei *Chimenti et al.* berichteten Fragmente. Einzig TROLL2 und TROLL7 liegen in einem anderen Bereich. Die bevorzugt an die p97-N Domäne bindenden Strukturen lassen sich demnach anhand ihrer physikochemischen Eigenschaften und der daraus resultierenden Lage charakterisieren. Des Weiteren zeigt sich, dass die hier durchgeführten Screenings vergleichbare Ergebnisse mit den in der Literatur

bekanntes Fragmentscreenings zeigen, was die Aussagekraft beider Screenings bekräftigt. Inwiefern die Lage im chemischen Raum einer möglichen Bindestelle zugeordnet werden kann, bleibt hier mit nur einer Kristallstruktur von TROLL2 und den NMR-Indizien für SHP9 ungeklärt. Es zeigt sich darüberhinaus, dass die eingesetzte UiB Bibliothek diesen Bereich nicht optimal mit Fragmenten abdeckt. Fast die Hälfte der Fläche innerhalb der Ellipse ist hier durch kein Fragment besetzt und nur der obere Teil der Ellipse zeigt eine Abdeckung mit 5 und mehr Fragmenten pro Hexagon. Es wäre hier sicherlich interessant, ob eine gezielte Anreicherung dieses Bereichs mehr Treffer zu Tage fördern würde, als eine zufällige Auswahl an Fragmenten. Die eingesetzte virtuelle Bibliothek (Abbildung 6.71 **(b)**) deckt diesen Bereich um einiges besser ab und könnte daher für weitere Untersuchungen herangezogen werden.

6.5.2 Dreidimensionale Strukturanalyse

Zur weiteren Aufklärung von eventuellen Gemeinsamkeiten der Fragmente wurden mittels der *flexiblen Aligment*-Funktion in MOE ausgewählte Fragmente überlagert. Die überlagerten Strukturen wurden anschließend mit der *Consensus-Pharmakophorsuche*-Funktion (Suchradius: 1.5 Å) in MOE weiter analysiert. Auch hier ließen sich Gemeinsamkeiten unter den Fragmenten dieser Arbeit, wie auch mit den bei *Chimenti et al.* berichteten Fragmenten, finden.

Abbildung 6.72 zeigt die mittels der *Consensus-Pharmakophorsuche* erstellten Pharmakophormodelle. Die Überlagerungen von TROLL14 mit ID4^{Chimenti} **(a)**, TROLL12 **(b)** und VIK20 **(c)** zeigen ähnliche funktionelle Gruppen. Alle zeigen zwei aromatische Funktionen in Form eines Phenylrings, sowie eines weiteren Heterozyklus. Zwei der vier Moleküle enthalten eine zusätzliche hydrophobe Gruppe in *para*-Stellung am Phenylring. Alle Fragmente besitzen einen Wasserstoffbrückendonator in Form einer primären Hydroxyl- bzw. Aminfunktion am Heterozyklus. Zusätzlich zeigen die Fragmente ein an ein „L“ erinnerndes Erscheinungsbild der Solvens-zugänglichen Moleküloberfläche **(g)**.

Eine Auswertung der gefundenen 3-Aminopyrazol-Strukturen zeigt, dass sich nur TROLL18 mit ID2^{Chimenti} **(d)** mit guter Übereinstimmung überlagern lässt. Die Aminfunktion von TROLL8 kann hier nicht zur Deckung gebracht werden. Dagegen zeigt TROLL8 mit TROLL2, (TROLL7) eine Übereinstimmung von ähnlichen Eigenschaften an den gleichen Positionen **(e)**. Das Amin von TROLL8 und das Amin von TROLL2, (TROLL7) können mit einem Wasserstoffbrückendonator beschrieben werden. Auch zeigen der Wasserstoffbrückendonator bzw. -akzeptor

des Pyrazolrings und des sekundären Alkohols eine gute Übereinstimmung. Beide Fragmente enthalten darüber hinaus aromatische bzw. hydrophobe Gruppen an den gleichen Positionen. Alle Moleküle weisen eine längliche Form der Solvens-zugänglichen Moleküloberfläche auf (**h**). TROLL6 und ID3^{Chimenti} zeigen ebenfalls ähnliche Eigenschaften in Form einer aromatischen Gruppe und einer anionischen Gruppe (**f**), zusammen mit einer vergleichbaren Solvens-zugänglichen Moleküloberfläche (**i**).

Alle Moleküle liegen nicht in direkter Nachbarschaft im chemischen Raum. Die aufgestellten Pharmakophormodelle werden zudem nicht gänzlich von den Ergebnissen der *mixed-solvent* Simulationen gestützt, da die Fragmente hier teilweise nicht an die gleichen Bindestellen binden. TROLL8 bindet nicht wie TROLL2 bzw. TROLL7 in die SHP-II Bindestelle. Dagegen bindet TROLL12 an die SHP-II Bindestelle. Die in der Kristallstruktur gefundenen Interaktionen von TROLL2 befinden sich dagegen mit dem aufgestellten Pharmakophormodell in Einklang. Trotz der genannten Diskrepanzen zwischen den Pharmakophormodellen und den Ergebnissen der *mixed-solvent* Simulationen könnten die Modelle für weitere Pharmakophor-Screenings eingesetzt werden, um gezielt neue Bibliotheken für weitere computerbasierte oder experimentelle Untersuchungen zu erstellen. Auf diesem Wege neu gefundene Liganden an p97-N oder die kristallographisch bestimmten Bindeposen der bereits gefundenen Moleküle könnten die hier aufgestellten Thesen weiter untermauern und verfeinern.

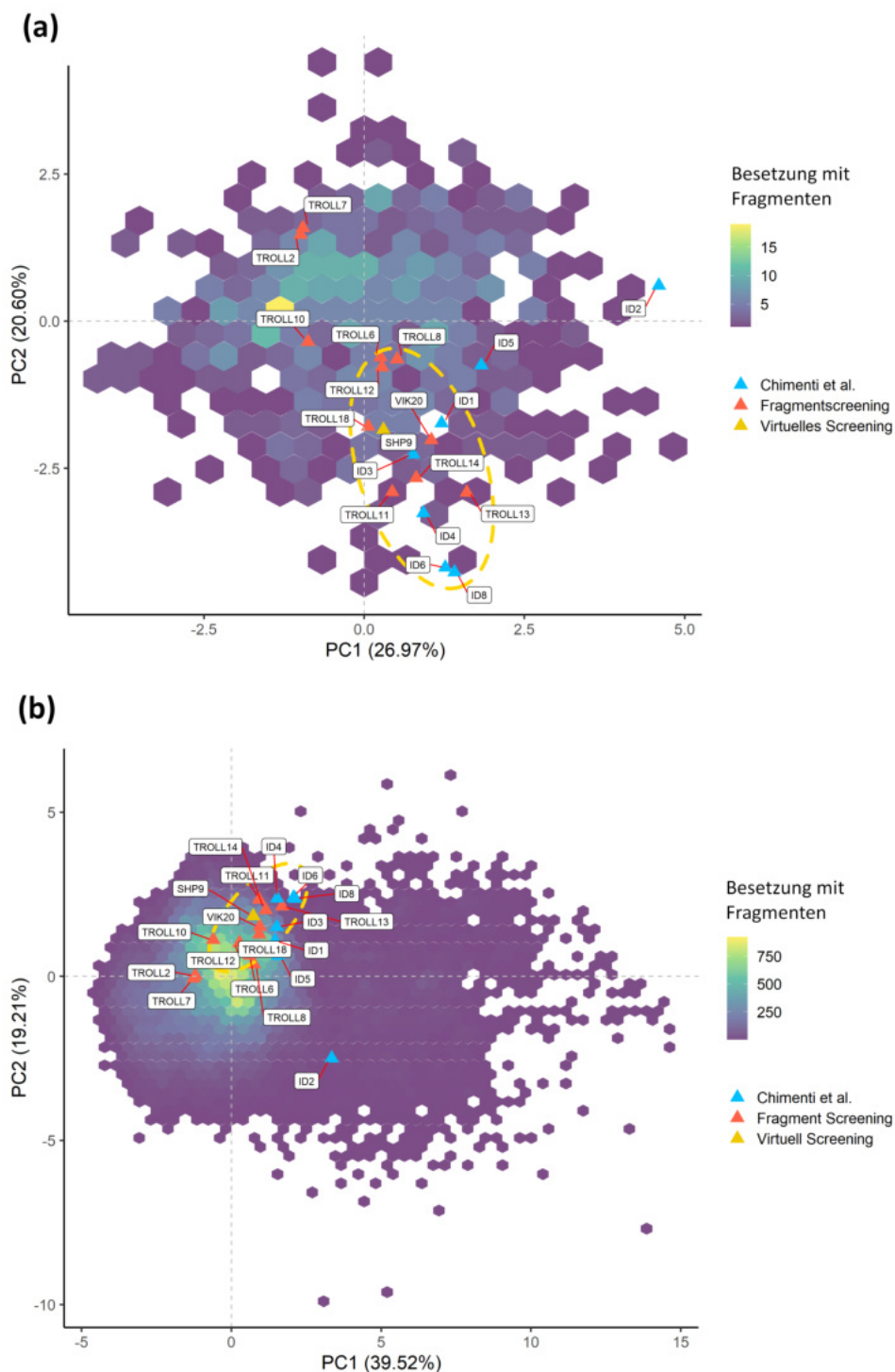


Abbildung 6.71: PCA der an p97-N bindenden Fragmente und deren Lage im Bezug zu benutzten Fragmentbibliotheken

(a) Lage innerhalb der UiB Fragmentbibliothek. (b) Lage innerhalb der virtuellen Fragmentbibliothek. In Rot sind alle Treffer des BLI-Screenings gezeigt, in Blau die bei *Chimenti et al.* berichteten Fragmente. Gelb zeigt den Treffer des virtuellen Screenings. Die gezeigten Hexagonale zeigen an, mit wie vielen Fragmenten der Bereich des chemischen Raums jeweils besetzt ist. Für weitere Erläuterungen siehe Text.

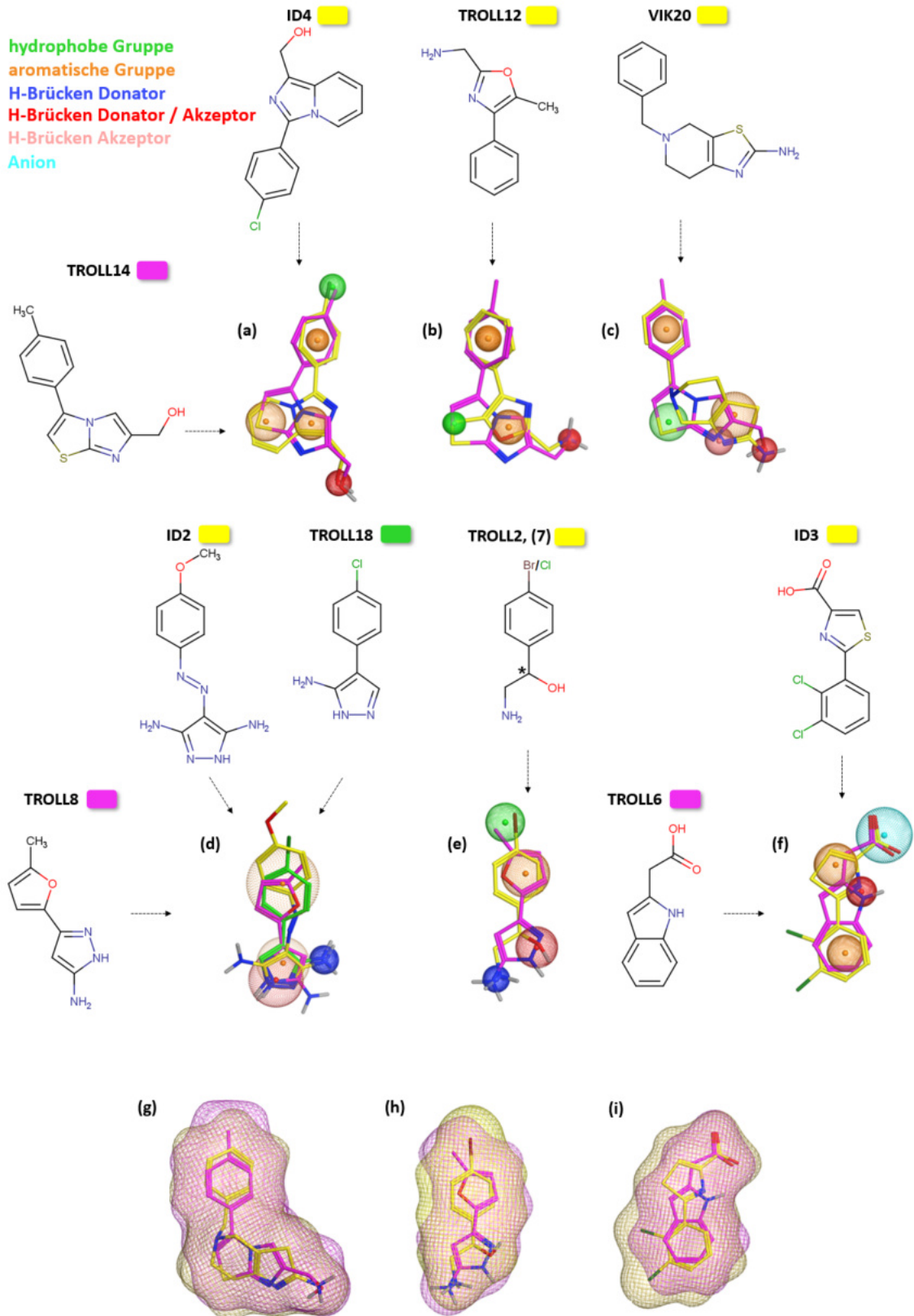


Abbildung 6.72: Ligandbasierte Pharmakophore Modelle

Anhand der Fragmente erstellte ligandbasierte Pharmakophormodelle. Für eine detaillierte Beschreibung der Modelle siehe Text.

6.5.3 Ligandeffizienz-Indizes

Im Folgenden sollen die ausgewählten Fragmente auf der Basis ihrer errechneten Ligandeffizienz-Indizes eingeordnet und mit den in der Literatur berichteten Werten für Fragmentscreenings verglichen werden. Die Grundlage der Berechnung der Indizes bilden die Werte $K_D^{Langmuir}$ für die N-Domäne, die Schweratomanzahl HA , sowie der $clogP_{MOE}$ -Wert jedes Fragments.

Tabelle 6.23 zeigt die erhaltenen Mittelwerte für das hier durchgeführte (**fett**) und die bei *Chimenti et al.* (in Klammern und *kursiv*) berichteten Fragmentscreenings mit der N-Domäne. Gegenübergestellt sind sie den in der Literatur referierten Mittelwerten von erfolgreichen Fragmentscreening-Kampagnen. Die (*kursiv*) gezeigten Werte geben hier die Werte der aus den Fragmenten erhaltenen, optimierten Leitstrukturen wieder und zeigen, in welche Richtung sich die Indizes bei der Optimierung entwickelten. Die Mittelwerte stammen aus einer 2013 von György F. & György K. [311] veröffentlichten retrospektiven Analyse von 145 berichteten fragmentbasierten Studien an 83 Zielproteinen.

Die in dieser Arbeit gefundenen Verbindungen zeigen in großer Übereinstimmung Mittelwerte wie die von György F. & György K. berichteten Fragmentscreenings, während die Indizes für das von *Chimenti et al.* berichtete Screening deutlich abweichen. Es lässt sich festhalten, dass das hier durchgeführte Screening als erfolgreich eingestuft werden kann und Verbindungen enthält, die anhand ihrer Indizes in einem sehr guten Bereich zur weiteren Optimierung liegen. Das von *Chimenti et al.* berichtete Screening schneidet deutlich schlechter ab. Anhand der Indizes lässt sich sehen, dass die dort berichteten Fragmente deutlich lipophiler sind, was für eine Optimierung schlechtere Voraussetzungen darstellt (siehe die Indizes LELP oder LLE_{AT}). Dies deckt sich auch mit der Lage der Fragmente von *Chimenti et al.* im lipophileren Bereich des chemischen Raums.

Welche Fragmente im Einzelnen nun für eine weitere Optimierung günstige Indizes zeigen, beantwortet Tabelle 6.24. Die Fragmente TROLL2, 7, 8, 10²³ und 12 weisen einen LLE_{AT} und LE von größer 3.0 auf. Auch die LELP- bzw. LLE-Werte liegen in einem günstigen Startbereich, um die Fragmente mit noch weiteren lipophilen Gruppen vergrößern zu können. Die genannten Fragmente zeigten auch in den *mixed-solvent* Simulationen (bis auf TROLL10, das aufgrund der nicht vorhandenen STD-Signale nicht simuliert wurde) die besten Resultate. Die dort identifizierten Bindestellen zusammen mit der für TROLL2 zusätzlich vorliegenden

²³TROLL10 wurde, trotz des fehlenden STD-Signals, hier in die Analyse mit aufgenommen.

Kristallstruktur, machen hier ein rationales Optimieren möglich.

Tabelle 6.25 zeigt im Einzelnen die Indizes der von *Chimenti et al.* berichteten Fragmente. Nur das Fragment ID5 weist akzeptable Werte auf. Interessant sind hier auch die Resultate für XN, das mit einer Molekülgröße von 354 Da bereits im Bereich von Leitstrukturen liegt. XN zeigt im Verhältnis zu seiner Affinität bereits eine hohe Lipophilie und somit suboptimale Werte für LLE und LELP im Vergleich zu den Mittelwerten der optimierten Fragmente. Auch der molekülgrößenunabhängige SILE-Index ergibt einen Wert von nur 1.96 für XN im Vergleich zu 2.6 für optimierte Leitstrukturen. Anhand dieser Werte lässt sich sagen, dass eine weitere Optimierung von XN, vor allem durch eine Einführung polarer Gruppen erfolgen müsste, um der bereits hohen Lipophilie entgegenzusteuern. Ein solcher Weg der Optimierung ist allerdings oftmals schwierig und die Einführung von zusätzlichen polaren Gruppen kann leicht zu einem Aktivitätsverlust führen [312, 313].

Tabelle 6.23: Vergleich der Mittelwerte verschiedener LE Indizes der ausgewählten Fragmente mit in der Literatur berichteten Fragmentenscreenings. *Mittelwert optimiert^{Lit.}* gibt die Mittelwerte der aus optimierten Fragmenten erhaltenen Leitstrukturen an. **Fett** gedruckt sind die Mittelwerte des in dieser Arbeit durchgeführten Fragmentenscreenings, *kursiv* gedruckt sind die Mittelwerte des bei *Chimenti et al.* [61] berichteten Fragmentenscreenings.

Deskriptor	Mittelwert Fragmente ^{Lit.}	<i>Mittelwert optimiert^{Lit.}</i>	Mittelwert p97-N Screening
MW	211	389	202, (243)
HA	15	28	15, (17)
pK _D /pIC ₅₀	4.1	7.2	3.5, (3.0)
logP	1.8	3.2	1.75, (2.4)
LE	0.37	<i>0.34</i>	0.35, (0.26)
SILE	1.8	2.6	1.64, (1.36)
LLE	2.1	4.1	1.86, (0.73)
LELP	4.6	9	5.43, (9.68)
LLE _{AT}	0.32	0.3	0.3, (0.17)

Tabelle 6.24: Indizes ausgewählter Fragmente des N-Domänen Screenings

ID	$clogP$	MW	HA	K_D^a	ΔG^b	LE	SILE	LLE	LLE _{AT}	LLEP
TROLL2	1.39	216	11	56	5.75	0.52	2.07	2.86	0.47	2.67
TROLL6	2.04	175	13	160	5.13	0.39	1.76	1.76	0.30	5.23
TROLL7	1.18	172	11	223	4.93	0.45	1.78	2.47	0.42	2.62
TROLL8	1.09	200	12	245	4.88	0.41	1.71	2.53	0.40	2.65
TROLL10	0.96	161	12	276	4.81	0.40	1.69	2.59	0.41	2.42
TROLL11	2.08	223	17	323	4.72	0.28	1.49	1.41	0.22	7.44
TROLL12	1.09	118	14	354	4.66	0.33	1.56	2.36	0.34	3.31
TROLL13	2.42	252	19	409	4.58	0.24	1.40	0.97	0.18	10.09
TROLL14	2.29	244	17	409	4.58	0.27	1.45	1.10	0.20	8.46
TROLL18	2.78	194	13	626	4.33	0.33	1.48	0.42	0.16	8.44
VIK20	1.80	245	17	147	5.18	0.30	1.64	2.03	0.27	5.99
SHP9	1.92	220	15	208	4.98	0.33	1.63	1.76	0.27	5.82
Mittelwert	1.61	200	15	370	4.82	0.35	1.61	1.95	0.31	5.11

^a K_D in μM ^b ΔG in kcal mol^{-1}

Tabelle 6.25: Indizes der bei *Chimenti et al.* berichteten Fragmenten, sowie von XN

ID	$\log P_c$	MW	HA	K_D^a	ΔG^b	LE	SILE	LLE	LLE _{AT}	LLEP
Xanthohumol										
XN	5.10	354	26	8	6.89	0.27	1.92	-0.003	0.11	18.89
Fragmente <i>Chimenti et al.</i>										
ID1	2.20	239	16	890	4.12	0.26	1.33	0.85	0.18	8.45
ID2	1.76	232	17	1300	3.90	0.23	1.23	1.13	0.20	7.66
ID3	3.33	274	16	1960	3.66	0.23	1.18	-0.62	0.06	14.48
ID4	2.70	259	18	250	4.87	0.27	1.51	0.90	0.18	10.01
ID5	1.61	233	16	241	4.89	0.31	1.57	2.01	0.28	5.19
ID8	2.94	222	17	840	4.16	0.24	1.31	0.13	0.12	12.26
Mittelwert ^{ohne XN}	2.42	243	17	914	4.27	0.26	1.36	0.73	0.17	9.68

^a K_D in μM ^b ΔG in kcal mol^{-1}

6.5.4 Relevanz der Fragmente für einen möglichen PPI-Inhibitor der SHP-Bindestelle

Welche Relevanz besitzen die gefundenen Fragmente für die Entwicklung eines möglichen PPI-Inhibitors mit Angriff an der SHP-Bindestelle? Abbildung 6.73 zeigt die Bindung des SHP-Motivs von UFD1 (schwarz) und dazu überlagert die gefundenen Bindestellen aus den *mixed-solvent* Simulationen, sowie der Kristallstruktur von TROLL2. Wie zu sehen ist, fallen die Bindestellen bis auf jene von TROLL8 mit der UFD1-Bindestelle zusammen. In der SHP-I Region wurden in den *mixed-solvent* Simulationen im Bereich der nativen L235^{UFD1}-Binderegion die aromatischen Systeme von TROLL2 und 7 gefunden. SHP9 interagiert über eine Wasserstoffbrücke und $\pi - \pi$ -Wechselwirkungen analog der Aminosäure N233^{UFD1} mit p97-N. In der SHP-II Bindestelle stellt die Bindung der aromatischen Systeme der Fragmente eine Imitation der Bindung von F228^{UFD1} dar.

Die gefundenen Fragmente TROLL2, 7,12 und SHP9 liegen in identifizierten *hot-spot* Bereichen der p97-N-SHP-Bindung. Basierend auf der Tatsache, dass die Fragmente genau an die Regionen binden, an denen auch die für die Bindung wichtigen Bindestellen der Aminosäuren L235^{UFD1} und F228^{UFD1} liegen (siehe Abschnitt 6.1.2), zeigt, dass die gefundenen Verbindungen eine hohe Relevanz für die Entwicklung eines PPI-Inhibitors besitzen. Auch TROLL8 stellt insofern eine interessante Verbindung dar, als dass es in der unmittelbaren Umgebung der UFD1-Bindestelle liegt. Ein Wachsen in Richtung oder Verknüpfen mit den anderen gefundenen Fragmenten stellt hier eine lohnende Strategie für TROLL8 dar.

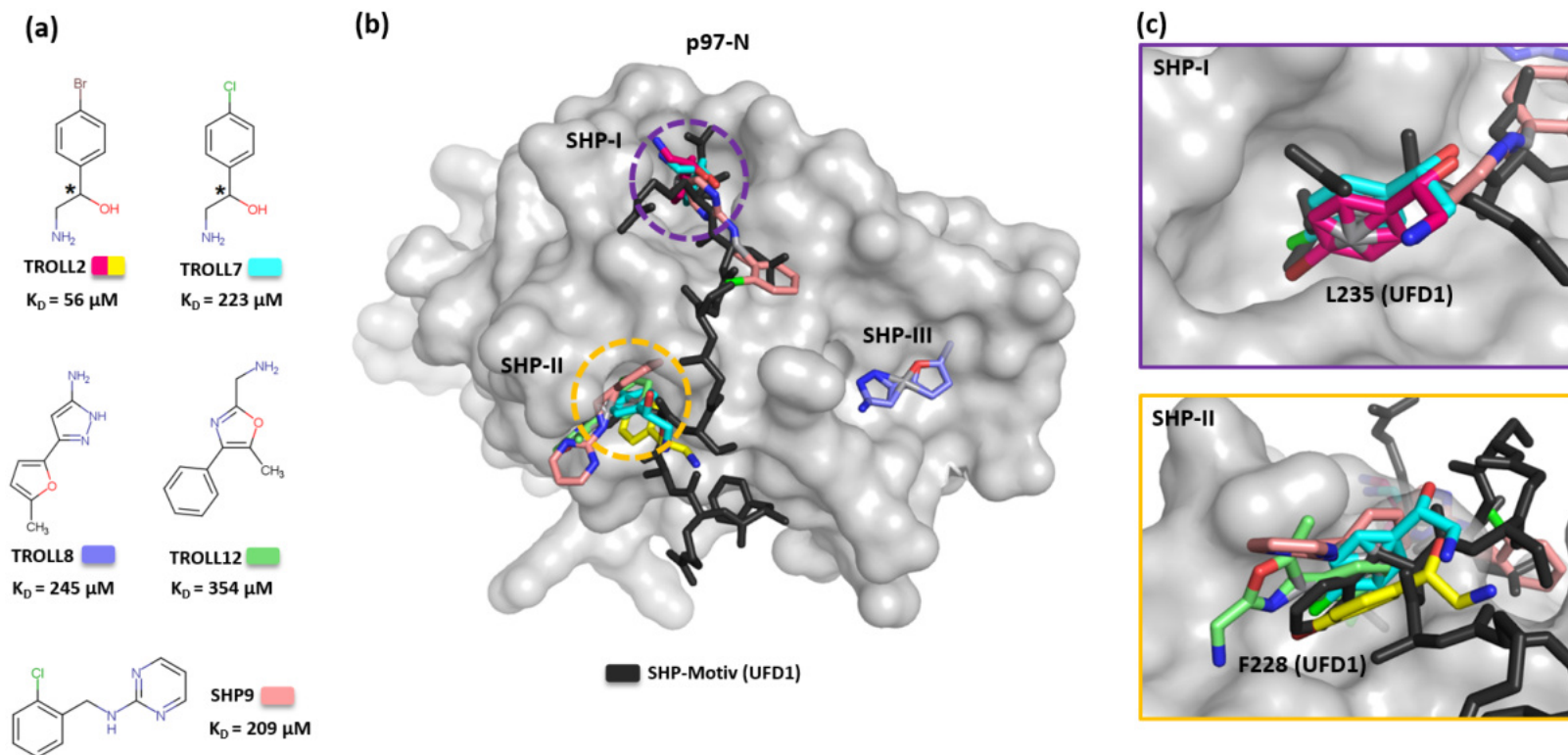


Abbildung 6.73: Bindestellen der Fragmente in Bezug auf die Bindung des SHP-Motivs von UFD1

(a) In dieser Arbeit identifizierte Fragmente die eine Relevanz für die Entwicklung eines PPI-Inhibitors mit Angriff an der SHP-Bindestelle an p97-N besitzen und deren Affinitäten. (b) Überlagerung der Kristallstruktur 5B6C mit den in den *mixed-solvent* Simulationen und der Röntgenstrukturanalyse gefundenen Bindestellen der Fragmente. Für TROLL2 ist die SHP-II Bindepote der Kristallstruktur gelb und die gefundene SHP-I Bindepote in den *mixed-solvent* Simulationen violett gezeigt. Alle Fragmente außer TROLL8 interagieren mit den gleichen *hot-spots* mit denen auch das SHP-Motiv von UFD1 wechselwirkt. TROLL8 liegt in unmittelbarer Nachbarschaft und könnte durch entsprechendes Wachsen ebenfalls zu einem interessanten Molekül optimiert werden. (c) Detailansicht der SHP-I bzw. SHP-II Bindestellen. Die hydrophoben Gruppen von TROLL2 und 7 imitieren die Wechselwirkung von L235^{UFD1} in der SHP-I Region. Alle Fragmente in der SHP-II Bindestelle imitieren die aromatische Wechselwirkung von L228^{UFD1}.

Kapitel 7

Zusammenfassung & Ausblick

7.1 Zusammenfassung

Die in Kapitel 3 genannten Ziele dieser Arbeit konnten weitestgehend erreicht werden. Basierend auf der Analyse der Komplex-Kristallstrukturen der SHP-Motive von UFD1 (PDB: 5B6C) und Derlin-1 (PDB: 5GLF) mit p97-N wurde ein Protokoll für ein virtuelles Screening erstellt, mit dessen Hilfe zehn Verbindungen ausgewählt und experimentell untersucht werden konnten. Zwei der zehn Verbindungen zeigten in BLI- und STD-NMR-Messungen eine Bindung an p97. Für eines dieser Fragmente (SHP9) konnte die Bindung in der favorisierten SHP-I Binderegion durch STD-NMR-Messungen an p97-Mutanten, sowie durch *mixed-solvent* Simulationen untermauert werden. Des Weiteren wurden BLI- und STD-NMR-Assays zur Detektion der Bindung von Fragmenten an p97 erfolgreich etabliert. Das durchgeführte Fragment-Screening mit p97-ND1 zeigte einen sehr guten Z-Faktor von 0.67, und ermöglichte die Festlegung der Screening-Bedingung für die Messungen an der isolierten N-Domäne. 651 Fragmenten wurden mit der N-Domäne vermessen, wovon 19 Verbindungen eine Konzentrationsabhängigkeit der Signale zeigten. Durch die detaillierte Analyse der Sensorgramme und die weitere Vermessung im orthogonalen STD-NMR-Assay wurden hiervon 11 dieser Verbindungen favorisiert.

Eine durchgeführte PCA der 11 Fragmente basierend auf ausgewählten physikochemischen Deskriptoren zeigte, dass die hier gefundenen Verbindungen in dem fast gleichen chemischen Raum zu finden sind, wie die bei *Chimenti et al.* [61] berichteten Fragmente. Des Weiteren liegen die Mittelwerte unterschiedlicher Ligandindizes in den gleichen Bereichen, wie die in der Literatur berichteten Mittelwerte von erfolgreich durchgeführten Fragmentenscreenings. Darüber hinaus zeigen die gefundenen Verbindungen deutlich bessere Startvoraussetzungen für eine weitere Optimierung als die bei *Chimenti et al.* berichteten Fragmente. In

daraufhin durchgeführten *mixed-solvent* Simulationen konnten für vier Fragmente (TROLL2, 7, 8, 12) mögliche Bindestellen und Bindeposen an p97-N identifiziert werden und zeigten in unterschiedlichen Ligandeffizienz-Indizes die besten Werte für eine mögliche Optimierung, sowie Affinitäten zwischen 54 μM und 354 μM .

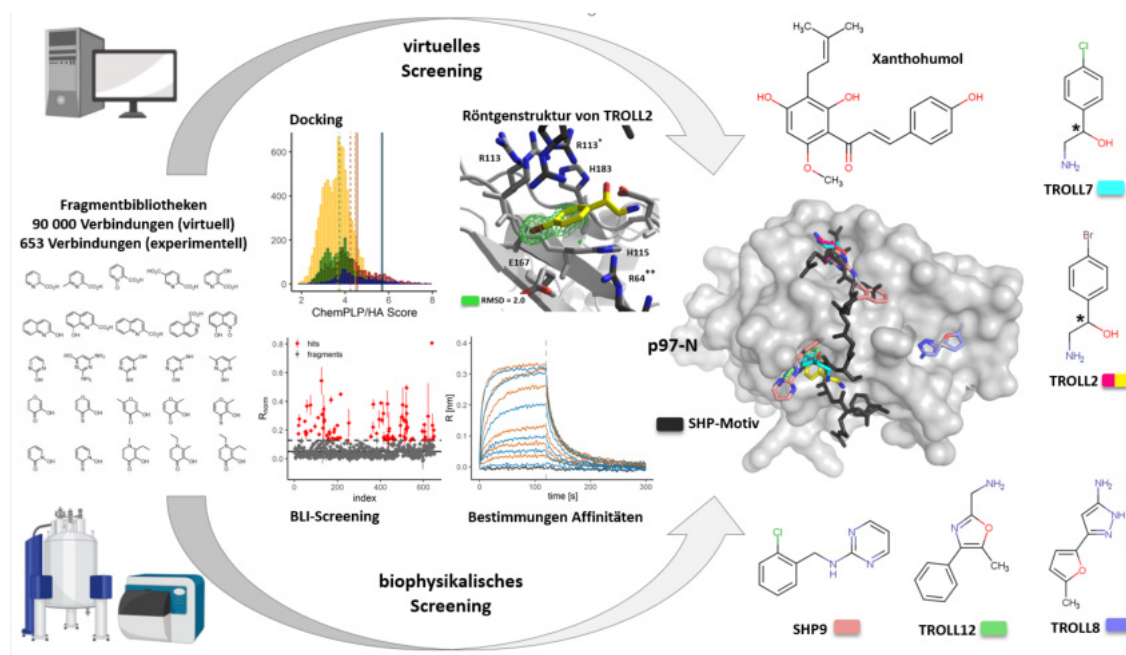


Abbildung 7.1: Grafische Zusammenfassung der Arbeit

In dieser Arbeit wurde ein fragmentbasierter Ansatz zur Identifizierung von Startstrukturen für einen zukünftigen PPI-Inhibitor mit Angriff an der SHP-Bindestelle von p97-N verwendet. Hierzu wurde ein virtuelles und ein biophysikalisches Screening von Fragmentbibliotheken durchgeführt. Mögliche Bindestellen wurden computerbasiert und experimentell bestimmt. Fünf bindende Fragmente konnten mit Affinitäten zwischen 54 und 354 μM identifiziert werden. Die Bindestellen der Fragmente wurden durch *mixed solvent* Simulationen theoretisch vorhergesagt und konnten für TROLL2 kristallographisch und für SHP9 durch STD-NMR Messungen bestätigt werden. Darüber hinaus konnte die Affinität des bereits bekannten Naturstoffs Xanthohumol mit 8 μM bestimmt werden.

Es konnte eine reproduzierbare Kristallisationsbedingung für ein p97-ND1^{L198W}-Konstrukt etabliert werden, die zukünftige kristallographische Analysen und Screenings ermöglicht. Mit diesen Bedingungen war es möglich, die Kristallstruktur des Fragments TROLL2 in Komplex mit p97 zu erhalten. Die in der Kristallstruktur gefundene Bindestelle sowie der Bindemodus stimmten im hohen Maße mit denen der *mixed-solvent* Simulation überein. Die anhand der STD-NMR-Messungen, der Röntgenstrukturanalysen und der *mixed-solvent* Simulation erhaltenen Bindestellen der fünf Fragmente (TROLL2, 7, 8, 12 und SHP9) liegen in oder in direkter Nachbarschaft zu der Bindestelle des SHP-Motivs und stellen somit potentielle Startstrukturen für die weitere Entwicklung eines PPI-Inhibitors

dar. Darüber hinaus konnte die Bindung von Xanthohumol an die N-Domäne gezeigt und deren Affinität mit einem K_D -Wert rund 8 μM bestimmt werden.

7.2 Ausblick

Für das weitere Vorgehen zur Entwicklung eines PPI Inhibitors mit Angriff an der SHP-Bindestelle steht die experimentelle Bestätigung der in den *mixed-solvent* Simulationen gefundenen Bindestellen an erster Stelle. Dies bildet die Grundlage für das weitere rationale Design der Fragmente. Hierzu sind kristallographische Untersuchungen sowie ^1H - ^{15}N -HSQC-Experimente mit der N-Domäne in Planung und bereits in Arbeit. Eingehende thermodynamische Analysen der Fragmente in ITC-Experimenten und kinetische BLI-Messungen sollten zur weiteren Charakterisierung der ausgewählten Fragmente durchgeführt werden. Zusätzlich sollten Studien zu Strukturwirkungsbeziehungen (SAR) an den erfolgversprechendsten Fragmenten begonnen werden. Diese wären anhand der Datelage dieser Arbeit die Fragmente TROLL2, 7, 8, 12 und SHP9. Grundlage solcher SAR-Studien können die in dieser Arbeit aufgestellten Pharmakophormodelle sowie Analysen des chemischen Raums bilden, um entsprechende Substanzbibliotheken zusammenzustellen. Des Weiteren sollten auf Basis der *mixed-solvent* Simulationen und zukünftiger Kristallstrukturen synthetische Wege zur Verknüpfung und zum Ausbau dieser Fragmente gefunden werden. Ähnliche Vorgehensweisen gelten auch für Xanthohumol. Hier bedarf es ebenfalls der Aufklärung der Bindestelle durch kristallographische oder NMR-Untersuchungen, die parallel von SAR-Studien unterstützt werden sollten. In *in vitro* Kompetitionsassays sollten optimierte Verbindungen auf ihre Funktionalität hin untersucht werden, um sicherzustellen, dass das Ziel der Inhibierung der Kofaktorbinding an p97 erreicht wurde. Die Etablierung eines solchen Assays ist ein zentraler nächster Schritt dieses Projekts und ermöglicht nach erfolgreichem Ausgang, die Verbindungen in zellbasierten Assays auf Zytotoxizität in gesunden Zellen und Inhibition des Wachstums entarteter Zellen hin zu untersuchen.

Anhang A

Verteilung unterschiedlicher Deskriptoren der eingesetzten Fragmentbibliotheken

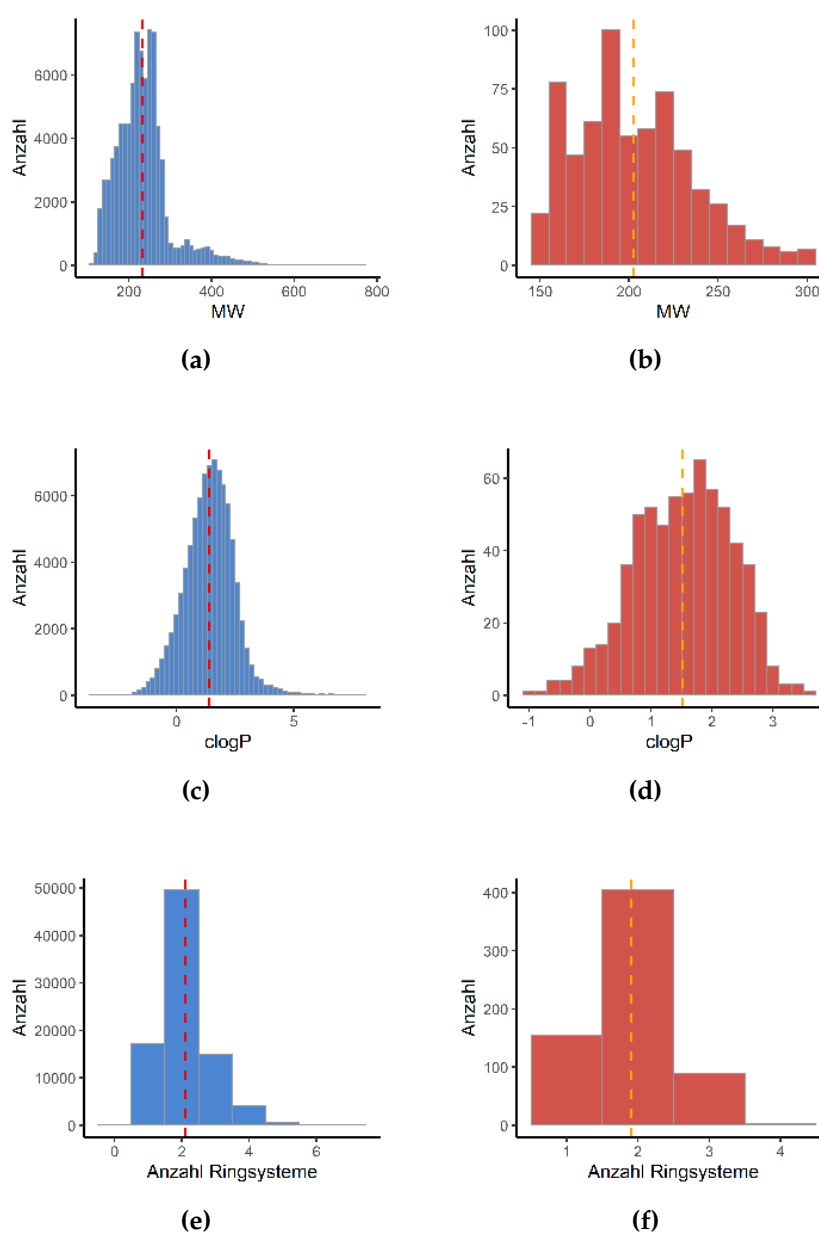


Abbildung A.1: Verteilung verschiedener Moleküldeskriptoren. Virtuelle Fragment Bibliothek (blau), reale Fragment Bibliothek (rot) mit den jeweiligen eingezeichneten Mittelwerten.

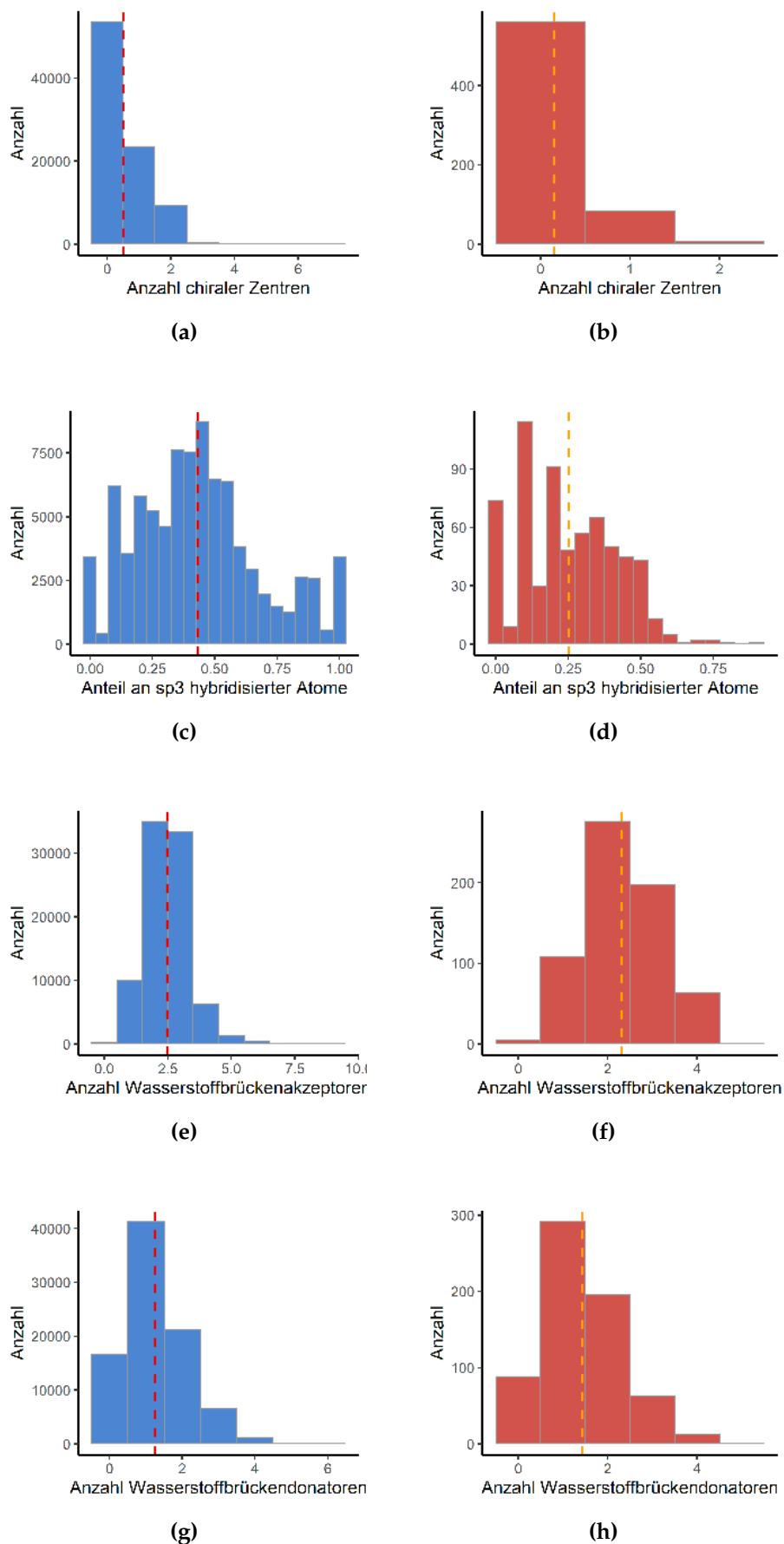


Abbildung A.2: Verteilung verschiedener Moleküldeskriptoren. Virtuelle Fragment Bibliothek (blau), reale Fragment Bibliothek (rot) mit den jeweiligen eingezeichneten Mittelwerten.

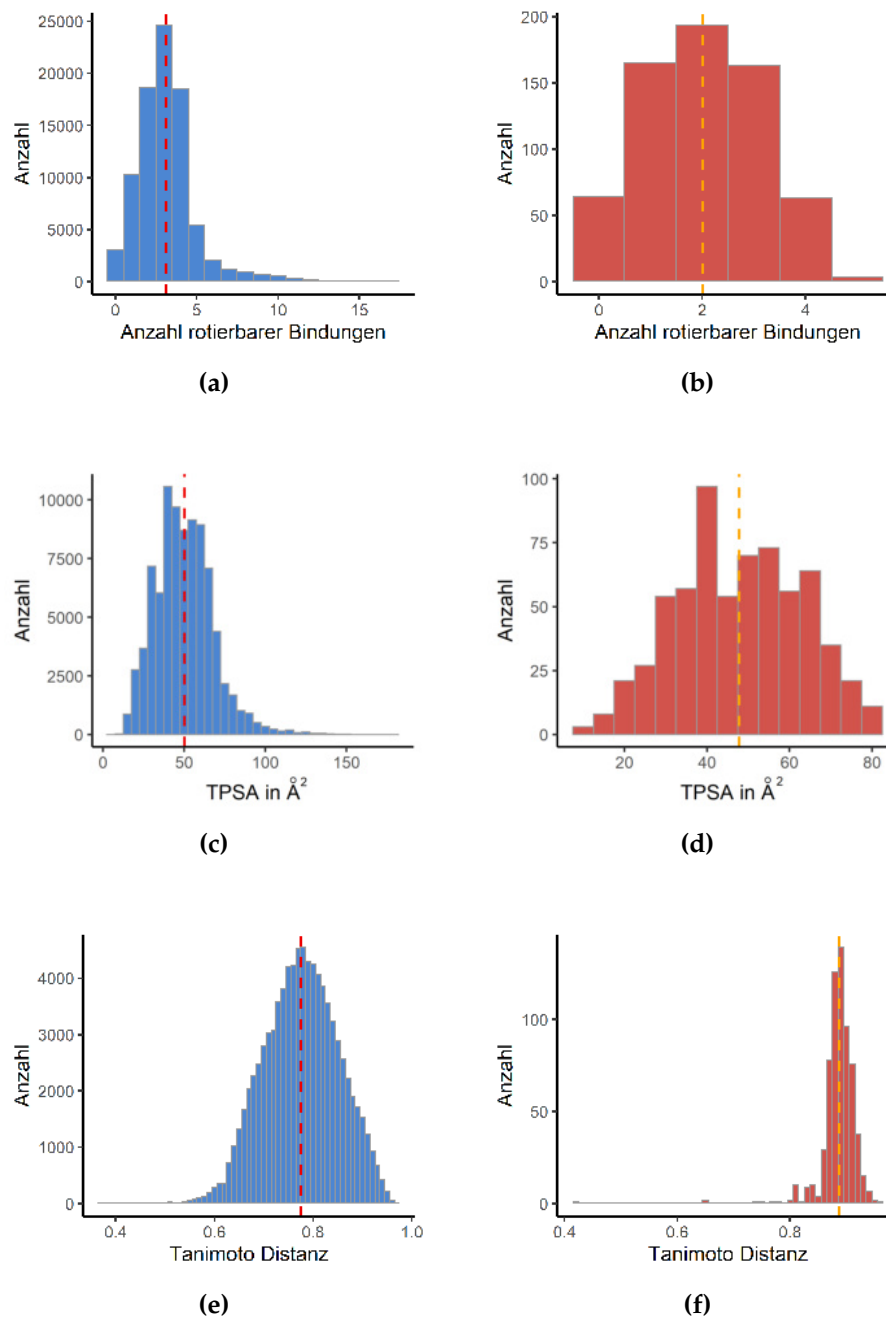


Abbildung A.3: Verteilung verschiedener Moleküldeskriptoren. Virtuelle Fragment Bibliothek (blau), reale Fragment Bibliothek (rot) mit den jeweiligen eingezeichneten Mittelwerten.

Anhang B

Docking und MD-Posen der für die experimentelle Validierung ausgewählten Fragmente des virtuellen Screenings an der N-Domäne

Gezeigt sind in den folgenden Abbildungen die jeweilige Dockingpose, die vorgefundenen MD-Posen, sowie der zeitliche Verlauf der RMSD-Werte, der für die experimentellen Untersuchungen ausgewählten Fragmente des virtuellen Screenings.

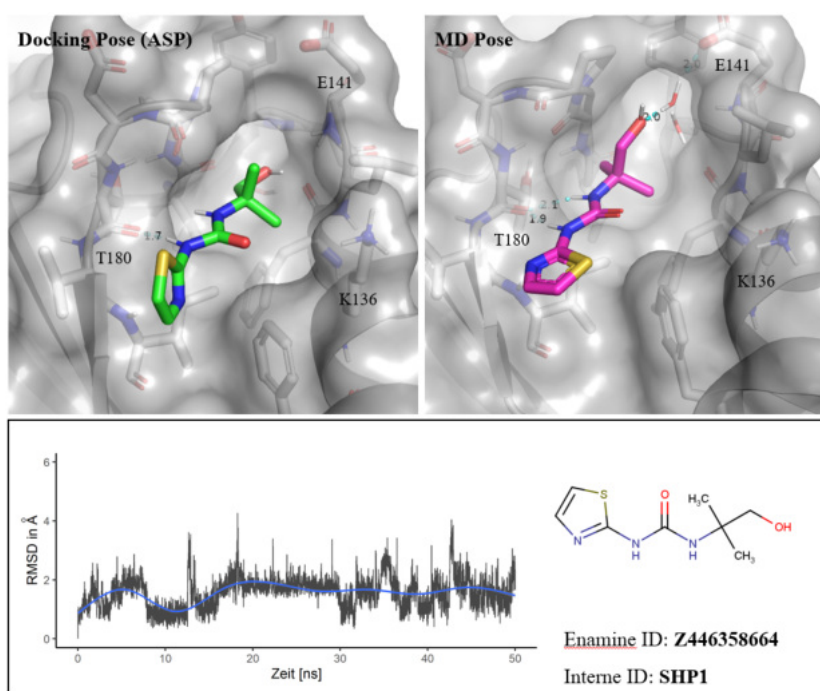


Abbildung B.1: Docking und MD-Profil der Verbindung SHP1

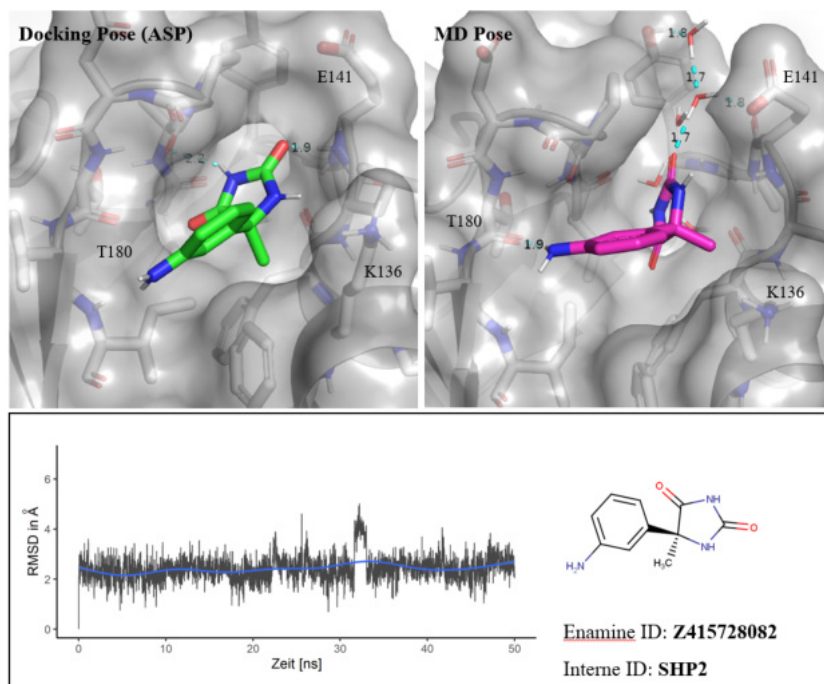


Abbildung B.2: Docking und MD-Profil der Verbindung SHP2

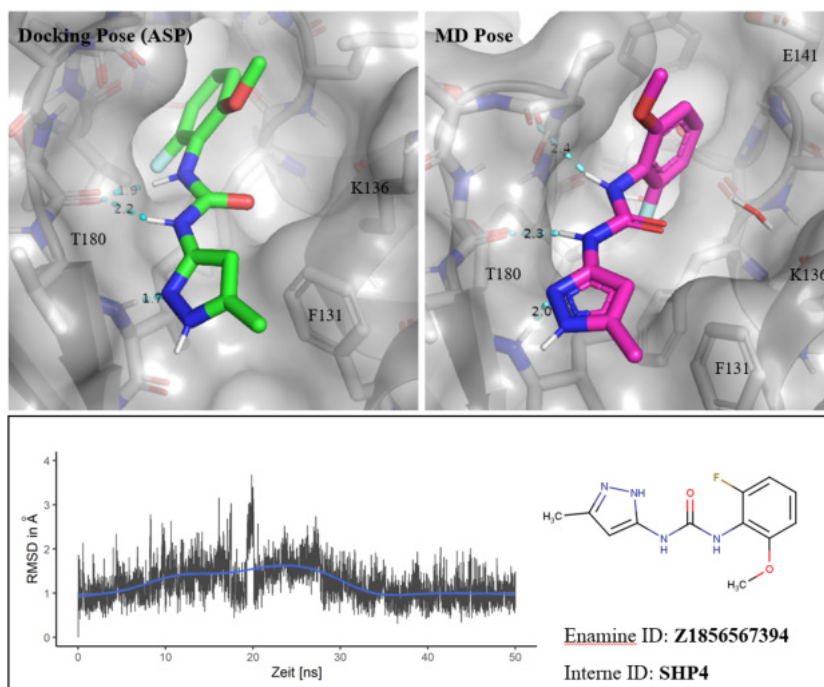


Abbildung B.3: Docking und MD-Profil der Verbindung SHP4

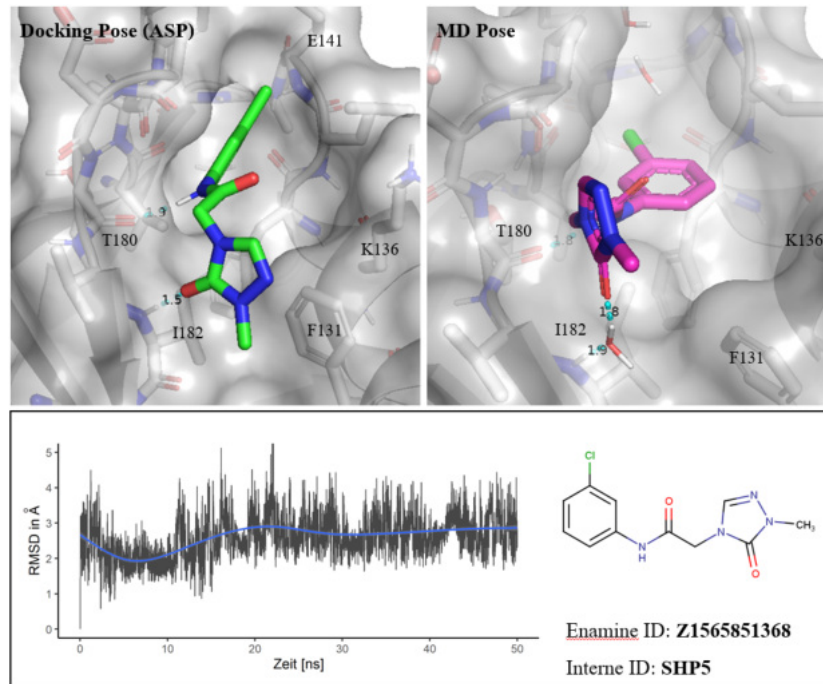


Abbildung B.4: Docking und MD-Profil der Verbindung SHP5

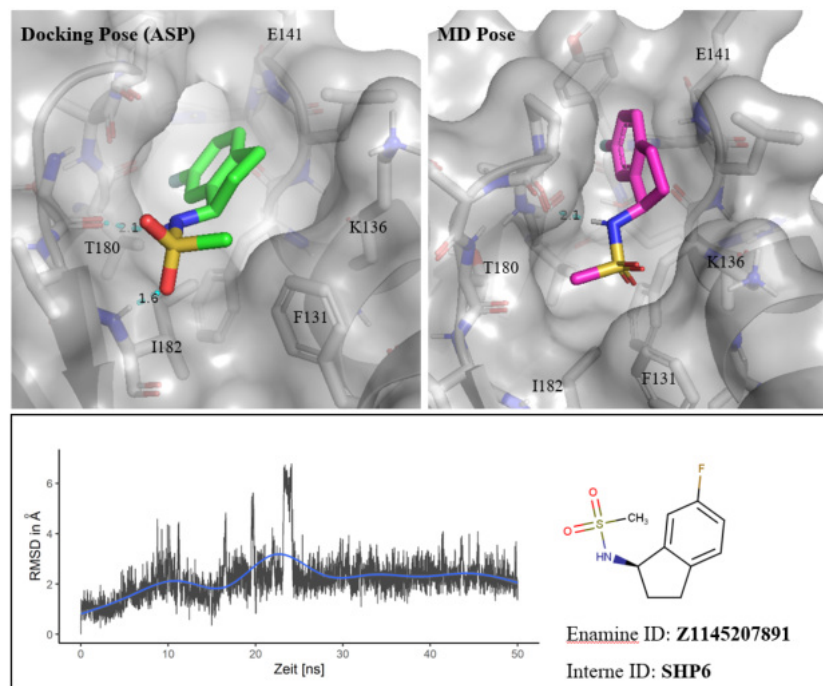


Abbildung B.5: Docking und MD-Profil der Verbindung SHP6

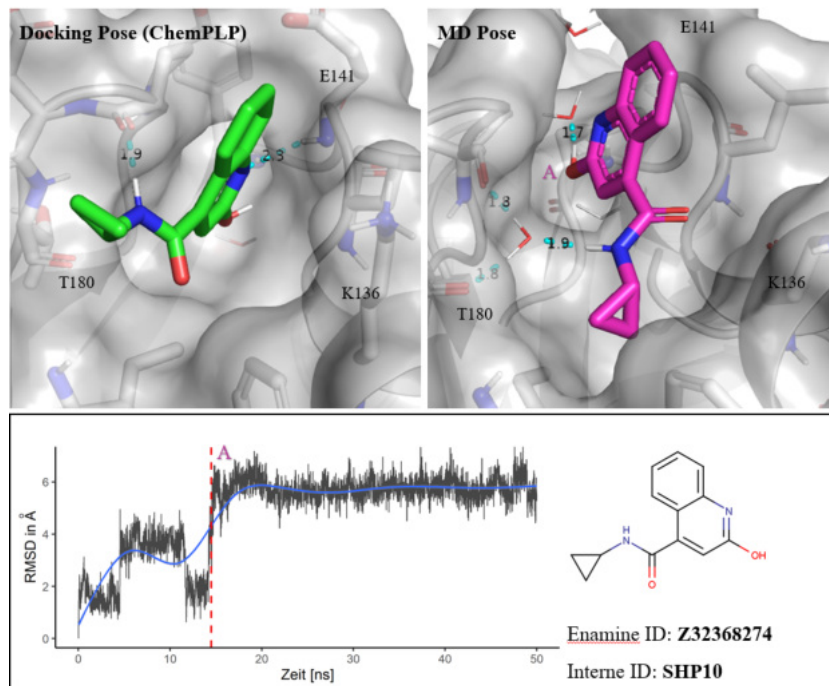


Abbildung B.6: Docking und MD-Profil der Verbindung SHP10

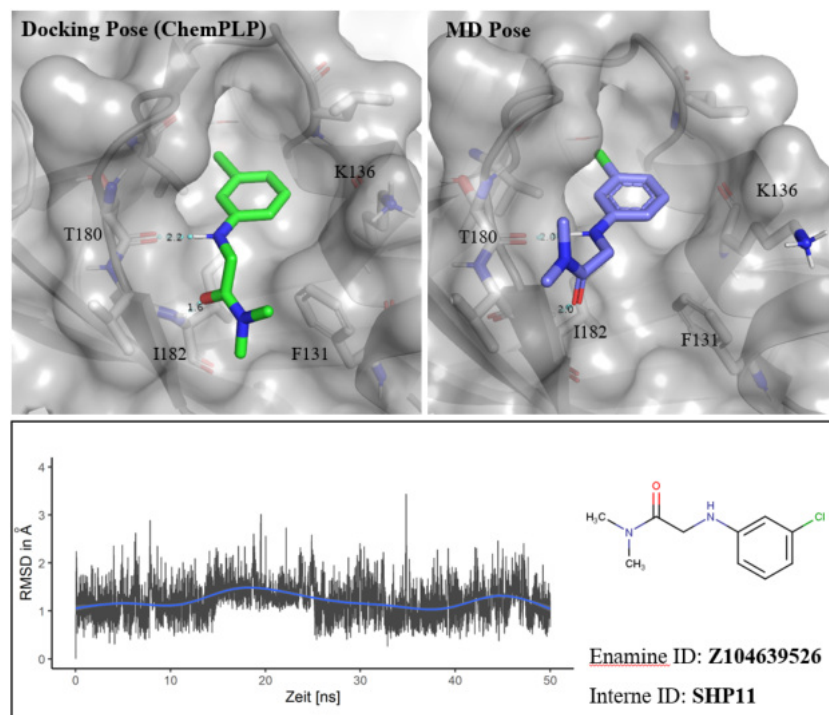


Abbildung B.7: Docking und MD-Profil der Verbindung SHP11

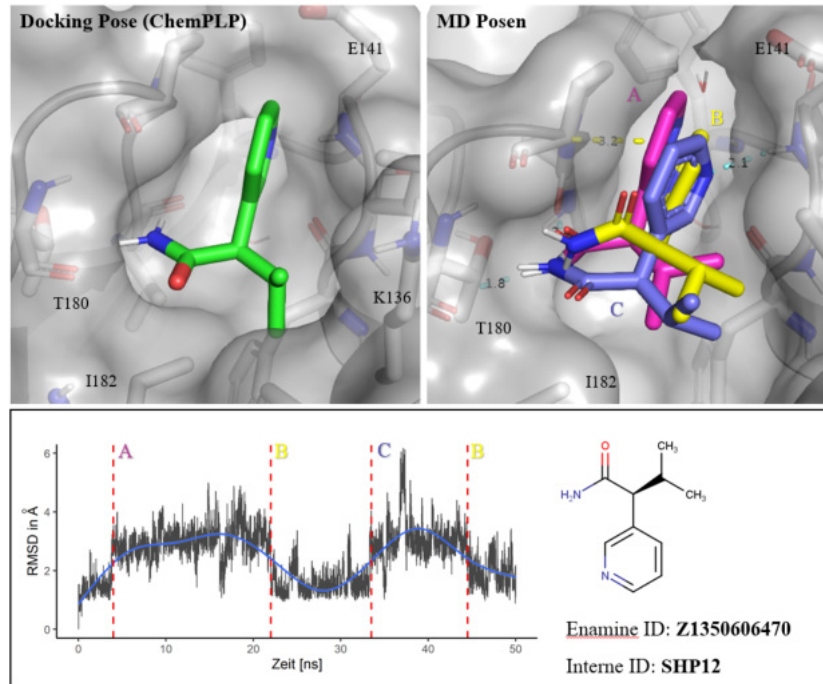


Abbildung B.8: Docking und MD-Profil der Verbindung SHP12

Anhang C

Retrospektive Betrachtung der Dockingergebnisse

Die Ergebnisse des Dockings wurden retrospektive noch einmal auf die Frage analysiert, ob ein consensusbasierter Ansatz andere Strukturen favorisiert hätte. Hierbei wurden die Score-Werte pro Schweratom für die jeweilige Scoringfunktion, sowie die vom Rescoring erhaltenen DSX-Score pro Schweratom zunächst normiert, und anschließend ein Consensus-Score durch $Score_{consensus} = Score_{norm}ASP/ChemPLP + Score_{norm}DSX$ gebildet. Hierbei zeigte sich vor allem für die Fragmente, die mit der ChemPLP-Scoringfunktion gescored wurden, neue Strukturen. Die mit diesem Consensus-Score höchst bewerteten Moleküle zeigten alle anstatt eines aromatischen Systems, hydrierte Strukturen mit einem hohen Anteil an sp^3 hybridisierten Kohlenstoffatomen. Alle füllten sterisch gesehen sehr gut die Bindetasche aus und wiesen meistens eine polare Interaktion mit P178 bzw. T180 auf. Fragmente die mit der ASP-Scoringfunktion gescored wurden, wiesen diese neuen Strukturen dagegen nicht auf. Der chemische Raum den diese Moleküle einnehmen ist leider nicht von der UiB Fragmentbibliothek abgedeckt, so dass sich anhand des durchgeführten biophysikalischen Fragmentscreenings diese Beobachtung nicht bestätigen lässt. Zukünftige erneute Analysen und MD-Simulationen sind hier im Fortgang dieses Projektes auf Grundlage dieser Daten von Nöten.

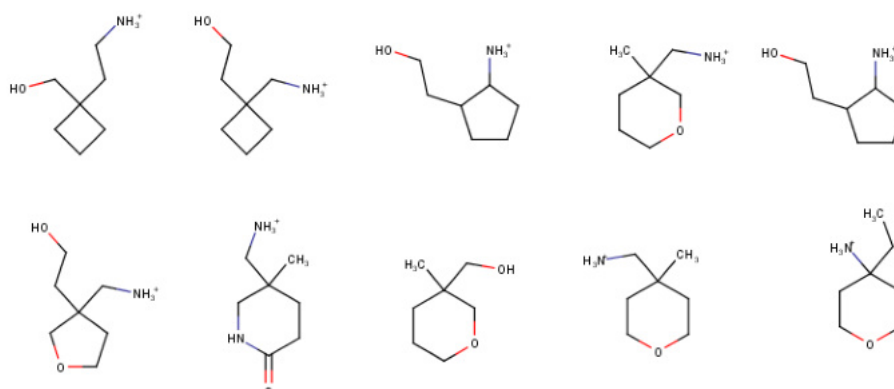


Abbildung C.1: Besten Treffer Consensus-Score

Von rechts nach links sind die Strukturen der 10 besten Fragmente nach Consensus-Score zu sehen.

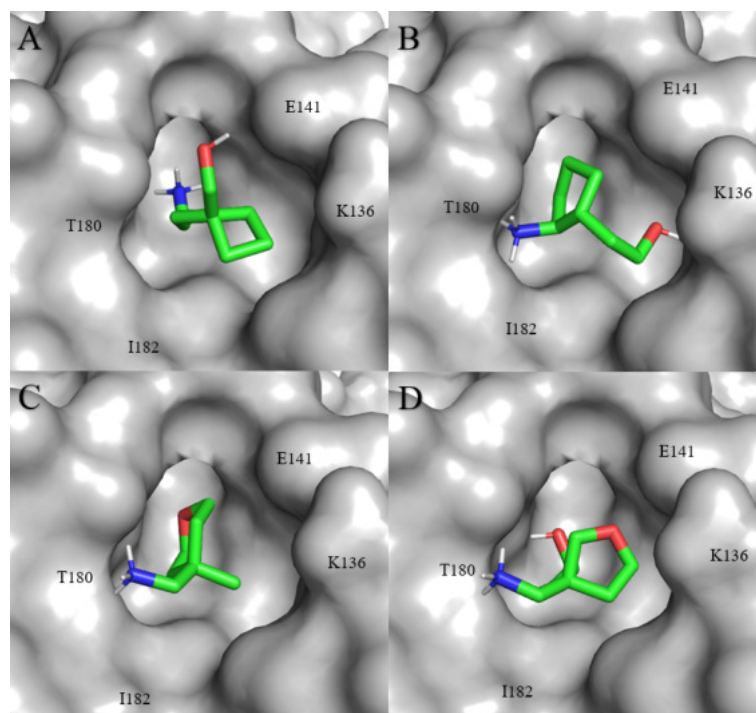


Abbildung C.2: Besten Treffer Consensus-Score
 Von rechts nach links sind die Strukturen der 10 besten Fragmente nach Consensus Score zu sehen.

Tabelle C.1: Übersicht Fragmente Consensus-Score

ID	DSX/HA	ChemPLP/HA	ChemPLP _{HA} norm.	DSX _{HA} norm.	Score _{con}
Z2456312994	-7.82	8.12	1.00	0.95	1.95
Z1443588953	-7.66	7.93	0.95	0.91	1.86
Z1742055021	-7.72	7.77	0.91	0.92	1.83
Z1079127074	-7.68	6.93	0.69	0.91	1.60
Z1037341778	-8.05	6.37	0.55	1.00	1.55
Z1398461930	-7.90	6.40	0.56	0.96	1.52
Z1431002654	-7.16	7.02	0.71	0.79	1.50
Z2418194148	-7.61	6.59	0.60	0.90	1.50
Z1270138957	-7.29	6.86	0.67	0.82	1.49
Z2472857066	-7.47	6.65	0.62	0.86	1.48

Anhang D

Sensorgramme der Xanthohumol (XN)-Messungen mit p97

D.1 Messungen mit der N-Domäne (SSA-Sensoren)

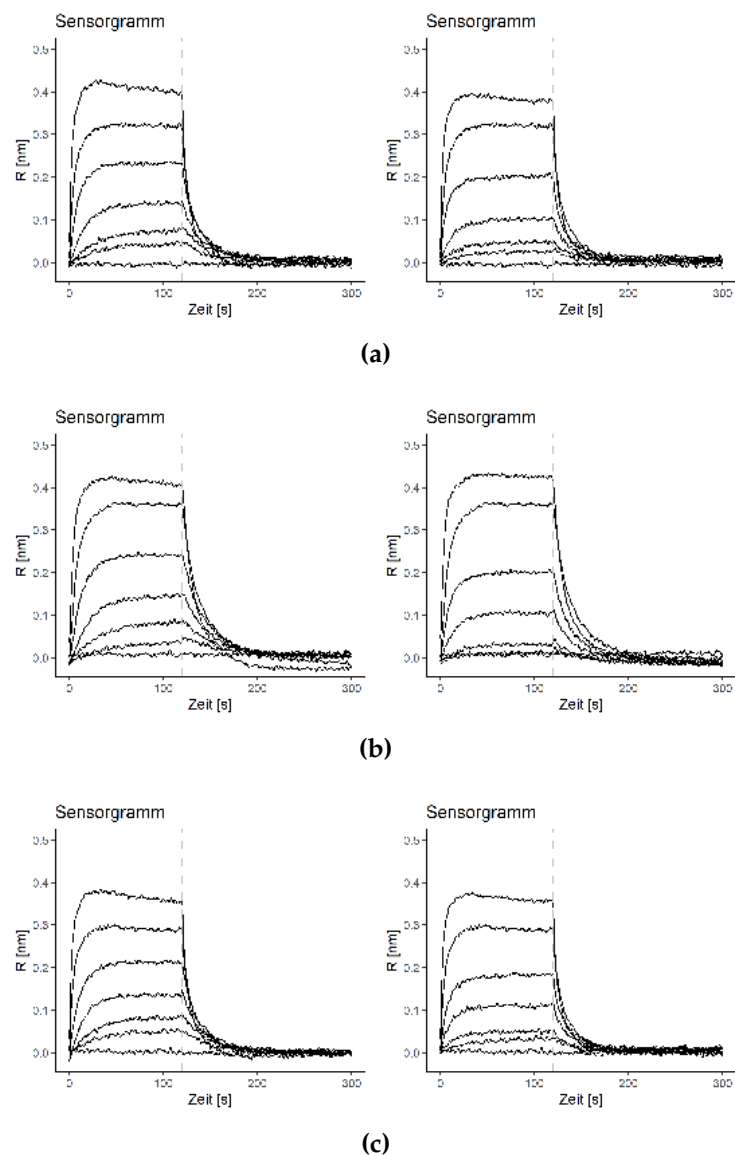


Abbildung D.1: Sensorgramme aller sechs Messungen, der 1:1-Verdünnungsreihen von 20 bis 0.625 μM XN mit p97-N. (a) Messungen 1. Sensor, (b) Messungen 2. Sensor, (c) Messungen 3. Sensor.

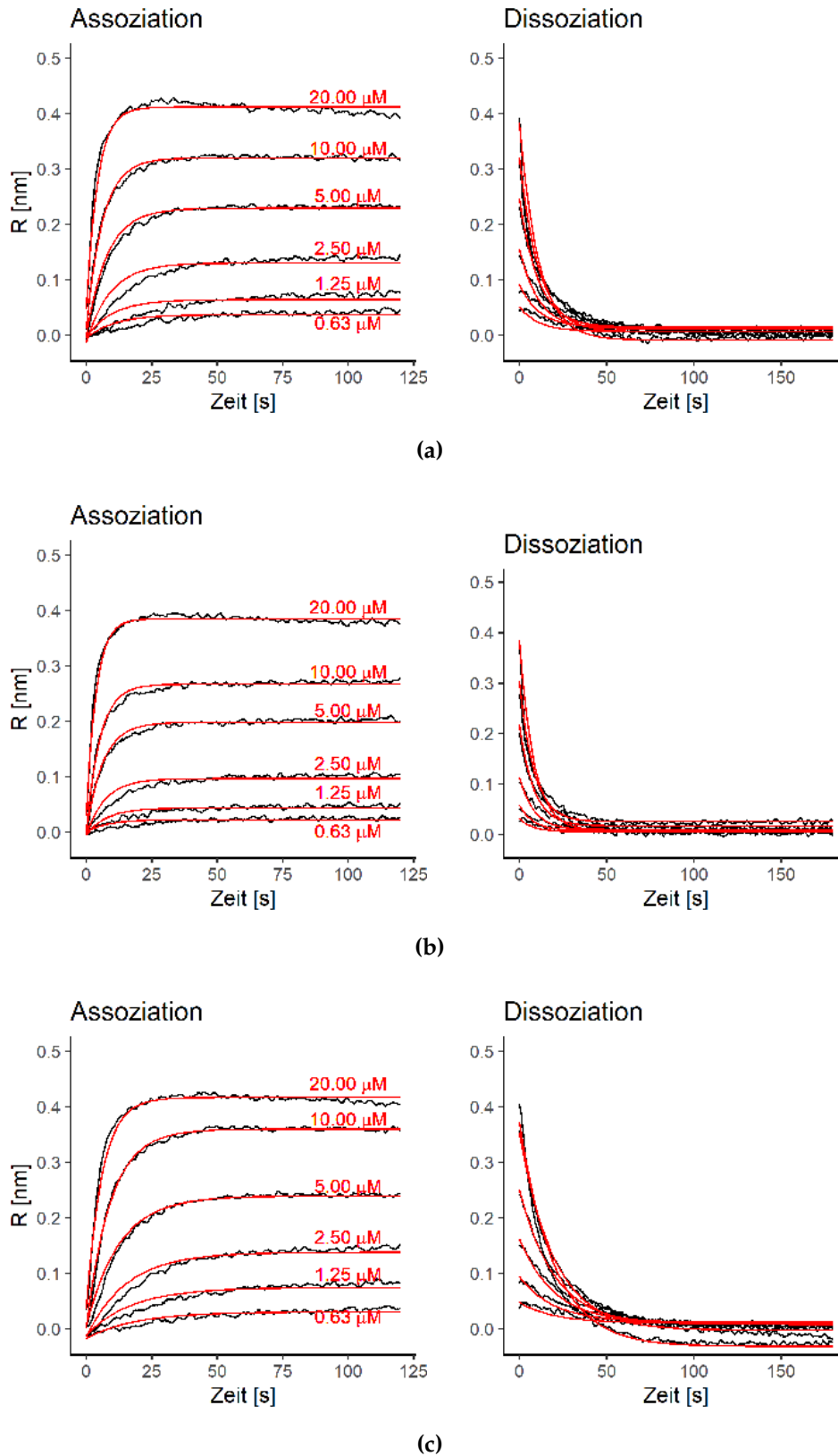


Abbildung D.2: Überlagerung der simulierten Kurven basierend auf einer globalen Regression eines 1:1-Bindungsmodells von XN-Messungen mit p97-N.

(a) 1. Messungen 1. Sensor, (b) 2. Messungen 1. Sensor, (c) 1. Messungen 2. Sensor.

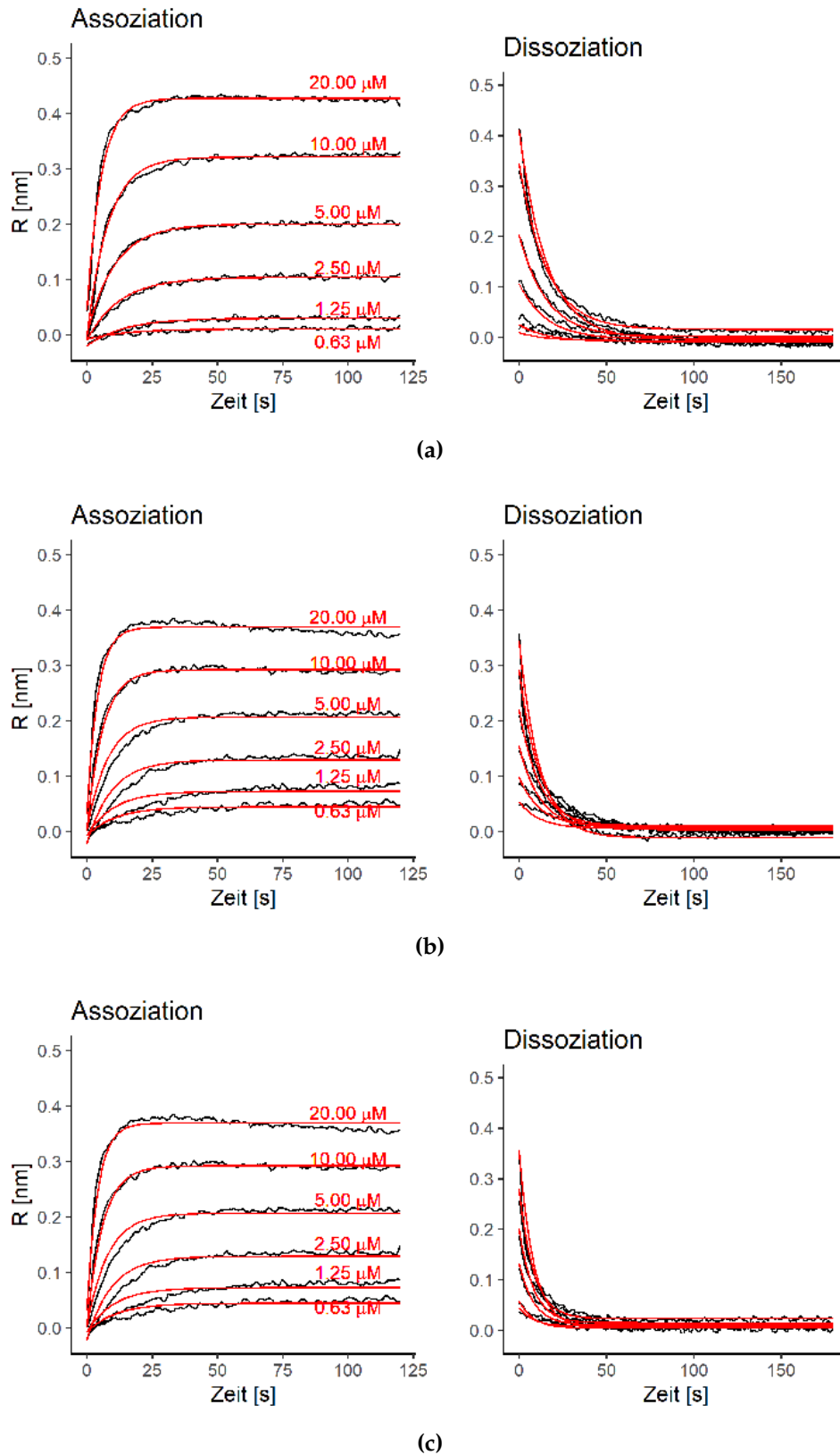


Abbildung D.3: Überlagerung der simulierten Kurven basierend auf einer globalen Regression eines 1:1-Bindungsmodells XN-Messungen mit p97-N.

(a) 2. Messungen 2. Sensor, (b) 1. Messungen 3. Sensor, (c) 2. Messungen 3. Sensor.

D.2 Messungen mit der ND1-Domäne (SSA-Sensoren)

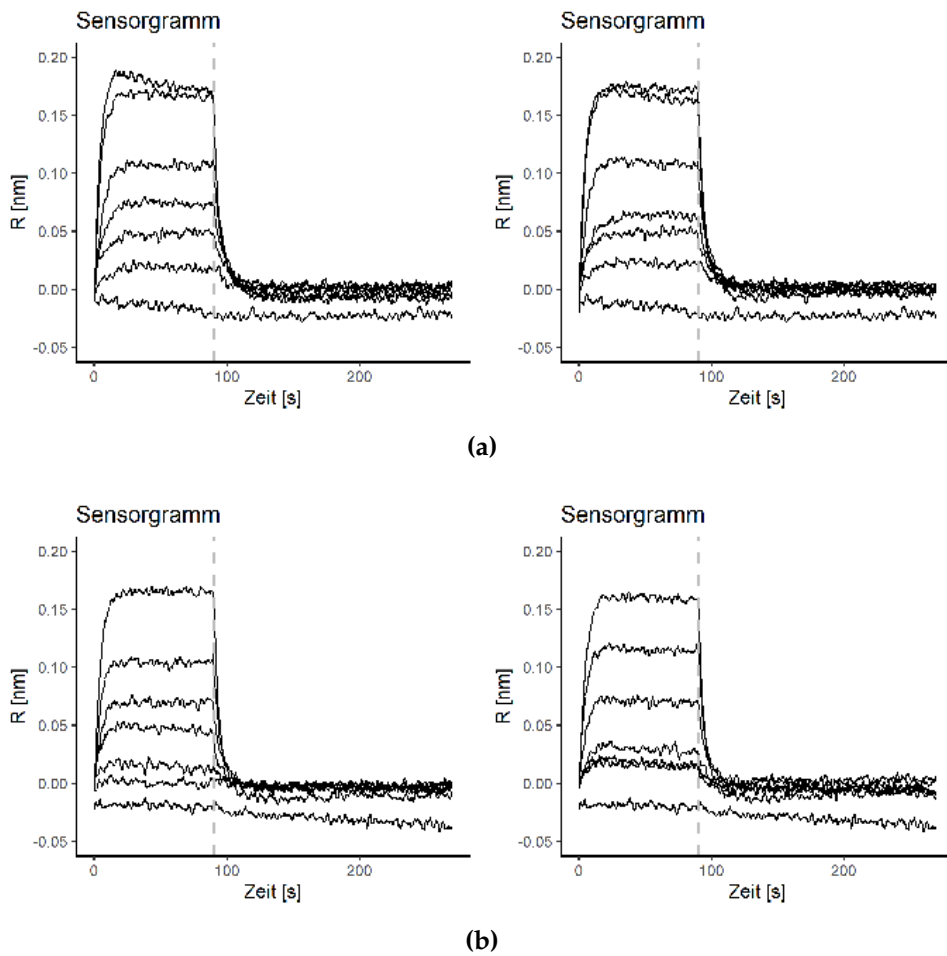
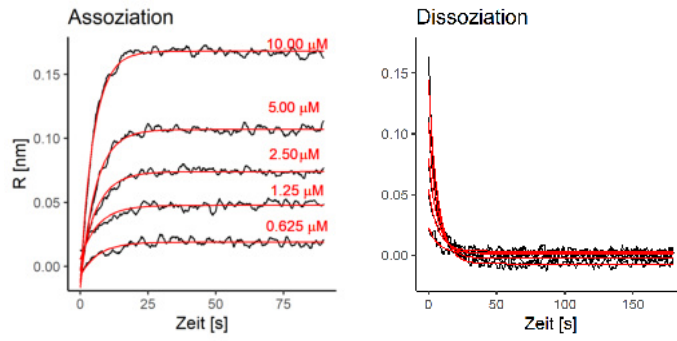
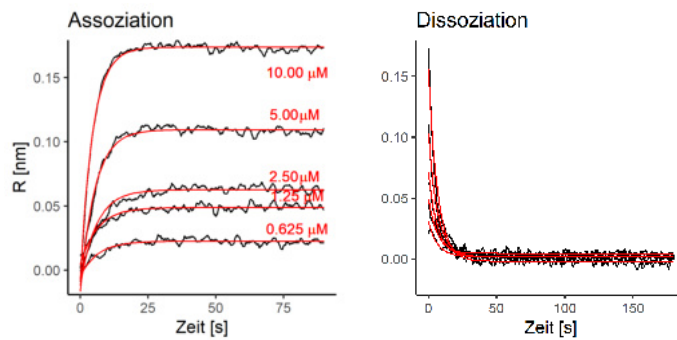


Abbildung D.4: Sensorgramme aller vier Messungen, der 1:1-Verdünnungsreihen von 20 bis 0.625 μM XN-Messungen mit p97-ND1.

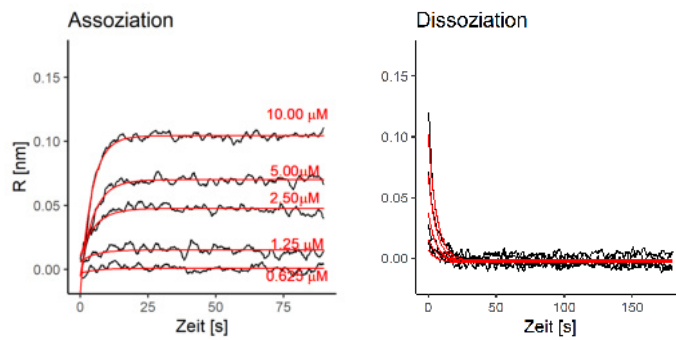
(a) Messungen 1. Sensor, (b) Messungen 2. Sensor.



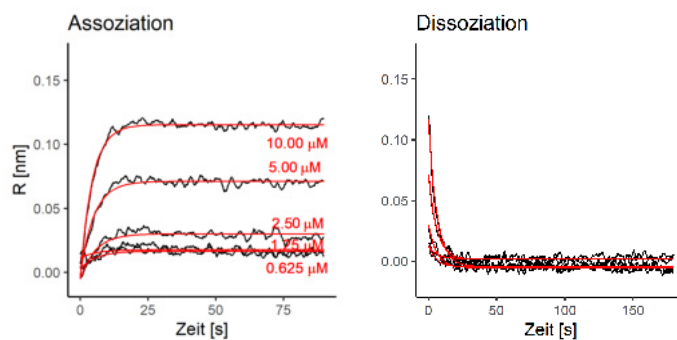
(a)



(b)



(c)



(d)

Abbildung D.5: Überlagerung der simulierten Kurven basierend auf einer globalen Regression eines 1:1-Bindungsmodells von XN mit p97-ND1 (SSA-Sensoren).

(a) 1. Messungen 1. Sensor, (b) 2. Messungen 1. Sensor, (c) 1. Messungen 2. Sensor., (d) 2. Messungen 2. Sensor.

D.3 Messungen mit der ND1-Domäne (SAX-Sensoren)

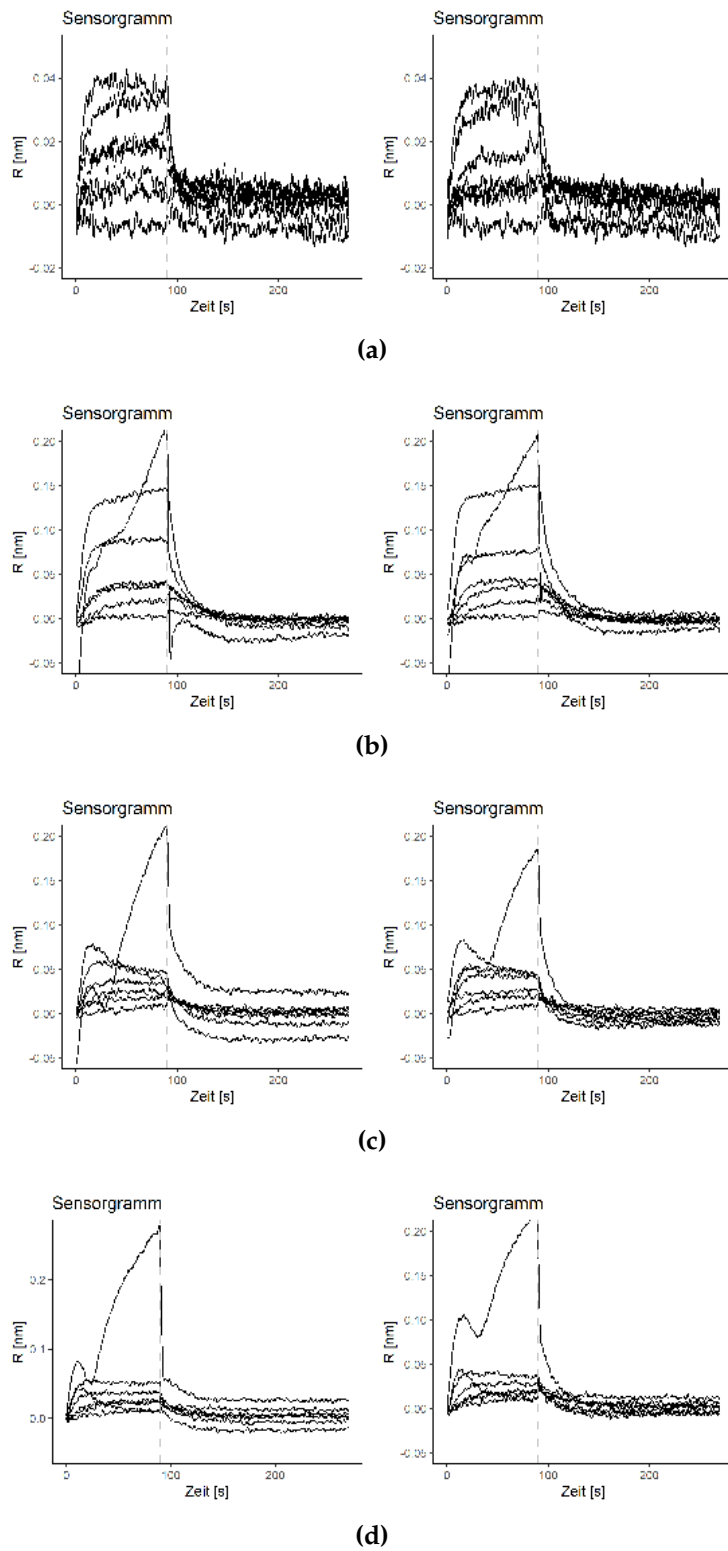


Abbildung D.6: Sensorgramme aller vier Messungen, der 1:1-Verdünnungsreihen von 20 bis $0.625 \mu\text{M}$ XN mit p97-ND1. (a) Messungen 1. Sensor, (b) Messungen 2. Sensor, (c) Messungen 3. Sensor, (d) Messungen 4. Sensor.

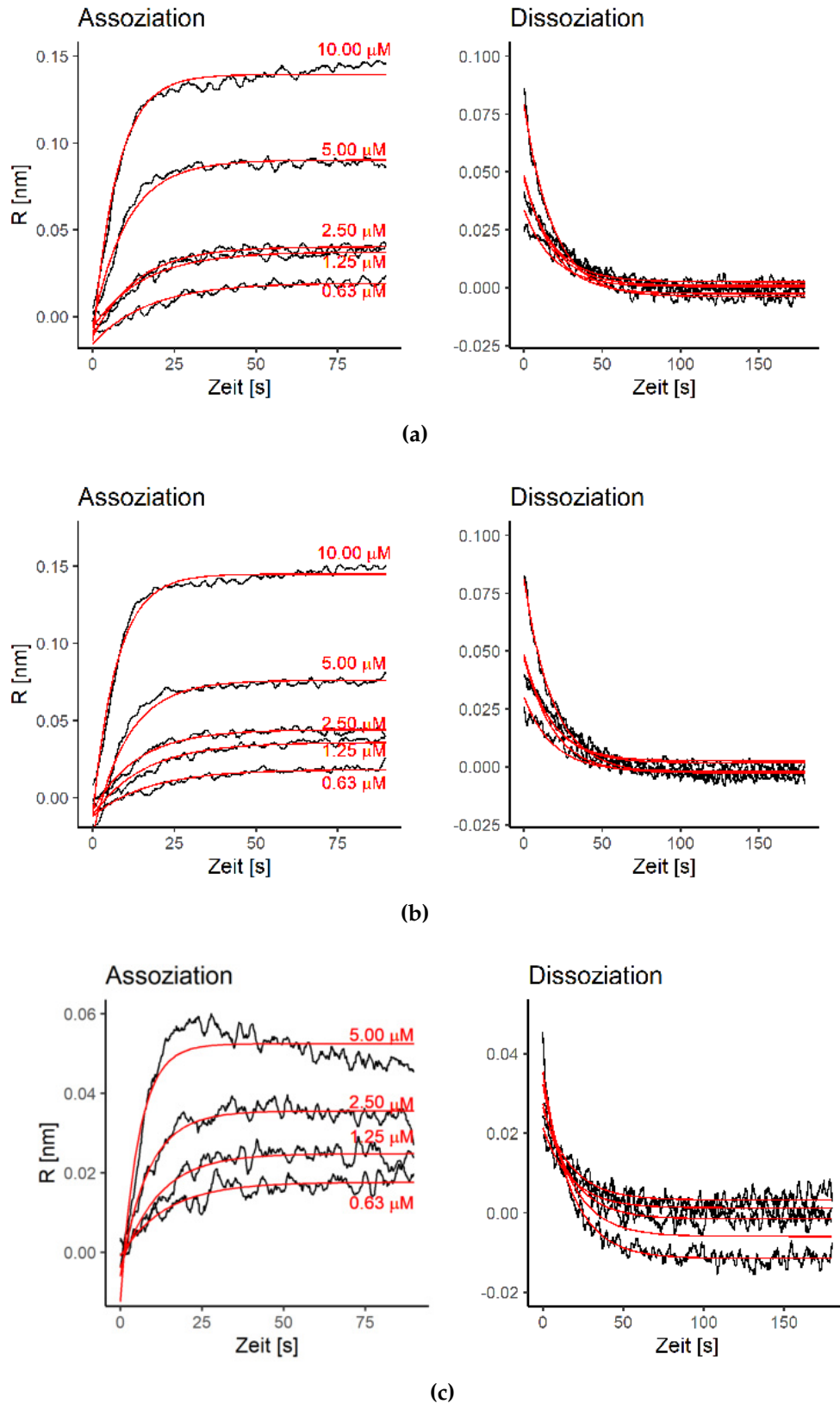


Abbildung D.7: Überlagerung der simulierten Kurven basierend auf einer globalen Regression eines 1:1-Bindungsmodells von XN-Messungen mit p97-ND1 (SAX-Sensoren).

(a) 1. Messungen 2. Sensor, (b) 2. Messungen 2. Sensor, (c) 1. Messungen 3. Sensor.

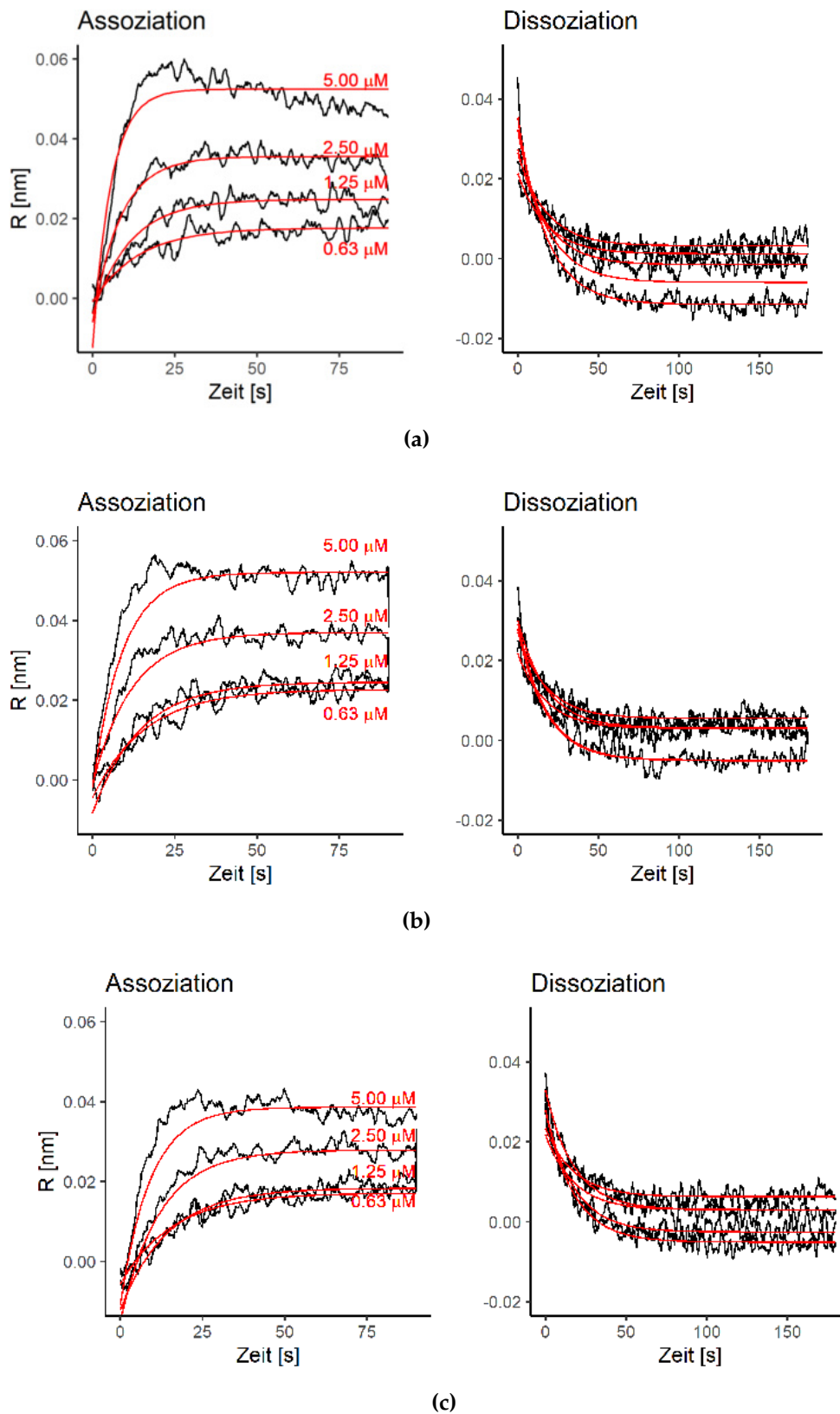


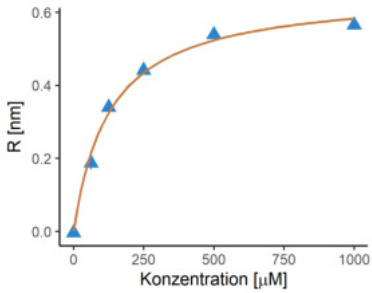
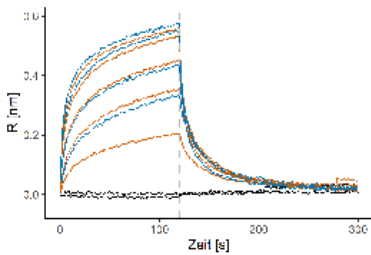
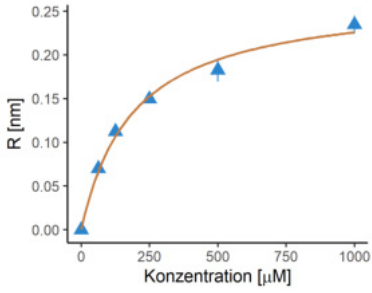
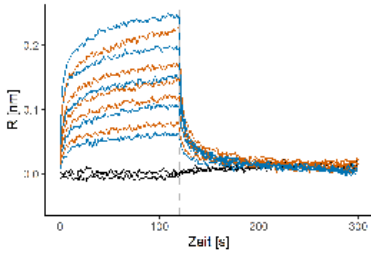
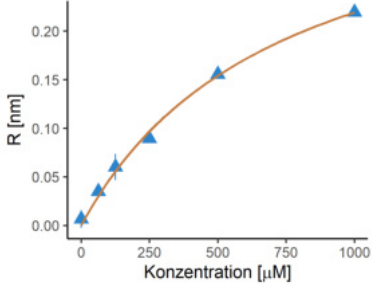
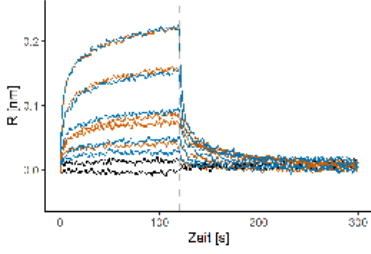
Abbildung D.8: Überlagerung der simulierten Kurven basierend auf einer globalen Regression eines 1:1-Bindungsmodells von XN-Messungen mit p97-ND1 (SAX-Sensoren).

(a) 2. Messungen 3. Sensor, (b) 1. Messungen 4. Sensor, (c) 2. Messungen 4. Sensor.

Anhang E

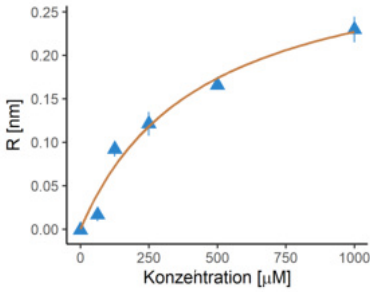
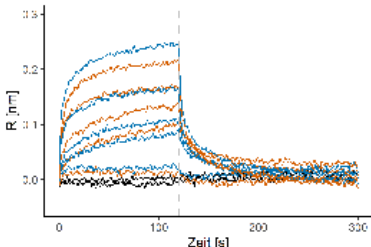
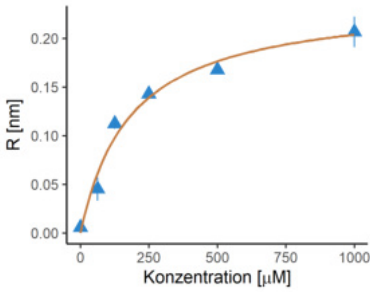
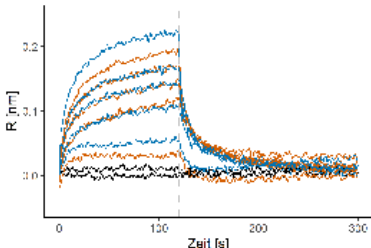
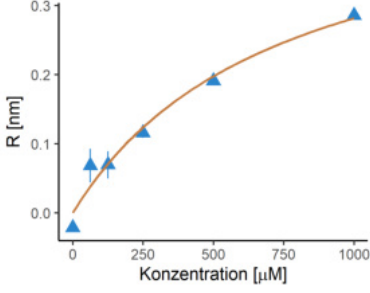
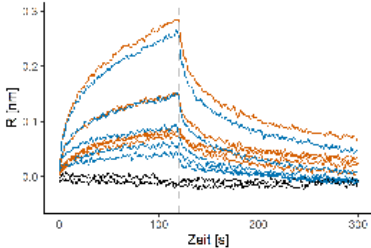
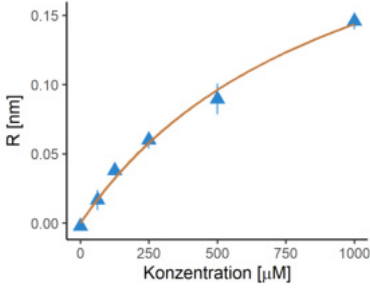
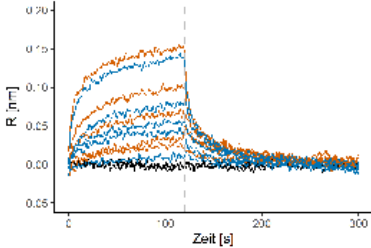
Sensorgramme und Langmuir-Modelle des Screenings mit p97-ND1

Tabelle E.1: Sensorgramme und Langmuir-Modelle aller identifizierten Treffer

Langmuir-Modell	Sensorgramme	Ergebnisse Modell
		<p>ID_{Screen}: G1-B6 ID_{intern}: VIK1 K_D: 115 ± 13 [$\mu\text{mol L}^{-1}$] R_{max}: 0.642 [nm] R^2: 0.997</p>
		<p>ID_{Screen}: A1-D2 ID_{intern}: VIK20 K_D: 192 ± 25 [$\mu\text{mol L}^{-1}$] R_{max}: 0.267 [nm] R^2: 0.998</p>
		<p>ID_{Screen}: G1-B5 ID_{intern}: VIK40 K_D: 707 ± 32 [$\mu\text{mol L}^{-1}$] R_{max}: 0.374 [nm] R^2: 0.999</p>

Fortsetzung auf nächster Seite

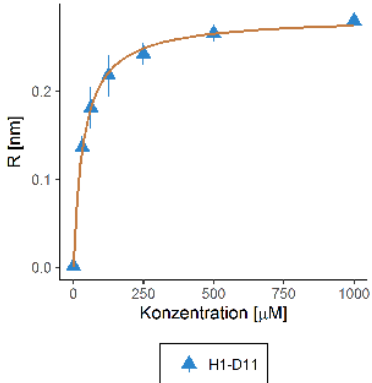
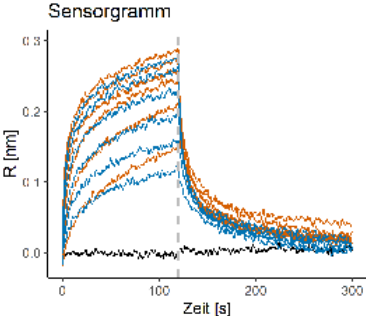
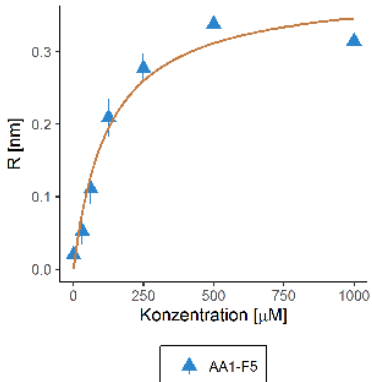
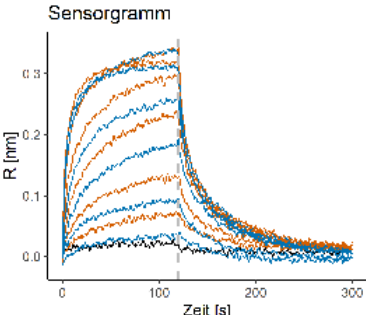
Tabelle E.1 Fortsetzung von letzter Seite

Langmuir-Modell	Sensorgramme	Ergebnisse Modell
		<p> ID_{Screen}: G1-F11 ID_{intern}: VIK50 K_D: 457 ± 227 [$\mu\text{mol L}^{-1}$] R_{max}: 0.319 [nm] R^2: 0.985 </p>
		<p> ID_{Screen}: A1-G10 ID_{intern}: VIK80 K_D: 112 ± 16 [$\mu\text{mol L}^{-1}$] R_{max}: 0.205 [nm] R^2: 0.988 </p>
		<p> ID_{Screen}: B1-B4 ID_{intern}: VIK90 K_D: 963 ± 409 [$\mu\text{mol L}^{-1}$] R_{max}: 0.561 [nm] R^2: 0.989 </p>
		<p> ID_{Screen}: E1-B7 ID_{intern}: VIK100 K_D: 785 ± 149 [$\mu\text{mol L}^{-1}$] R_{max}: 0.258 [nm] R^2: 0.994 </p>

Anhang F

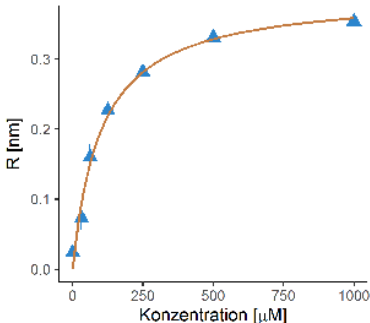
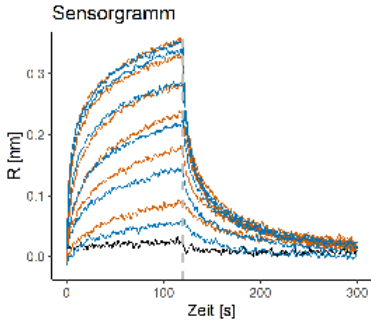
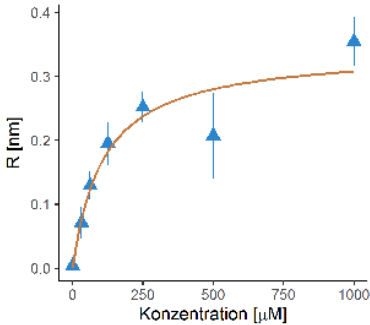
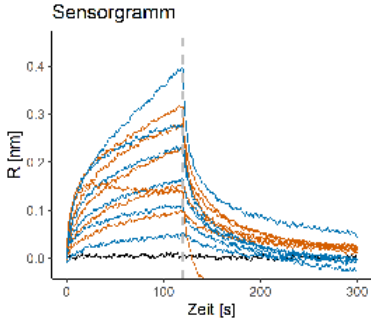
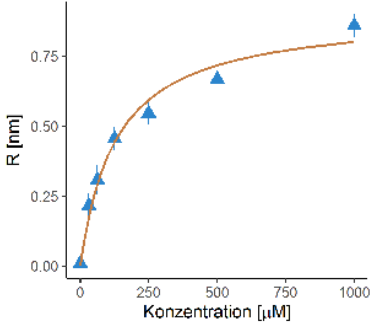
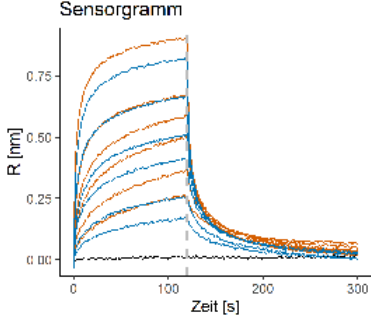
Sensorgramme und Langmuir-Modelle des Screenings mit p97-N

Tabelle F.1: Sensorgramme aller im N-Domänen Screening identifizierten Treffer

Langmuir-Modell	Sensorgramme	Ergebnisse Modell
 <p style="text-align: center;">▲ H1-D11</p>		<p>ID_{Screen}: H1-D11 ID_{intern}: TROLL1</p> <p>$K_D: 37 \pm 3 [\mu\text{mol L}^{-1}]$ $R_{max}: 0.288 [\text{nm}]$ $R^2: 0.991$</p>
 <p style="text-align: center;">▲ AA1-F5</p>		<p>ID_{Screen}: AA1-F5 ID_{intern}: TROLL2</p> <p>$K_D: 56 \pm 46 [\mu\text{mol L}^{-1}]$ $R_{max}: 0.374 [\text{nm}]$ $R^2: 0.592$</p>

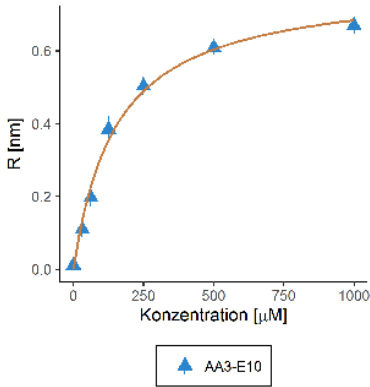
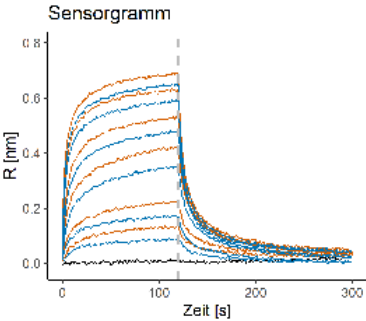
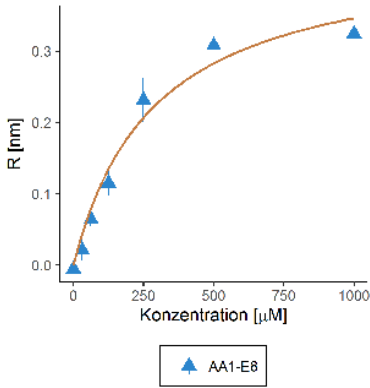
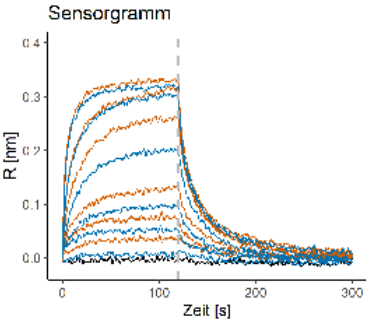
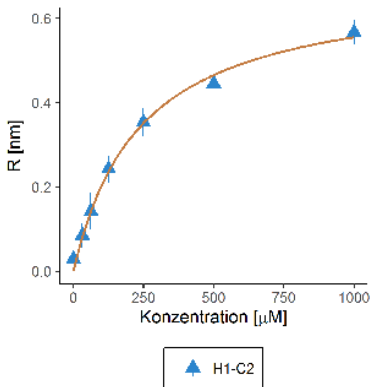
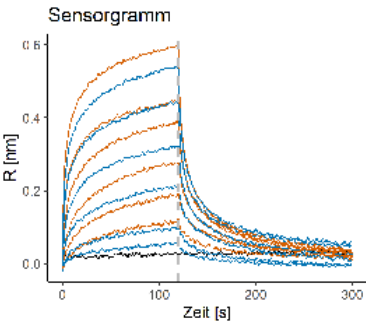
Fortsetzung auf nächster Seite

Tabelle F.1 Fortsetzung von letzter Seite

Langmuir-Modell	Sensorgramme	Ergebnisse Modell
 <p style="text-align: center;">▲ G1-B5</p>		<p>ID_{Screen}: G1-B5 ID_{intern}: VIK40</p> <p>$K_D: 93 \pm 4 [\mu\text{mol L}^{-1}]$ $R_{max}: 0.387 [\text{nm}]$ $R^2: 0.996$</p>
 <p style="text-align: center;">▲ AA1-F2</p>		<p>ID_{Screen}: AA1-F2 ID_{intern}: TROLL4</p> <p>$K_D: 124 \pm 29 [\mu\text{mol L}^{-1}]$ $R_{max}: 0.376 [\text{nm}]$ $R^2: 0.946$</p>
 <p style="text-align: center;">▲ A1-D2</p>		<p>ID_{Screen}: A1-D2 ID_{intern}: VIK20</p> <p>$K_D: 147 \pm 32 [\mu\text{mol L}^{-1}]$ $R_{max}: 0.956 [\text{nm}]$ $R^2: 0.961$</p>

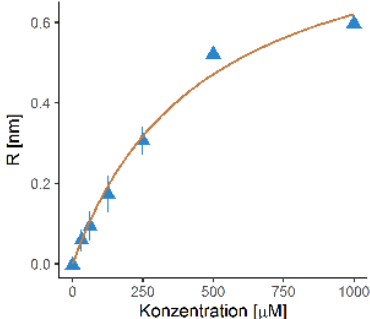
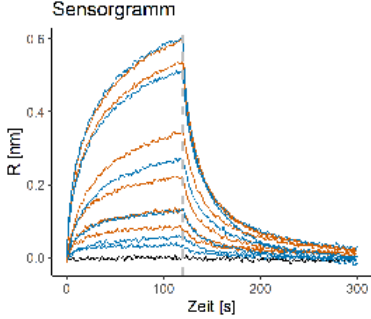
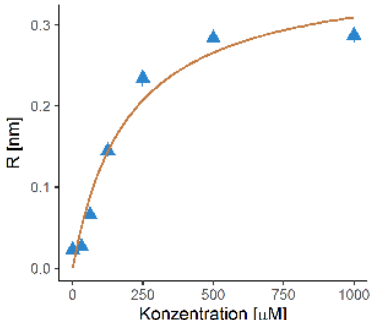
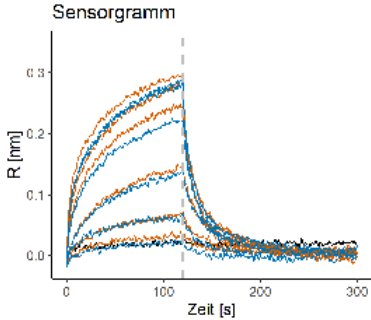
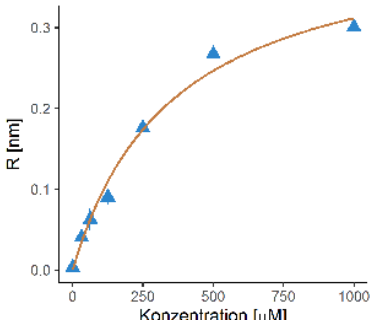
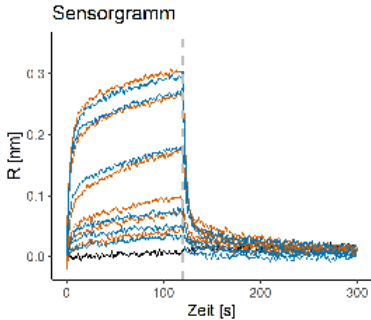
Fortsetzung auf nächster Seite

Tabelle F.1 Fortsetzung von letzter Seite

Langmuir-Modell	Sensorgramme	Ergebnisse Modell
		<p>ID_{Screen}: AA3-E10 ID_{intern}: TROLL6</p> <p>$K_D: 160 \pm 22 [\mu\text{mol L}^{-1}]$ $R_{max}: 0.793 [\text{nm}]$ $R^2: 0.99$</p>
		<p>ID_{Screen}: AA1-E8 ID_{intern}: TROLL7</p> <p>$K_D: 223 \pm 83 [\mu\text{mol L}^{-1}]$ $R_{max}: 0.416 [\text{nm}]$ $R^2: 0.949$</p>
		<p>ID_{Screen}: H1-C2 ID_{intern}: TROLL8</p> <p>$K_D: 245 \pm 43 [\mu\text{mol L}^{-1}]$ $R_{max}: 0.663 [\text{nm}]$ $R^2: 0.989$</p>

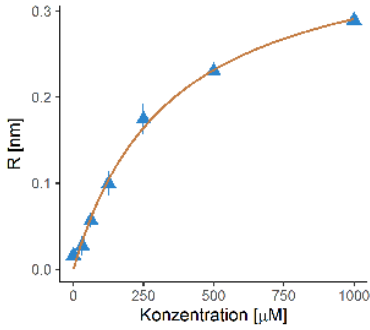
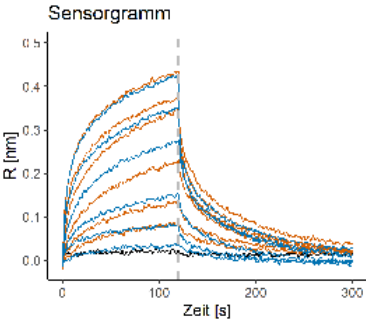
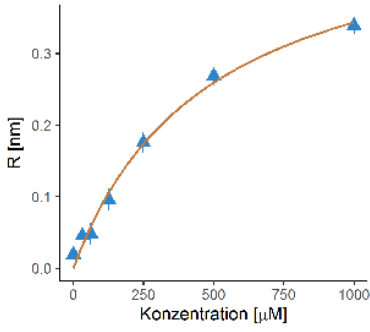
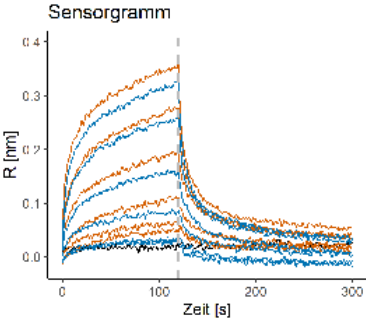
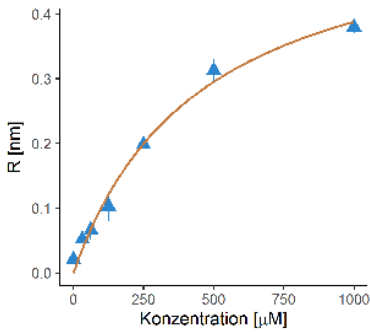
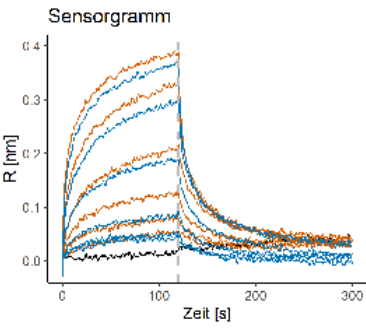
Fortsetzung auf nächster Seite

Tabelle F.1 Fortsetzung von letzter Seite

Langmuir-Modell	Sensorgramme	Ergebnisse Modell
 <p style="text-align: center;">▲ H1-D10</p>		<p>ID_{Screen}: H1-D10 ID_{intern}: TROLL9</p> <p>$K_D: 268 \pm 65 \text{ [}\mu\text{mol L}^{-1}\text{]}$ $R_{max}: 0.757 \text{ [nm]}$ $R^2: 0.978$</p>
 <p style="text-align: center;">▲ H1-E6</p>		<p>ID_{Screen}: H1-E6 ID_{intern}: TROLL10</p> <p>$K_D: 276 \pm 89 \text{ [}\mu\text{mol L}^{-1}\text{]}$ $R_{max}: 0.438 \text{ [nm]}$ $R^2: 0.983$</p>
 <p style="text-align: center;">▲ F1-G6</p>		<p>ID_{Screen}: F1-G6 ID_{intern}: TROLL11</p> <p>$K_D: 323 \pm 48 \text{ [}\mu\text{mol L}^{-1}\text{]}$ $R_{max}: 0.409 \text{ [nm]}$ $R^2: 0.978$</p>

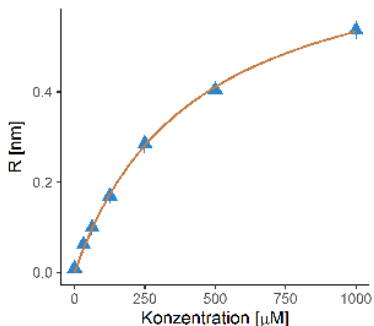
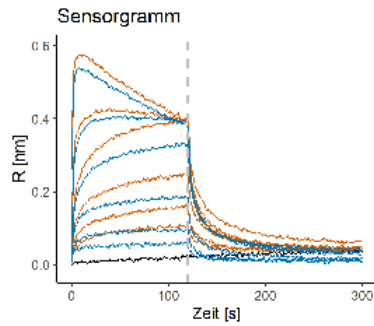
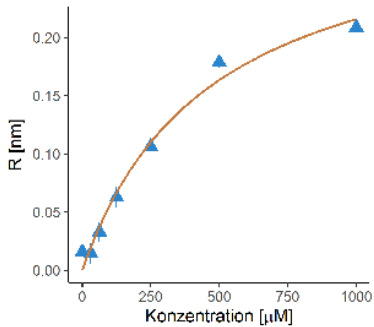
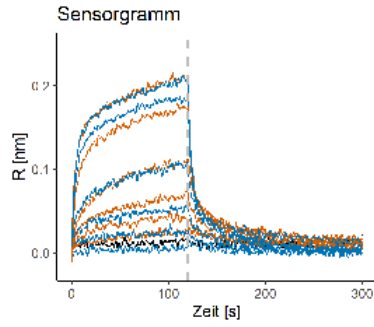
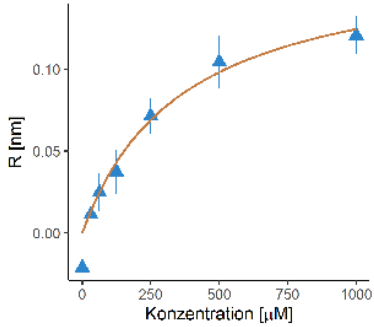
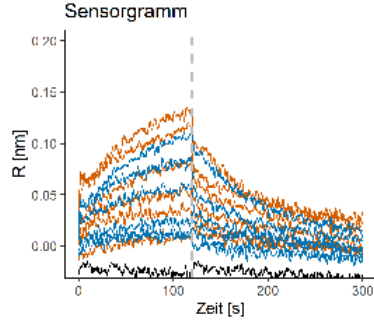
Fortsetzung auf nächster Seite

Tabelle F.1 Fortsetzung von letzter Seite

Langmuir-Modell	Sensorgramme	Ergebnisse Modell
 <p style="text-align: center;">▲ G1-G3</p>		<p>ID_{Screen}: G1-G3 ID_{intern}: TROLL12</p> <p>$K_D: 354 \pm 17 [\mu\text{mol L}^{-1}]$ $R_{max}: 0.392 [\text{nm}]$ $R^2: 0.999$</p>
 <p style="text-align: center;">▲ H1-E3</p>		<p>ID_{Screen}: H1-E3 ID_{intern}: TROLL13</p> <p>$K_D: 409 \pm 81 [\mu\text{mol L}^{-1}]$ $R_{max}: 0.478 [\text{nm}]$ $R^2: 0.989$</p>
 <p style="text-align: center;">▲ G1-G5</p>		<p>ID_{Screen}: G1-G5 ID_{intern}: TROLL14</p> <p>$K_D: 409 \pm 61 [\mu\text{mol L}^{-1}]$ $R_{max}: 0.535 [\text{nm}]$ $R^2: 0.991$</p>

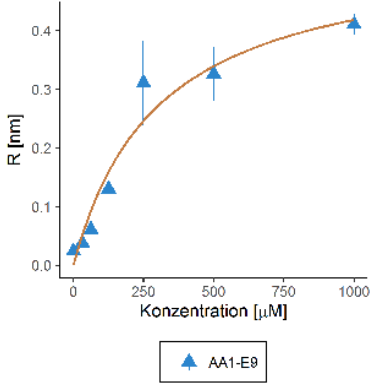
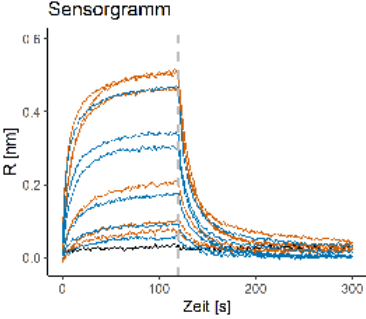
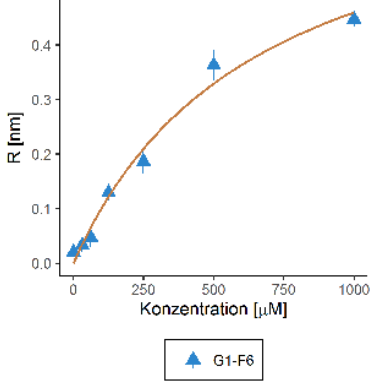
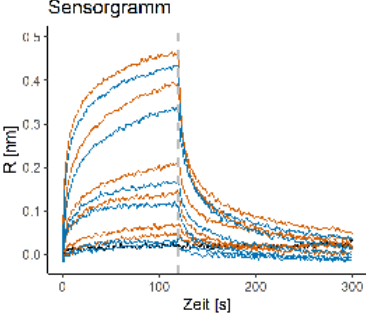
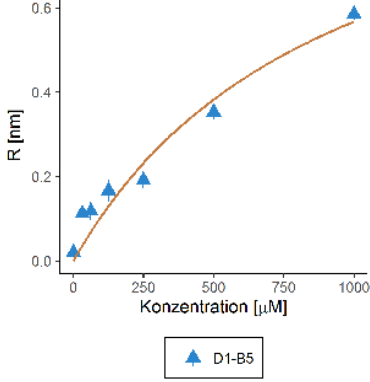
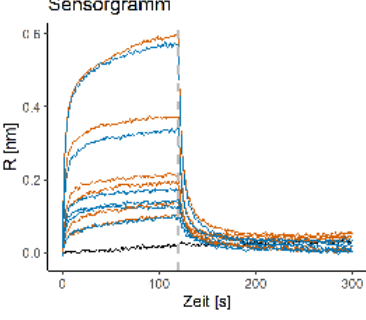
Fortsetzung auf nächster Seite

Tabelle F.1 Fortsetzung von letzter Seite

Langmuir-Modell	Sensorgramme	Ergebnisse Modell
 <p style="text-align: center;">▲ AA1-D2</p>		<p>ID_{Screen}: AA1-D2 ID_{intern}: TROLL15</p> <p>K_D: 425 ± 31 [$\mu\text{mol L}^{-1}$] R_{max}: 0.754 [nm] R^2: 0.998</p>
 <p style="text-align: center;">▲ AA1-C5</p>		<p>ID_{Screen}: AA1-C5 ID_{intern}: TROLL16</p> <p>K_D: 440 ± 53 [$\mu\text{mol L}^{-1}$] R_{max}: 0.300 [nm] R^2: 0.993</p>
 <p style="text-align: center;">▲ D1-D10</p>		<p>ID_{Screen}: D1-D10 ID_{intern}: TROLL17</p> <p>K_D: 444 ± 56 [$\mu\text{mol L}^{-1}$] R_{max}: 0.18 [nm] R^2: 0.992</p>

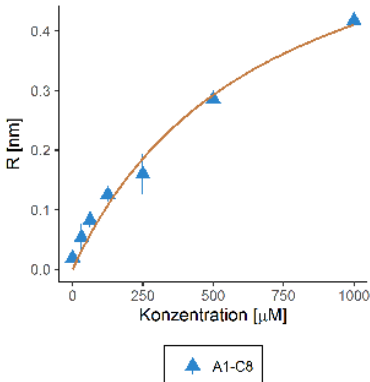
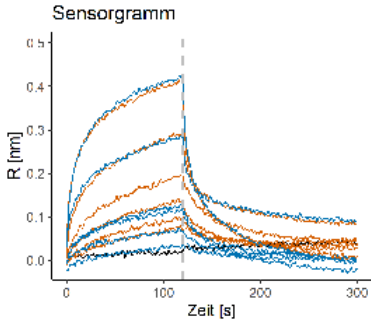
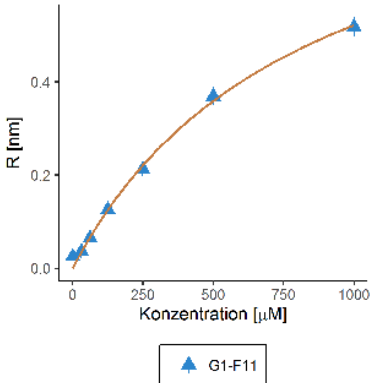
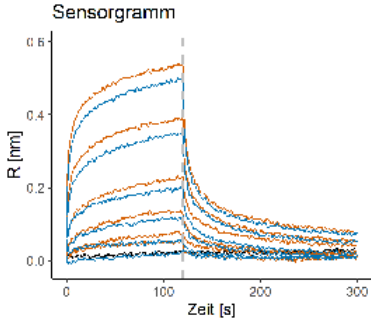
Fortsetzung auf nächster Seite

Tabelle F.1 Fortsetzung von letzter Seite

Langmuir-Modell	Sensorgramme	Ergebnisse Modell
 <p>▲ AA1-E9</p>		<p>ID_{Screen}: AA1-E9 ID_{intern}: TROLL18</p> <p>$K_D: 626 \pm 89 [\mu\text{mol L}^{-1}]$ $R_{max}: 0.679 [\text{nm}]$ $R^2: 0.975$</p>
 <p>▲ G1-F6</p>		<p>ID_{Screen}: G1-F6 ID_{intern}: TROLL19</p> <p>$K_D: 642 \pm 125 [\mu\text{mol L}^{-1}]$ $R_{max}: 0.738 [\text{nm}]$ $R^2: 0.992$</p>
 <p>▲ D1-B5</p>		<p>ID_{Screen}: D1-B5 ID_{intern}: TROLL20</p> <p>$K_D: 714 \pm 525 [\mu\text{mol L}^{-1}]$ $R_{max}: 0.978 [\text{nm}]$ $R^2: 0.908$</p>

Fortsetzung auf nächster Seite

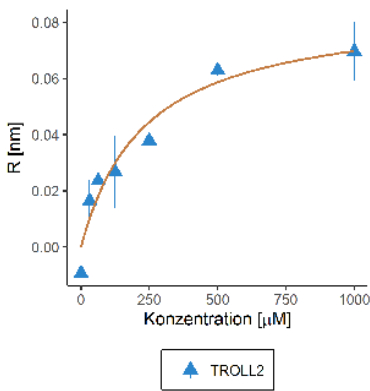
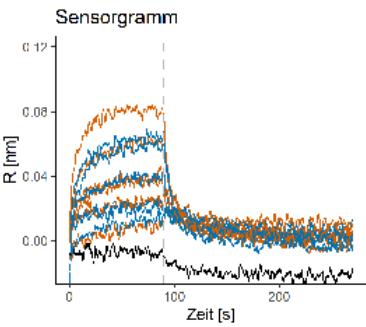
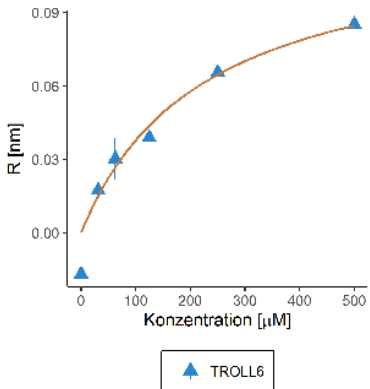
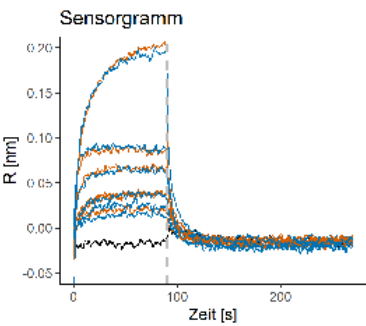
Tabelle F.1 Fortsetzung von letzter Seite

Langmuir-Modell	Sensorgramme	Ergebnisse Modell
		<p>ID_{Screen}: A1-C8 ID_{intern}: TROLL21</p> <p>K_D: 821 ± 100 [$\mu\text{mol L}^{-1}$] R_{max}: 0.757 [nm] R^2: 0.991</p>
		<p>ID_{Screen}: G1-F11 ID_{intern}: VIK50</p> <p>K_D: 853 ± 57 [$\mu\text{mol L}^{-1}$] R_{max}: 0.996 [nm] R^2: 0.999</p>

Anhang G

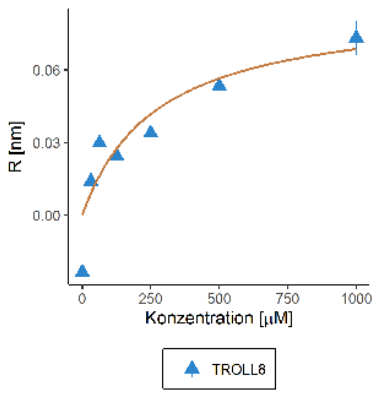
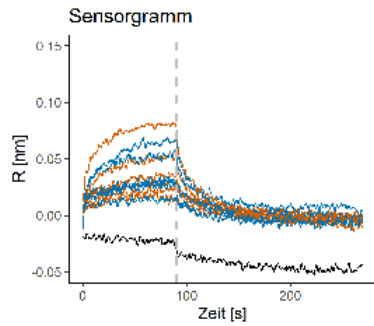
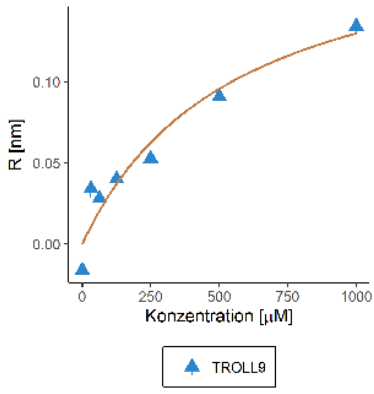
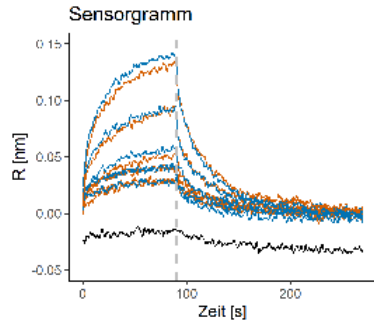
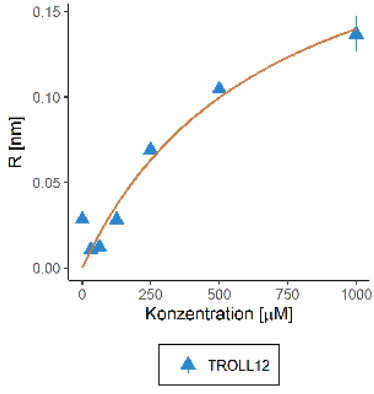
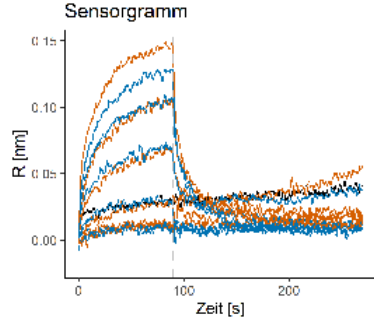
Sensorgramme und Langmuir-Modelle der Vermessung ausgewählter Treffer des p97-N-Screenings mit p97-ND1

Tabelle G.1: Sensorgramme und Langmuir-Modelle der Vermessung ausgewählter Treffer des p97-N Screenings mit p97-ND1

Langmuir-Modell	Sensorgramme	Ergebnisse Modell
 <p style="text-align: center;">▲ TROLL2</p>		<p>ID_{Screen}: AA1-F5 ID_{intern}: TROLL2</p> <p>K_D: 187 ± 90 [$\mu\text{mol L}^{-1}$] R_{max}: 0.078 [nm] R^2: 0.829</p>
 <p style="text-align: center;">▲ TROLL6</p>		<p>ID_{Screen}: AA3-E10 ID_{intern}: TROLL6</p> <p>K_D: 255 ± 36 [$\mu\text{mol L}^{-1}$] R_{max}: 0.129 [nm] R^2: 0.99</p>

Fortsetzung auf nächster Seite

Tabelle G.1 Fortsetzung von letzter Seite

Langmuir-Modell	Sensorgramme	Ergebnisse Modell
 <p>▲ TROLL8</p>		<p>ID_{Screen}: H1-C2 ID_{intern}: TROLL8</p> <p>K_D: 113 ± 13 [$\mu\text{mol L}^{-1}$] R_{max}: 0.065 [nm] R^2: 0.986</p>
 <p>▲ TROLL9</p>		<p>ID_{Screen}: H1-D10 ID_{intern}: TROLL9</p> <p>K_D: 362 ± 44 [$\mu\text{mol L}^{-1}$] R_{max}: 0.157 [nm] R^2: 0.98</p>
 <p>▲ TROLL12</p>		<p>ID_{Screen}: G1-G3 ID_{intern}: TROLL12</p> <p>K_D: 726 ± 72 [$\mu\text{mol L}^{-1}$] R_{max}: 0.257 [nm] R^2: 0.996</p>

Fortsetzung auf nächster Seite

Tabelle G.1 Fortsetzung von letzter Seite

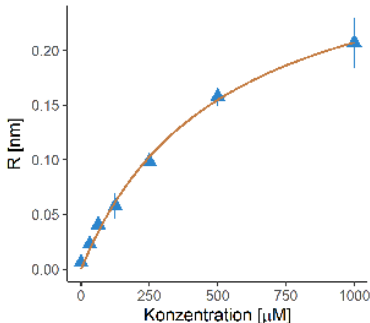
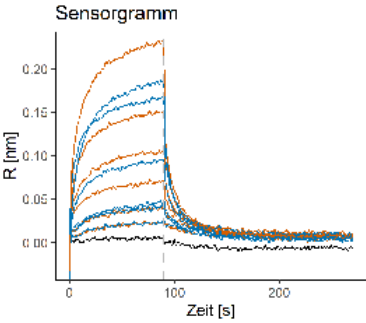
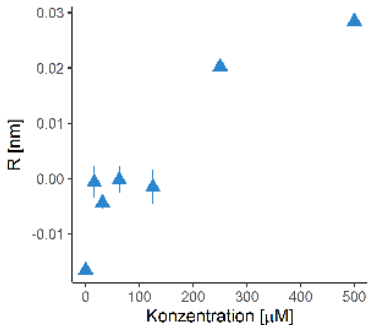
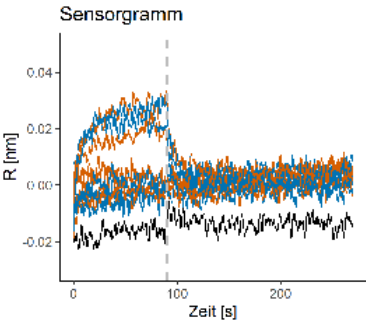
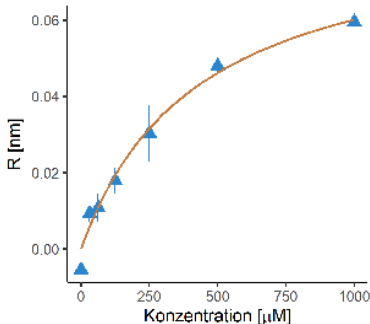
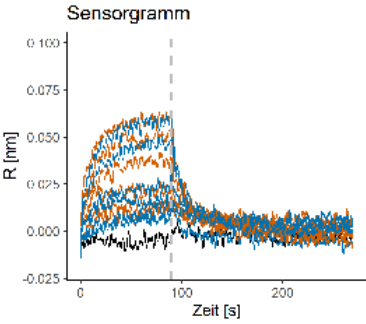
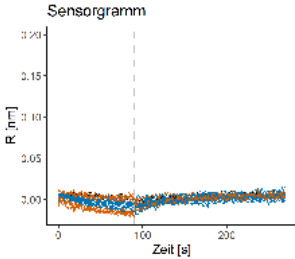
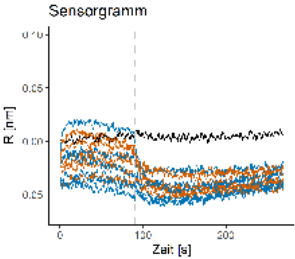
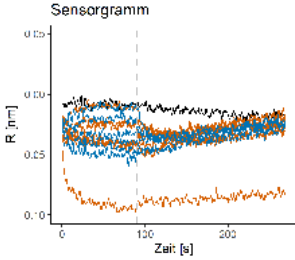
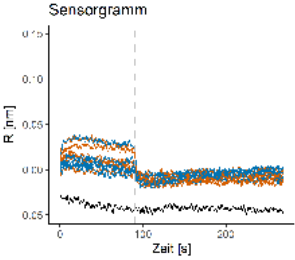
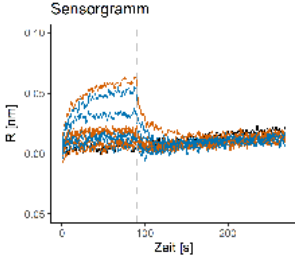
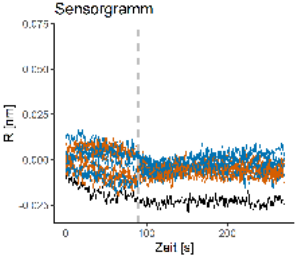
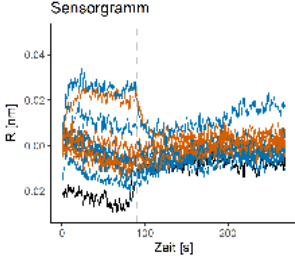
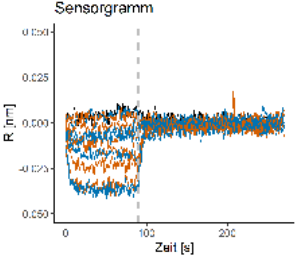
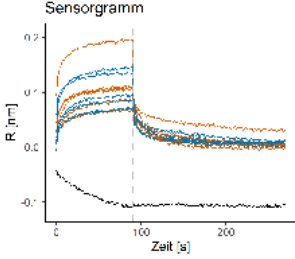
Langmuir-Modell	Sensorgramme	Ergebnisse Modell
 <p data-bbox="443 728 560 763">▲ TROLL14</p>		<p data-bbox="1098 383 1284 412">ID_{Screen}: G1-G5</p> <p data-bbox="1098 421 1321 450">ID_{intern}: TROLL14</p> <p data-bbox="1098 499 1396 528">$K_D: 290 \pm 33 [\mu\text{mol L}^{-1}]$</p> <p data-bbox="1098 537 1300 566">$R_{max}: 0.231 [\text{nm}]$</p> <p data-bbox="1098 575 1209 604">$R^2: 0.989$</p>
 <p data-bbox="443 1162 560 1198">▲ TROLL16</p>		<p data-bbox="1098 817 1305 846">ID_{Screen}: AA1-C5</p> <p data-bbox="1098 855 1321 884">ID_{intern}: TROLL16</p> <p data-bbox="1098 934 1412 963">$K_D: 386 \pm 161 [\mu\text{mol L}^{-1}]$</p> <p data-bbox="1098 972 1300 1001">$R_{max}: 0.051 [\text{nm}]$</p> <p data-bbox="1098 1010 1209 1039">$R^2: 0.903$</p>
 <p data-bbox="443 1597 560 1632">▲ TROLL19</p>		<p data-bbox="1098 1252 1284 1281">ID_{Screen}: G1-F6</p> <p data-bbox="1098 1290 1321 1319">ID_{intern}: TROLL19</p> <p data-bbox="1098 1368 1396 1397">$K_D: 351 \pm 57 [\mu\text{mol L}^{-1}]$</p> <p data-bbox="1098 1406 1284 1435">$R_{max}: 0.08 [\text{nm}]$</p> <p data-bbox="1098 1444 1209 1473">$R^2: 0.994$</p>

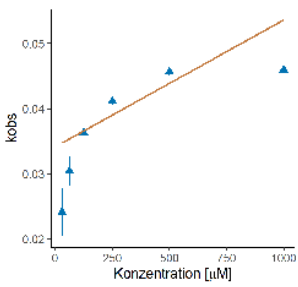
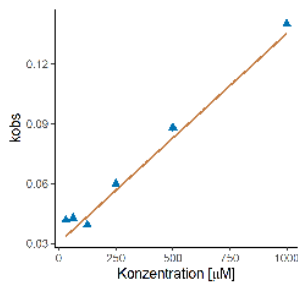
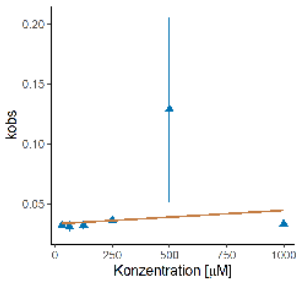
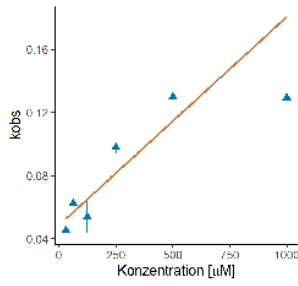
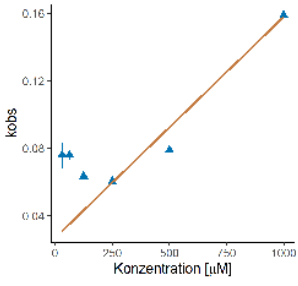
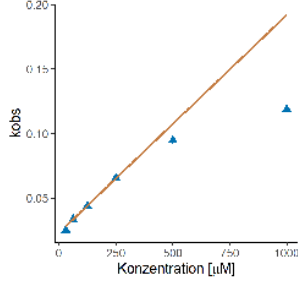
Tabelle G.2: Sensorgramme aller Fragmente diese nicht mit dem Langmuir-Modell modeliert werden konnten

Sensorgramme	ID	Sensorgramme	ID
	ID _{Screen} : H1-D11 ID _{intern} : TROLL1		ID _{Screen} : AA1-F2 ID _{intern} : TROLL4
	ID _{Screen} : AA1-E8 ID _{intern} : TROLL7		ID _{Screen} : H1-E6 ID _{intern} : TROLL10
	ID _{Screen} : F1-G6 ID _{intern} : TROLL11		ID _{Screen} : H1-E3 ID _{intern} : TROLL13
	ID _{Screen} : AA1-D2 ID _{intern} : TROLL15		ID _{Screen} : D1-D10 ID _{intern} : TROLL17
	ID _{Screen} : AA1-E9 ID _{intern} : TROLL18		

Anhang H

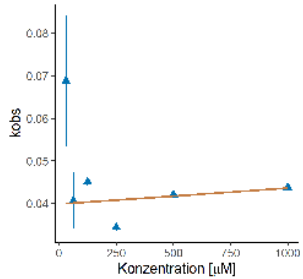
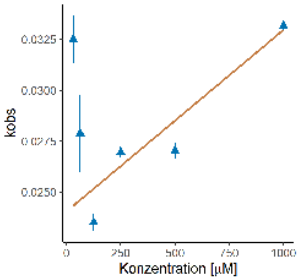
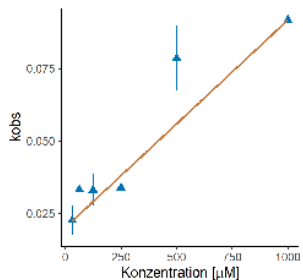
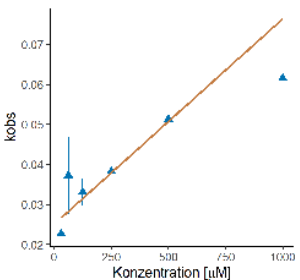
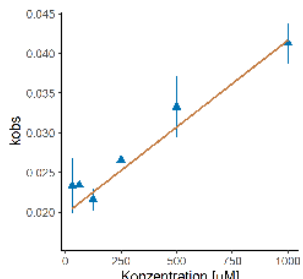
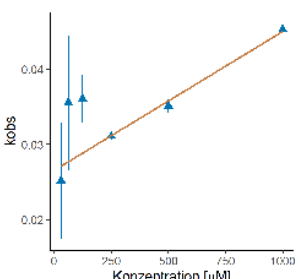
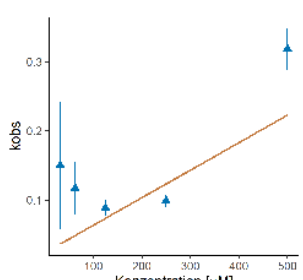
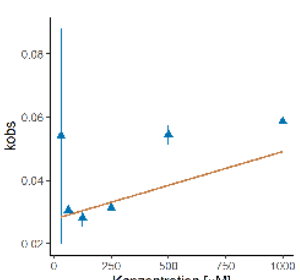
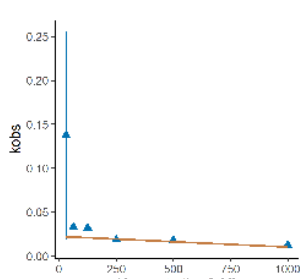
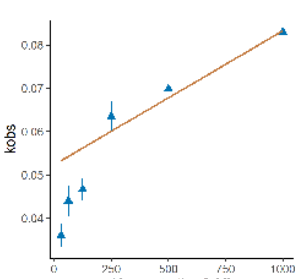
Kinetische Auswertung der Fragmente anhand ihrer empirischen Geschwindigkeitskonstanten

Tabelle H.1: Lineare Regressionen der aus der kinetischen Auswertung erhaltenen k_{obs} in Abhängigkeit von der Konzentration

Regression	ID	Regression	ID
	ID _{intern} : TROLL1		ID _{intern} : TROLL2
	ID _{intern} : TROLL4		ID _{intern} : TROLL6
	ID _{intern} : TROLL7		ID _{intern} : TROLL8

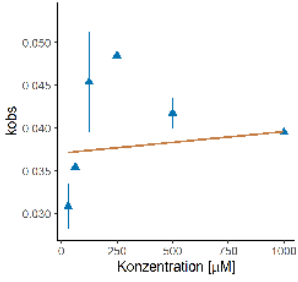
Fortsetzung auf nächster Seite

Tabelle H.1 Fortsetzung von letzter Seite

Regression	ID	Regression	ID
	ID _{intern} : TROLL9		ID _{intern} : TROLL10
	ID _{intern} : TROLL11		ID _{intern} : TROLL12
	ID _{intern} : TROLL13		ID _{intern} : TROLL14
	ID _{intern} : TROLL15		ID _{intern} : TROLL16
	ID _{intern} : TROLL17		ID _{intern} : TROLL18

Fortsetzung auf nächster Seite

Tabelle H.1 Fortsetzung von letzter Seite

Regression	ID	Regression	ID
	ID _{intern} : TROLL19	-	-

Anhang I

Eingesetzte Filter und Einstellungen des *Conformational Imports* in MOE

I.1 Eingesetzte Filter in der Fragmentbibliothek Generierung

Definition reaktiver Gruppen

```
sdfilter -nonreactive
```

```
{ The molecule must not contain groups in a conservative list of reactive groups:
acrylamide, acyclic diketyl, acyl halide, aldehyde, aliphatic imine, aliphatic ketone,
aliphatic nitro, aliphatic (thio)ester, alkyl halide, anhydride, azide, aziridine,
beta-heterounsubstituted carbonyl, epoxide, halopyrimidine, heteroallyl, iso(thio)cyanate,
maleimide, michael acceptor, perhalo ketone, phosphonate ester, phospho-, sulfonate
ester, thiol, thio(urea), O-O/N-N/O-S/O-N single bonds, and transition metal.}
```

Definitionen unerwünschter Protonierungs-, und Tautomerisierungszustände (SMARTS-Definitionen)

```
-expr "F<7
-expr "[#G7!F]<4
-expr "[N+](=[OX1])[OX1]<3
-expr "[CX4Q4](F)(F)[CX4Q4](F)(F)[CX4Q4](F)(F)-0
-expr "[CX4Q2]!@[CX4Q2]!@[CX4Q2]!@[CX4Q2]!@[CX4Q2]-0
-expr "[OX2Q2][SX4]([OX1])([OX1])0-0
-expr "[N;!$(* [#G6])]=[CX3][#G6]-0
-expr "[OX2Q2][N+Q3](=[OX1])[OX1]-0
-expr "[#6+!$(C(N)(N))]-0
-expr "[N+X3](=C[NX3])OC=0-0
-expr "[NX3](C=[N+X3])OC=0-0
-expr "[O!i]!@[N!$(*[i])]-0
-expr "[#6-]-0
-expr "[#6-]-0
```

```
-expr "[#6++] -0  
-expr "[0+] -0  
-expr "[#6r6]1-***-***-1=*-0  
-expr "[#6r6]1-***-***-1=*-0  
-expr "[!D0!D1!D2!D3!D4] -0
```

I.2 Einstellungen *Conformation Import*

```
conf_Import input file  
outputFile: output file  
outputFileSplit: 700,  
outputFileOpenDBV: 0,  
outputRefine: 0,  
outputPacked: 0,  
outputStrainLimit: 4,  
outputConformationLimit: 300,  
mode: 'normal',  
newFragmentFile: 'frag/conf_unk.mdb',  
newFragmentReset: 0,  
logFilteredCompounds : 0,  
logmdb: 'log/conf_log.mdb',  
logOpenDBV: 0,  
logmdbkey: 0,  
filterAcidBase: 0,  
descexpr: '',  
forceTransAmide: 1,  
forceTransEster: 1,  
forceChair: 1,  
forceTransConjugatedEster: 1,  
forceTransVinyl: 1,  
forceOriginalCC: 1,  
forceOriginalCN: 0,  
forceOriginalNN: 0,  
forceOriginalChirality: 0,  
forceTransUrea: 1,  
fragStrainLimit: 4,  
skelStrainLimit: 6,
```

```
mmStrainLimit: 7,  
mmFailureLimit: 30,  
mmIterationLimit: 500,  
mmIterationLimitMM: 200,  
mmGradientTestMM: 0.001,  
mmSuperposeRMSD: 0.30,  
mmConformationLimit: 300,  
saveFragmentSkeleton: 0,  
noStochasticSearch: 0,  
sampleFailures: 1,  
timeout: 600,  
silent: 0  
;
```


Anhang J

Protokolle der AMBER MD-Simulationen

J.1 Energieminimierung in *sander*

Energieminimierung Protein-Ligand-Komplex

```
&cntrl
imin=1, maxcyc=500, ncyc=100,
cut=9999.0,
ntb=0,
ntc=1,
ntf=1,
ntpr=10,
igb=1
/
```

Energieminimierung Solvens

```
&cntrl
imin = 1,
maxcyc = 1000,
ncyc = 200,
ntb = 1,
ntr=1, restraintmask="!:WAT & !@H=", restraint_wt=500.0

cut = 8.0
ntpr=10
/
```

J.2 Äquilibrierung

Heizen des Solvens von 100 auf 300 Kelvin

```
&cntrl
imin = 0, irest = 0, ntx = 1,
ntb = 1, cut = 8.0, ntr = 1,
ntc = 2, ntf = 2,
tempi = 100.0, temp0 = 300.0,
ntt = 3, gamma_ln = 1.0,
nstlim = 250000, dt = 0.002
ntpr = 5000, ntwx = 5000, ntwr = 1000
ntr=1, restraintmask="!:WAT & !@H=", restraint_wt=100.0
nmropt=1,
ig=-1
/
```

Kühlen des Solvens von 300 auf 100 Kelvin

```
&cntrl
imin = 0, irest = 1, ntx = 5,
cut = 8.0, ntb = 1,
ntr = 1,
ntc = 2, ntf = 2,
tempi = 300.0, temp0 = 100.0,
ntt = 3, gamma_ln = 1.0,
nstlim = 125000, dt = 0.002
ntpr = 5000, ntwx = 5000, ntwr = 5000
ntr=1, restraintmask="!:WAT & !@H=", restraint_wt=10.0
nmropt=1,
ig=-1
/
```

Heizen des Protein-Ligand-Komplexes und des Solvens von 100 auf 300 Kelvin

```
&cntrl
imin=0, ntx=5, irest=0,
ntf=2, ntc=2,
tempi=100.0, temp0=300.0,
cut=8.0, ntb=1, ntp=0,
ntt=3, gamma_ln=1.0,
```



```

nstlim = 250000, dt = 0.002,
ntpr = 5000, ntwx = 5000, ntwr = 5000
ntr=1, restraintmask="!:WAT & !@H=", restraint_wt=1.0
nmropt=1,
ig=-1
/

```

Äquilibrieren des Drucks

```

&cntrl
imin = 0, irest = 1, ntx = 5,
ntb = 2, pres0 = 1.0, ntp = 1,
taup = 5.0,
cut = 8.0, ntr = 0,
ntc = 2, ntf = 2,
tempi = 300.0, temp0 = 300.0,
ntt = 3, gamma_ln = 1.0,
nstlim = 250000, dt = 0.002,
ntpr = 5000, ntwx = 5000, ntwr = 5000
iwrap = 1
/

```

J.3 Produktionsläufe

25 ns bzw. 50 ns Simulationen

```

&cntrl imin = 0, irest = 1, ntx = 5,
ntb = 2, pres0 = 1.0, ntp = 1,
taup = 2.0,
cut = 8.0, ntr = 0,
ntc = 2, ntf = 2,
tempi = 300.0, temp0 = 300.0,
ntt = 3, gamma_ln = 1.0,
nstlim = 12500000, dt = 0.002, # 25 ns
nstlim = 25000000, dt = 0.002, # 50 ns
ntpr = 500, ntwx = 500, ntwr = 500000
/

```

100 ns Simulation zur Analyse der Wassereigenschaften in GIST

```
&cntrl
imin = 0,  irest = 0,  ntx = 1,
ntb = 2,  pres0 = 1.0,  ntp = 1,  taup = 2.0,
cut = 8.0,  ig = -1,
ntc = 2,  ntf = 2,
tempi = 300.0,  temp0 = 300.0,
ntt = 3,  gamma_ln = 1.0,
nstlim = 50000000,  dt = 0.002,
ntpr = 5000,  ntwx = 5000,  ntwr = 500000,
iwrap=1,  nscm = 1000
ntr=1,  restraintmask="!:WAT & !@H=",  restraint_wt=10.0
/
```

Literatur

1. Wyk, B.-E., Wink, C. & Wink, M. in *Handbuch der Arzneipflanzen* 8–13 (2015).
2. J. P. Erzberger & J. M. Berger. Evolutionary relationships and structural mechanisms of AAA+ proteins. *Annual Review of Biophysics and Biomolecular Structure* **35**, 93–114 (2006).
3. Ogura, T. & Wilkinson, A. J. AAA+ superfamily ATPases: Common structure-diverse function. *Genes to Cells* **6**, 575–597 (2001).
4. Hanson, P. I. & Whiteheart, S. W. AAA+ proteins: Have engine, will work. *Nature Reviews Molecular Cell Biology* **6**, 519–529 (2005).
5. Moir, D., Stewart, S. E., Osmond, B. C. & Botstein, D. Cold-sensitive cell-division-cycle mutants of yeast: isolation, properties, and pseudoreversion studies. *Genetics* **100**, 547–563 (1982).
6. Barthelme, D. & Sauer, R. T. Origin and Functional Evolution of the Cdc48/p97/VCP AAA+ Protein Unfolding and Remodeling Machine. *Journal of Molecular Biology* **428**, 1861–1869 (2016).
7. Dalal, S. & Hanson, P. I. Membrane traffic: What drives the AAA motor? *Cell* **104**, 5–8 (2001).
8. Zhang, X., Shaw, A., Bates, P. A., Newman, R. H., Gowen, B., Orlova, E., Gorman, M. A., Kondo, H., Dokurno, P., Lally, J., Leonard, G., Meyer, H., Van Heel, M. & Freemont, P. S. Structure of the AAA ATPase p97. *Molecular Cell* (2000).
9. DeLaBarre, B. & Brunger, A. T. Complete structure of p97/valosin-containing protein reveals communication between nucleotide domains. *Nature Structural Biology* **10**, 856–863 (2003).
10. Hänzelmann, P. & Schindelin, H. Structural Basis of ATP Hydrolysis and Intersubunit Signaling in the AAA+ ATPase p97. *Structure* **24**, 127–139 (2016).
11. Buchberger, A., Schindelin, H. & Hänzelmann, P. Control of p97 function by cofactor binding. *FEBS Letters* **589**, 2578–2589 (2015).

12. Hänzelmann, P. & Schindelin, H. The Structural and Functional Basis of the p97/Valosin-containing Protein (VCP)-interacting Motif (VIM) mutually exclusive binding of cofactors to the N-terminal domain of p97. *Journal of Biological Chemistry* **286**, 38679–38690 (2011).
13. Banerjee, S., Bartesaghi, A., Merk, A., Rao, P., Bulfer, S. L., Yan, Y., Green, N., Mroczkowski, B., Neitz, R. J., Wipf, P., Falconieri, V., Deshaies, R. J., Milne, J. L. S., Huryn, D., Arkin, M. & Subramaniam, S. 2.3 Å resolution cryo-EM structure of human p97 and mechanism of allosteric inhibition. *Science (New York, N.Y.)* **351**, 871–5 (2016).
14. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. The Protein Data Bank. *Nucleic Acids Research* **28**, 235–242 (Jan. 2000).
15. Brandman, O. & Hegde, R. S. Ribosome-associated protein quality control. *Nature Structural and Molecular Biology* **23**, 7–15 (2016).
16. Stach, L. & Freemont, P. S. The AAA+ ATPase p97, a cellular multitool. *Biochemical Journal* **474**, 2953–2976 (2017).
17. Ye, Y., Tang, W. K., Zhang, T. & Xia, D. A Mighty “Protein Extractor” of the Cell: Structure and Function of the p97/CDC48 ATPase. *Frontiers in Molecular Biosciences* **4**, 39 (2017).
18. Swatek, K. N. & Komander, D. Ubiquitin modifications. *Nature Publishing Group*, 399–422 (2016).
19. Boom, J. V. D. & Meyer, H. Review VCP / p97-Mediated Unfolding as a Principle in Protein Homeostasis and Signaling. *Molecular Cell* **69**, 182–194 (2017).
20. Wu, X. & Rapoport, T. A. Mechanistic insights into ER-associated protein degradation. *Current Opinion in Cell Biology* **53**. Membrane Trafficking, 22–28 (2018).
21. Tabara, K., Iwata, Y. & Koizumi, N. in *The Plant Endoplasmic Reticulum : Methods and Protocols* (Hrsg. Hawes, C. & Kriechbaumer, V.) 223–230 (Springer New York, New York, NY, 2018).
22. Franz, A., Ackermann, L. & Hoppe, T. Ring of change: CDC48/p97 drives protein dynamics at chromatin. *Frontiers in Genetics* **7**, 1–14 (2016).
23. Xu, Y., Anderson, D. E. & Ye, Y. The HECT domain ubiquitin ligase HUWE1 targets unassembled soluble proteins for degradation. *Cell Discovery* **2**, 1–16 (2016).

24. Hänzelmann, P. & Schindelin, H. The Interplay of Cofactor Interactions and Post-translational Modifications in the Regulation of the AAA+ ATPase p97. *Frontiers in Molecular Biosciences* **4**, 21 (2017).
25. Buchberger, A., Howard, M. J., Proctor, M. & Bycroft, M. The UBX domain: a widespread ubiquitin-like module¹¹Edited by P. E. Wright. *Journal of Molecular Biology* **307**, 17–24 (2001).
26. Hänzelmann, P., Buchberger, A. & Schindelin, H. Hierarchical Binding of Cofactors to the AAA ATPase p97. *Structure* **19**, 833–843 (2011).
27. Lee, J. J., Park, J. K., Jeong, J., Jeon, H., Yoon, J. B., Kim, E. E. K. & Lee, K. J. Complex of fas-associated Factor 1 (FAF1) with valosin-containing protein (VCP)-Npl4-Ufd1 and polyubiquitinated proteins promotes endoplasmic reticulum-associated degradation (ERAD)S. *Journal of Biological Chemistry* **288**, 6998–7011 (2013).
28. Kim, K. H., Kang, W., Suh, S. W. & Yang, J. K. Crystal structure of FAF1 UBX domain in complex with p97/VCP N domain reveals a conformational change in the conserved FcisP touch-turn motif of UBX domain. *Proteins: Structure, Function, and Bioinformatics* **79**, 2583–2587 (2011).
29. Hui Li, Z., Wang, Y., Xu, M. & Jiang, T. Crystal structures of the UBX domain of human UBXD7 and its complex with p97 ATPase. *Biochemical and Biophysical Research Communications* **486**, 94–100 (2017).
30. Dreveny, I., Kondo, H., Uchiyama, K., Shaw, A., Zhang, X. & Freemont, P. S. Structural basis of the interaction between the AAA ATPase p97/VCP and its adaptor protein p47. *EMBO Journal* **23**, 1030–1039 (2004).
31. Ernst, R., Mueller, B., Ploegh, H. L. & Schlieker, C. The Otubain YOD1 Is a Deubiquitinating Enzyme that Associates with p97 to Facilitate Protein Dislocation from the ER. *Molecular Cell* **36**, 28–38 (2009).
32. Hao, Q., Jiao, S., Shi, Z., Li, C., Meng, X., Zhang, Z., Wang, Y., Song, X., Wang, W., Zhang, R., Zhao, Y., Wong, C. C. & Zhou, Z. A non-canonical role of the p97 complex in RIG-I antiviral signaling. *The EMBO Journal* **34**, 2903–2920 (2015).
33. Lim, J. J., Lee, Y., Ly, T. T., Kang, J. Y., Lee, J.-G., An, J. Y., Youn, H.-S., Park, K. R., Kim, T. G., Yang, J. K., Jun, Y. & Eom, S. H. Structural insights into the interaction of p97 N-terminus domain and VBM in rhomboid protease, RHBDL4. *Biochemical Journal* **473**, 2863–2880 (Sep. 2016).
34. Tang, W. K. & Xia, D. Mutations in the human AAA+ chaperone p97 and related diseases. *Frontiers in Molecular Biosciences* **3**, 1–12 (2016).

35. Meyer, T. *Stephen Hawking – ein Nachruf aus ALS-Perspektive* <https://www.als-charite.de/stephen-hawking-ein-nachruf-aus-als-perspektive/> (2020).
36. Huryn, D. M., Kornfilt, D. J. & Wipf, P. P97: An Emerging Target for Cancer, Neurodegenerative Diseases, and Viral Infections. *Journal of Medicinal Chemistry* **63**, 1892–1907 (2020).
37. Pschyrembel, W. & der Pschyrembel-Redaktion des Verlages. *Pschyrembel Klinisches Wörterbuch* (De Gruyter, Berlin, Boston, 21 Sep. 2020).
38. Bastola, P., Oien, D. B., Cooley, M. & Chien, J. Emerging Cancer Therapeutic Targets in Protein Homeostasis. *AAPS Journal* **20** (2018).
39. Anderson, D. J., Moigne, R. L., Djakovic, S., Kumar, B., Rice, J., Wong, S., Wang, J., Yao, B., Valle, E., Soly, S. K. V., Soriano, F., Menon, M.-k., Wu, Z. Y., Kampmann, M., Chen, Y., Weissman, J. S., Aftab, B. T., Yakes, F. M., Shawver, L., Zhou, H.-j., Wustrow, D. & Rolfe, M. Through Disruption of Protein Homeostasis. *Cancer Cell* **28**, 653–665 (2016).
40. Gugliotta, G., Sudo, M., Cao, Q., Lin, D. C., Sun, H., Takao, S., Le Moigne, R., Rolfe, M., Gery, S., Müschen, M., Cavo, M. & Koeffler, H. P. Valosin-Containing Protein/p97 as a Novel Therapeutic Target in Acute Lymphoblastic Leukemia. *Neoplasia (United States)* **19**, 750–761 (2017).
41. Tang, W. K., Odzorig, T., Jin, W. & Xia, D. Structural basis of p97 inhibition by the site-selective anticancer compound CB-5083. *Molecular Pharmacology* **95**, 286–293 (2019).
42. Wong, H. H., Kumar, P., Tay, F. P. L., Moreau, D., Liu, D. X. & Bard, F. Genome-Wide Screen Reveals Valosin-Containing Protein Requirement for Coronavirus Exit from Endosomes. *Journal of Virology* **89**, 11116–11128 (2015).
43. Chapman, E., Maksim, N., De La Cruz, F. & La Clair, J. J. Inhibitors of the AAA+ chaperone p97. *Molecules* **20**, 3027–3049 (2015).
44. Sui, X., Pan, M. & Li, Y.-M. Insights into Design of p97-targeting Small Molecules from Structural Studies on p97 Functional Mechanism. *Current Medicinal Chemistry* **26**, 1–18 (2019).
45. Bursavich, M. G., Parker, D. P., Willardsen, J. A., Gao, Z. H., Davis, T., Ostannin, K., Robinson, R., Peterson, A., Cimborá, D. M., Zhu, J. F. & Richards, B. 2-Anilino-4-aryl-1,3-thiazole inhibitors of valosin-containing protein (VCP or p97). *Bioorganic and Medicinal Chemistry Letters* **20**, 1677–1679 (2010).

46. Chou, T. F., Brown, S. J., Minond, D., Nordin, B. E., Li, K., Jones, A. C., Chasse, P., Porubsky, P. R., Stoltz, B. M., Schoenen, F. J., Patricelli, M. P., Hodder, P., Rosen, H. & Deshaies, R. J. Reversible inhibitor of p97, DBeQ, impairs both ubiquitin-dependent and autophagic protein clearance pathways. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 4834–4839 (2011).
47. Chou, T.-F., Li, K., Frankowski, K. J., Schoenen, F. J. & Deshaies, R. J. Structure–Activity Relationship Study Reveals ML240 and ML241 as Potent and Selective Inhibitors of p97 ATPase. *ChemMedChem* **8**, 297–312 (2013).
48. Therapeutics, C. *Cleave Therapeutics Announces Commencement of a Phase 1 Clinical Study of CB-5339, A Valosin-Containing Protein (VCP)/p97 Inhibitor, in Patients with Acute Myeloid Leukemia (AML) and Myelodysplastic Syndrome (MDS)* <https://pipelinereview.com/index.php/2020072875424/Small-Molecules/Cleave-Therapeutics-Announces-Commencement-of-a-Phase-1-Clinical-Study-of-CB-5339-A-Valosin-Containing-Protein-VCP/p97-Inhibitor-in-Patients-with-Acute-Myeloid-Leukemia-A.html> (2020).
49. Chou, T. F. & Deshaies, R. J. Quantitative cell-based protein degradation assays to identify and classify drugs that target the ubiquitin-proteasome system. *Journal of Biological Chemistry* **286**, 16546–16554 (2011).
50. Magnaghi, P., D'Alessio, R., Valsasina, B., Avanzi, N., Rizzi, S., Asa, D., Gasparri, F., Cozzi, L., Cucchi, U., Orrenius, C., Polucci, P., Ballinari, D., Perre-ra, C., Leone, A., Cervi, G., Casale, E., Xiao, Y., Wong, C., Anderson, D. J., Galvani, A., Donati, D., O'Brien, T., Jackson, P. K. & Isacchi, A. Covalent and allosteric inhibitors of the ATPase VCP/p97 induce cancer cell death. *Nature Chemical Biology* **9**, 548–559 (2013).
51. Tillotson, J., Bashyal, B. P., Kang, M., Shi, T., De La Cruz, F., Gunatillaka, A. A. & Chapman, E. Selective inhibition of p97 by chlorinated analogues of dehydrocurvularin. *Organic and Biomolecular Chemistry* **14**, 5918–5921 (2016).
52. Ding, R., Zhang, T., Wilson, D. J., Xie, J., Williams, J., Xu, Y., Ye, Y. & Chen, L. Discovery of Irreversible p97 Inhibitors. *Journal of Medicinal Chemistry* **62**, 2814–2829 (2019).
53. Polucci, P., Magnaghi, P., Angiolini, M., Asa, D., Avanzi, N., Badari, A., Bertrand, J., Casale, E., Cauteruccio, S., Cirila, A., Cozzi, L., Galvani, A., Jackson, P. K., Liu, Y., Magnuson, S., Malgesini, B., Nuvoloni, S., Orrenius, C., Sirtori, F. R., Riceputi, L., Rizzi, S., Trucchi, B., O'Brien, T., Isacchi,

- A., Donati, D. & D'Alessio, R. Alkylsulfanyl-1,2,4-triazoles, a New Class of allosteric valosine containing protein inhibitors. Synthesis and structure-activity relationships. *Journal of Medicinal Chemistry* **56**, 437–450 (2013).
54. Alvarez, C., Arkin, M. R., Bulfer, S. L., Colombo, R., Kovaliov, M., Laporte, M. G., Lim, C., Liang, M., Moore, W. J., Neitz, R. J., Yan, Y., Yue, Z., Huryn, D. M. & Wipf, P. Structure-Activity Study of Bioisosteric Trifluoromethyl and Pentafluorosulfanyl Indole Inhibitors of the AAA ATPase p97. *ACS Medicinal Chemistry Letters* **6**, 1225–1230 (2015).
55. Burnett, J. C., Lim, C., Peyser, B. D., Samankumara, L. P., Kovaliov, M., Colombo, R., Bulfer, S. L., LaPorte, M. G., Hermone, A. R., McGrath, C. F., Arkin, M. R., Gussio, R., Huryn, D. M. & Wipf, P. A threonine turnstile defines a dynamic amphiphilic binding motif in the AAA ATPase p97 allosteric binding site. *Organic and Biomolecular Chemistry* **15**, 4096–4114 (2017).
56. Laporte, M. G., Burnett, J. C., Colombo, R., Bulfer, S. L., Alvarez, C., Chou, T. F., Neitz, R. J., Green, N., Moore, W. J., Yue, Z., Li, S., Arkin, M. R., Wipf, P. & Huryn, D. M. Optimization of Phenyl Indole Inhibitors of the AAA+ ATPase p97. *ACS Medicinal Chemistry Letters* **9**, 1075–1081 (2018).
57. Pöhler, R., Krahn, J. H., van den Boom, J., Dobrynin, G., Kaschani, F., Egenweiler, H. M., Zenke, F. T., Kaiser, M. & Meyer, H. A Non-Competitive Inhibitor of VCP/p97 and VPS4 Reveals Conserved Allosteric Circuits in Type I and II AAA ATPases. *Angewandte Chemie - International Edition* **57**, 1576–1580 (2018).
58. Alvarez, C., Bulfer, S. L., Chakrasali, R., Chimenti, M. S., Deshaies, R. J., Green, N., Kelly, M., LaPorte, M. G., Lewis, T. S., Liang, M., Moore, W. J., Neitz, R. J., Peshkov, V. A., Walters, M. A., Zhang, F., Arkin, M. R., Wipf, P. & Huryn, D. M. Allosteric Indole Amide Inhibitors of p97: Identification of a Novel Probe of the Ubiquitin Pathway. *ACS Medicinal Chemistry Letters* **7**, 182–187 (2016).
59. Wang, Q., Shinkre, B. A., Gu Lee, J., Weniger, M. A., Liu, Y., Chen, W., Wiestner, A., Trenkle, W. C. & Ye, Y. The ERAD inhibitor eeyarestatin I is a bifunctional compound with a membrane-binding domain and a p97/VCP inhibitory group. *PLoS ONE* **5** (2010).
60. Segura-Cabrera, A., Tripathi, R., Zhang, X., Gui, L., Chou, T. F. & Komurov, K. A structure-and chemical genomics-based approach for repositioning of drugs against VCP/p97 ATPase. *Scientific Reports* **7**, 1–10 (2017).

61. Chimenti, M. S., Bulfer, S. L., Neitz, R. J., Renslo, A. R., Jacobson, M. P., James, T. L., Arkin, M. R. & Kelly, M. J. A fragment-based ligand screen against part of a large protein machine: The ND1 domains of the AAA+ ATPase p97/VCP. *Journal of Biomolecular Screening* **20**, 788–800 (2015).
62. He Chen, Q., liang Fu, M., miao Chen, M., Liu, J., jie Liu, X., qing He, G. & cheng Pu, S. Preparative isolation and purification of xanthohumol from hops (*Humulus lupulus* L.) by high-speed counter-current chromatography. *Food Chemistry* **132**, 619–623 (2012).
63. Kaminski, D. M., Gaweda, K., Arczewska, M., Senczyna, B., Gagos, M. A kinetic study of xanthohumol cyclization to isoxanthohumol - A role of water. *Journal of Molecular Structure* **1139**, 10–16 (2017).
64. Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. PII of original article: S0169-409X(96)00423-1. The article was originally published in *Advanced Drug Delivery Reviews* **23** (1997) 3–25.1. *Advanced Drug Delivery Reviews* **46**. Special issue dedicated to Dr. Eric Tomlinson, *Advanced Drug Delivery Reviews*, A Selection of the Most Highly Cited Articles, 1991-1998, 3–26 (2001).
65. Stevens, J. F. & Page, J. E. Xanthohumol and related prenylflavonoids from hops and beer: To your good health! *Phytochemistry* **65**, 1317–1330 (2004).
66. Liu, M., Hansen, P. E., Wang, G., Qiu, L., Dong, J., Yin, H., Qian, Z., Yang, M. & Miao, J. Pharmacological profile of xanthohumol, a prenylated flavonoid from hops (*Humulus lupulus*). *Molecules* **20**, 754–779 (2015).
67. Jiang, C. H., Sun, T. L., Xiang, D. X., Wei, S. S. & Li, W. Q. Anticancer activity and mechanism of xanthohumol: A prenylated flavonoid from hops (*Humulus lupulus* L.) *Frontiers in Pharmacology* **9**, 1–13 (2018).
68. Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., Light, Y., McGlinchey, S., Michalovich, D., Al-Lazikani, B. & Overington, J. P. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Research* **40**, D1100–D1107 (Sep. 2011).
69. Sasazawa, Y., Kanagaki, S., Tashiro, E., Nogawa, T., Muroi, M., Kondoh, Y., Osada, H. & Imoto, M. Xanthohumol impairs autophagosome maturation through direct inhibition of valosin-containing protein. *ACS Chemical Biology* **7**, 892–900 (2012).

70. Kanoh, N., Honda, K., Simizu, S., Muroi, M. & Osada, H. Photo-cross-linked small-molecule affinity matrix for facilitating forward and reverse chemical genetics. *Angewandte Chemie - International Edition* **44**, 3559–3562 (2005).
71. Scott, D. E., Bayly, A. R., Abell, C. & Skidmore, J. Small molecules, big targets: Drug discovery faces the protein-protein interaction challenge. *Nature Reviews Drug Discovery* **15**, 533–550 (2016).
72. Clackson, T. & Wells, J. A. A hot spot of binding energy in a hormone-receptor interface. *Science* **267**, 383–386 (1995).
73. Moreira, I. S., Fernandes, P. A. & Ramos, M. J. Hot spots-A review of the protein-protein interface determinant amino-acid residues. *Proteins: Structure, Function, and Bioinformatics* **68**, 803–812 (2007).
74. Sundberg, E. in *Computational Protein-Protein Interactions* 43–56 (CRC Press, 2009).
75. Smith, M. C. & Gestwicki, J. E. Features of protein-protein interactions that translate into potent inhibitors: topology, surface area and affinity. *Expert reviews in molecular medicine* **14**, 1–20 (2012).
76. Blundell, T. L., Sibanda, B. L., Montalvão, R. W., Brewerton, S., Chelliah, V., Worth, C. L., Harmer, N. J., Davies, O. & Burke, D. Structural biology and bioinformatics in drug design: Opportunities and challenges for target identification and lead discovery. *Philosophical Transactions of the Royal Society B: Biological Sciences* **361**, 413–423 (2006).
77. Macalino, S. J. Y., Gosu, V., Hong, S. & Choi, S. Role of computer-aided drug design in modern drug discovery. *Archives of Pharmacal Research* **38**, 1686–1701 (2015).
78. Brogi, S., Ramalho, T. C., Kuca, K., Medina-Franco, J. L. & Valko, M. Editorial: In silico Methods for Drug Design and Discovery. *Frontiers in Chemistry* **8**, 1–5 (2020).
79. Pinzi, L. & Rastelli, G. Molecular docking: Shifting paradigms in drug discovery. *International Journal of Molecular Sciences* **20** (2019).
80. Mak, K. K. & Pichika, M. R. Artificial intelligence in drug development: present status and future prospects. *Drug Discovery Today* **24**, 773–780 (2019).
81. Bender, A. & Brown, N. Special Issue: Cheminformatics in Drug Discovery. *ChemMedChem* **13**, 467–469 (2018).

82. Ertl, P., Rohde, B. & Selzer, P. Fast Calculation of Molecular Polar Surface Area as a Sum of Fragment-Based Contributions and Its Application to the Prediction of Drug Transport Properties. *Journal of Medicinal Chemistry* **43**. PMID: 11020286, 3714–3717 (2000).
83. Rekker, R. F., Laak, A. M. & Mannhold, R. On the Reliability of Calculated Log P-values: Rekker, Hansch/Leo and Suzuki Approach. *Quantitative Structure-Activity Relationships* **12**, 152–157 (1993).
84. Rácz, A., Bajusz, D. & Héberger, K. Life beyond the Tanimoto coefficient: Similarity measures for interaction fingerprints. *Journal of Cheminformatics* **10**, 1–12 (2018).
85. Todeschini, R. & Consonni, V. in *Handbook of Molecular Descriptors* (John Wiley & Sons, Ltd, 2000).
86. Leach, A. R. & Gillet, V. in *An Introduction To Chemoinformatics* 53–74 (Springer Netherlands, Dordrecht, 2007).
87. Van Drie, J. H. Monty Kier and the origin of the pharmacophore concept. *Internet Electronic Journal of Molecular Design* **11**, 271–279 (2012).
88. Wermuth, C. G., Ganellin, C. R., Lindberg, P. & Mitscher, L. a. Glossary for Chemists of Terms Used in Medicinal Chemistry. *Pure & Appl. Chem.* **70**, 1129–1143 (1998).
89. Markt, P., Schuster, D. & Langer, T. in *Virtual Screening* 115–152 (John Wiley & Sons, Ltd, 2011).
90. Goodford, P. J. A Computational Procedure for Determining Energetically Favorable Binding Sites on Biologically Important Macromolecules. *Journal of Medicinal Chemistry* **28**, 849–857 (1985).
91. *Handbuch-Abschnitt: Interactionpotential maps* 2019.01 (Chemical Computing Group ULC, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2021).
92. *Molecular Operating Environment (MOE)* 2019.01 (Chemical Computing Group ULC, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2021).
93. Markt, P., Schuster, D. & Langer, T. in *Virtual Screening* 115–152 (John Wiley & Sons, Ltd, 2011).
94. Kitchen, D. B., Decornez, H., Furr, J. R. & Bajorath, J. Docking and scoring in virtual screening for drug discovery: Methods and applications. *Nature Reviews Drug Discovery* **3**, 935–949 (2004).

95. Rognan, D. in *Virtual Screening* 153–176 (John Wiley & Sons, Ltd, 2011).
96. Du, X., Li, Y., Xia, Y. L., Ai, S. M., Liang, J., Sang, P., Ji, X. L. & Liu, S. Q. Insights into protein–ligand interactions: Mechanisms, models, and methods. *International Journal of Molecular Sciences* **17**, 1–34 (2016).
97. Mccammon, J. A., Miao, Y., Amaro, R. E., Baudry, J., Chodera, J. & Smith, J. C. Biophysical Perspective Ensemble Docking in Drug Discovery. *Biophysical Journal* **114**, 2271–2278 (2018).
98. Verdonk, M. L., Chessari, G., Cole, J. C., Hartshorn, M. J., Murray, C. W., Nissink, J. W. M., Taylor, R. D. & Taylor, R. Modeling water molecules in protein-ligand docking using GOLD. *Journal of Medicinal Chemistry* **48**, 6504–6515 (2005).
99. Uehara, S. & Tanaka, S. AutoDock-GIST: Incorporating thermodynamics of active-site water into scoring function for accurate protein-ligand docking. *Molecules* **21** (2016).
100. Jones, G., Willett, P., Glen, R. C., Leach, A. R. & Taylor, R. Development and validation of a genetic algorithm for flexible docking. *Journal of Molecular Biology* **267**, 727–748 (1997).
101. Jones, G., Willett, P. & Glen, R. C. A genetic algorithm for flexible molecular overlay and pharmacophore elucidation. *Journal of Computer-Aided Molecular Design* **9**, 532–549 (1995).
102. Raton, B. in *In Silico Drug Discovery and Design* 155–176 (CRC Press, 2016).
103. Pason, L. P. & Sotriffer, C. A. Empirical Scoring Functions for Affinity Prediction of Protein-ligand Complexes. *Molecular Informatics* **35**, 541–548 (2016).
104. Korb, O., Stützle, T. & Exner, T. E. Empirical scoring functions for advanced Protein-Ligand docking with PLANTS. *Journal of Chemical Information and Modeling* **49**, 84–96 (2009).
105. KORB, O., STÜTZLE, T e Exner, T. E. Empirical Scoring Functions for Advanced Protein-Ligand Docking with Plants. *J. Chem. Inf. Model.* **49**, 84–96 (2009).
106. Verdonk, M. L., Cole, J. C., Hartshorn, M. J., Murray, C. W. & Taylor, R. D. Improved protein-ligand docking using GOLD. *Proteins: Structure, Function and Genetics* **52**, 609–623 (2003).
107. Clark, M., Cramer, R. D. & Van Opdenbosch, N. Validation of the general purpose tripos 5.2 force field. *Journal of Computational Chemistry* **10**, 982–1012 (1989).

108. Mooij, W. T. & Verdonk, M. L. General and targeted statistical potentials for protein-ligand interactions. *Proteins: Structure, Function and Genetics* **61**, 272–287 (2005).
109. Neudert, G. & Klebe, G. DSX: A knowledge-based scoring function for the assessment of protein-ligand complexes. *Journal of Chemical Information and Modeling* **51**, 2731–2745 (2011).
110. Groom, C. R., Bruno, I. J., Lightfoot, M. P. & Ward, S. C. The Cambridge Structural Database. *Acta Crystallographica Section B* **72**, 171–179 (2016).
111. Feynman, R. P. *Feynman Vorlesungen über Physik* (Oldenbourg Wissenschaftsverlag, 2009).
112. Fischer, E. Einfluss der Configuration auf die Wirkung der Enzyme. *Berichte der deutschen chemischen Gesellschaft* **27**, 2985–2993 (1894).
113. McCammon, J. A. & Durrant, J. D. Molecular dynamics simulations and drug discovery. *BMC Biology* **9** (2011).
114. Hollingsworth, S. A. & Dror, R. O. Molecular Dynamics Simulation for All. *Neuron* **99**, 1129–1143 (2018).
115. Leach, A. R. in *Molecular Modelling, Principles and Applications* 165–252 (Pearson, 2001).
116. Case, D., Belfon, K., Ben-Shalom, I., Brozell, S., Cerutti, D., Cheatham, T., III, Cruzeiro, V., Darden, T., Duke, R., Giambasu, G., Gilson, M., Gohlke, H., Goetz, A., Harris, R., Izadi, S., Izmailov, S., Kasavajhala, K., Kovalenko, A., Krasny, R., Kurtzman, T., Lee, T., LeGrand, S., Li, P., Lin, C., Liu, J., Luchko, T., Luo, R., Man, V., Merz, K., Miao, Y., Mikhailovskii, O., Monard, G., Nguyen, H., Onufriev, A., Pan, F., Pantano, S., Qi, R., Roe, D., Roitberg, A., Sagui, C., Schott-Verdugo, S., Shen, J., Simmerling, C., N.R. Skrynnikov, Smith, J., Swails, J., Walker, R., Wang, J., Wilson, L., Wolf, R., Wu, X., Xiong, Y., Xue, Y., York, D. & Kollman, P. *Amber 2019 Reference Manual (Covers Amber18 and AmberTools19)* 166 (University of California, San Francisco, 2019).
117. Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E. & Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *Journal of Chemical Theory and Computation* **11**, 3696–3713 (2015).
118. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general Amber force field. *Journal of Computational Chemistry* **25**, 1157–1174 (2004).

119. Halgren, T. A. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry* **17**, 490–519 (1996).
120. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **79**, 926–935 (1983).
121. Jorgensen, W. L. & Tirado-Rives, J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society* **110**. PMID: 27557051, 1657–1666 (1988).
122. Jorgensen, W. L., Maxwell, D. S. & Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *Journal of the American Chemical Society* **118**, 11225–11236 (1996).
123. Leach, A. R. in *Molecular Modelling, Principles and Applications* 353–409 (Pearson, 2001).
124. Case, D., Cheatham, T. & Darden, T. e. a. The Amber biomolecular simulation programs. *J Comput Chem.* **26**, 1668–1688 (2005).
125. Ryckaert, J.-P., Ciccotti, G. & Berendsen, H. J. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics* **23**, 327–341 (1977).
126. Hopkins, C. W., Le Grand, S., Walker, R. C. & Roitberg, A. E. Long-time-step molecular dynamics through hydrogen mass repartitioning. *Journal of Chemical Theory and Computation* **11**, 1864–1874 (2015).
127. Grest, G. S. & Kremer, K. Molecular dynamics simulation for polymers in the presence of a heat bath. *Physical Review A* **33**, 3628–3631 (1986).
128. Martyna, G. J., Tobias, D. J. & Klein, M. L. Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics* **101**, 4177–4189 (1994).
129. Frenkel, D. & Smit, B. *Understanding Molecular Simulation* 2nd (Academic Press, Inc., USA, 2001).
130. Leach, A. R. *Molecular Modelling, Principles and Applications* (Pearson, 2001).
131. ESA. *How many stars are there in the Universe?* https://www.esa.int/Science_Exploration/Space_Science/Herschel/How_many_stars_are_there_in_the_Universe (2020).
132. Barker, A., Kettle, J. G., Nowak, T. & Pease, J. E. Expanding medicinal chemistry space. *Drug Discovery Today* **18**, 298–304 (2013).

133. Erlanson, D. A., Fesik, S. W., Hubbard, R. E., Jahnke, W. & Jhoti, H. Twenty years on: The impact of fragments on drug discovery. *Nature Reviews Drug Discovery* **15**, 605–619 (2016).
134. Hubbard, R. E. in *Fragment-based Drug Discovery Lessons and Outlook* 1–36 (John Wiley & Sons, Ltd, 2016).
135. Ruddigkeit, L., van Deursen, R., Blum, L. C. & Reymond, J.-L. Enumeration of 166 Billion Organic Small Molecules in the Chemical Universe Database GDB-17. *Journal of Chemical Information and Modeling* **52**. PMID: 23088335, 2864–2875 (2012).
136. Hann, M. M., Leach, A. R. & Harper, G. Molecular Complexity and Its Impact on the Probability of Finding Leads for Drug Discovery. *Journal of Chemical Information and Computer Sciences* **41**. PMID: 11410068, 856–864 (2001).
137. Brewer, M., Ichihara, O., Kirchhoff, C., Schade, M. & Whittaker, M. in *Fragment-Based Drug Discovery* 39–62 (John Wiley & Sons, Ltd, 2008).
138. Jacquemard, C. & Kellenberger, E. A bright future for fragment-based drug discovery: what does it hold? *Expert Opinion on Drug Discovery* **14**, 413–416 (2019).
139. Beglov, D. & R., B. in *Computational Protein-Protein Interactions* 253–277 (CRC Press, 2009).
140. Braisted, A. C., Oslob, J. D., Delano, W. L., Hyde, J., McDowell, R. S., Waal, N., Yu, C., Arkin, M. R. & Raimundo, B. C. Discovery of a potent small molecule IL-2 inhibitor through fragment assembly. *Journal of the American Chemical Society* **125**, 3714–3715 (2003).
141. Guo, W., Wisniewski, J. A. & Ji, H. Hot spot-based design of small-molecule inhibitors for protein-protein interactions. *Bioorganic and Medicinal Chemistry Letters* **24**, 2546–2554 (2014).
142. R&M. *Fragment-based Drug Discovery Market: Library and Service Providers, 2020-2030* https://www.researchandmarkets.com/reports/5025058/fragment-based-drug-discovery-market-library-and?utm_source=dynamic&utm_medium=CI&utm_code=xwqjxw&utm_campaign=1396257+-+Global+Fragment-based+Drug+Discovery+Market+2020+%3a+Library+and+Service+Providers&utm_exec=cari18cid (2020).
143. Erlanson, D. *Poll results: affiliation, fragment-finding methods, and library size* <http://practicalfragments.blogspot.com/2014/01/poll-results-affiliation-fragment.html> (2020).

144. Congreve, M., Carr, R., Murray, C. & Jhoti, H. A 'Rule of Three' for fragment-based lead discovery? *Drug Discovery Today* **8**, 876–877 (2003).
145. Davies, D. R., Mamat, B., Magnusson, O. T., Christensen, J., Haraldsson, M. H., Mishra, R., Pease, B., Hansen, E., Singh, J., Zembower, D., Kim, H., Kiselyov, A. S., Burgin, A. B., Gurney, M. E. & Stewart, L. J. Discovery of leukotriene A4 hydrolase inhibitors using metabolomics biased fragment crystallography. *Journal of Medicinal Chemistry* **52**, 4694–4715 (2009).
146. Consulting, C. M. *Fragment Collections* https://www.cambridgemedchemconsulting.com/resources/hit_identification/fragment_collections.html (2020).
147. Firth, N. C., Brown, N. & Blagg, J. Plane of best fit: A novel method to characterize the three-dimensionality of molecules. *Journal of Chemical Information and Modeling* **52**, 2516–2525 (2012).
148. Bower, J., Pugliese, A. & Drysdale, M. in *Fragment-based Drug Discovery Lessons and Outlook* 99–118 (John Wiley & Sons, Ltd, 2016).
149. Hall, R. J., Mortenson, P. N. & Murray, C. W. Efficient exploration of chemical space by fragment-based screening. *Progress in Biophysics and Molecular Biology* **116**, 82–91 (2014).
150. Hung, A. W., Ramek, A., Wang, Y., Kaya, T., Wilson, J. A., Clemons, P. A. & Young, D. W. Route to three-dimensional fragments using diversity-oriented synthesis. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 6799–6804 (2011).
151. Walse, B., Turnbull, A. P. & Boyd, S. M. in *Applied Biophysics for Drug Discovery* 29–59 (John Wiley & Sons, Ltd, 2017).
152. Sperandio, O., Reynès, C. H., Camproux, A. C. & Villoutreix, B. O. Rationalizing the chemical space of protein-protein interaction inhibitors. *Drug Discovery Today* **15**, 220–229 (2010).
153. Morelli, X., Bourgeas, R. & Roche, P. Chemical and structural lessons from recent successes in protein-protein interaction inhibition (2P2I). *Current Opinion in Chemical Biology* **15**, 475–481 (2011).
154. Hann, M. M. & Oprea, T. I. Pursuing the leadlikeness concept in pharmaceutical research. *Current Opinion in Chemical Biology* **8**, 255–263 (2004).
155. Jhoti, H., Williams, G., Rees, D. C. & Murray, C. W. The 'rule of three' for fragment-based drug discovery: Where are we now? *Nature Reviews Drug Discovery* **12**, 644 (2013).

156. Akella, L. B. & DeCaprio, D. Cheminformatics approaches to analyze diversity in compound screening libraries. *Current Opinion in Chemical Biology* **14**, 325–330 (2010).
157. Maggiora, G., Vogt, M., Stumpfe, D. & Bajorath, J. Molecular similarity in medicinal chemistry. *Journal of Medicinal Chemistry* **57**, 3186–3204 (2014).
158. Robson-Tull, J. Biophysical screening in fragment-based drug design: A brief overview. *Bioscience Horizons* **11**, 1–12 (2018).
159. Erlanson, D. *Poll results: affiliation and fragment-finding methods in 2019* <http://practicalfragments.blogspot.com/2019/12/poll-results-affiliation-and-fragment.html> (2020).
160. Coyle, J. & Walser, R. Applied Biophysical Methods in Fragment-Based Drug Discovery. *SLAS Discovery* **25**, 471–490 (2020).
161. Thorne, N., Auld, D. S. & Inglese, J. Apparent activity in high-throughput screening: origins of compound-dependent assay interference. *Current Opinion in Chemical Biology* **14**, 315–324 (2010).
162. Wilson, B. A., Thornburg, C. C., Henrich, C. J., Grkovic, T. & O’Keefe, B. R. Creating and screening natural product libraries. *Natural product reports* **37**, 893–918 (2020).
163. Davies, T. G. & Tickle, I. J. in *Fragment-Based Drug Discovery and X-Ray Crystallography* (Hrsg. Davies, T. G. & Hyvönen, M.) 33–59 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2012).
164. Badger, J. in *Structure-Based Drug Discovery* (Hrsg. Tari, L. W.) 161–177 (Humana Press, Totowa, NJ, 2012).
165. Erlanson, D. A., Davis, B. J. & Jahnke, W. Fragment-Based Drug Discovery: Advancing Fragments in the Absence of Crystal Structures. *Cell Chemical Biology* **26**, 9–15 (2019).
166. Harner, M. J., Frank, A. O. & Fesik, S. W. Fragment-based drug discovery using NMR spectroscopy. *Journal of Biomolecular NMR* **56**, 65–75 (2013).
167. Krimm, I. in *Applied Biophysics for Drug Discovery* 175–195 (John Wiley & Sons, Ltd, 2017).
168. Monaco, S., Tailford, L. E., Juge, N. & Angulo, J. Differential Epitope Mapping by STD NMR Spectroscopy To Reveal the Nature of Protein–Ligand Contacts. *Angewandte Chemie International Edition* **56**, 15289–15293 (2017).
169. Smith, D. A., Allerton, C., Kalgutkar, A. S., van de Waterbeemd, H. & Walker, D. K. in *Pharmacokinetics and Metabolism in Drug Design* 1–17 (John Wiley & Sons, Ltd, 2012).

170. Kuntz, I. D., Chen, K., Sharp, K. A. & Kollman, P. A. The maximal affinity of ligands. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 9997–10002 (1999).
171. Keserü, G. M. & Makara, G. M. The influence of lead discovery strategies on the properties of drug candidates. *Nature Reviews Drug Discovery* **8**, 203–212 (2009).
172. Mortenson, P. N. & Murray, C. W. Assessing the lipophilicity of fragments and early hits. *Journal of Computer-Aided Molecular Design* **25**, 663–667 (2011).
173. Nissink, J. W. M. Simple size-independent measure of ligand efficiency. *Journal of Chemical Information and Modeling* **49**, 1617–1622 (2009).
174. Leeson, P. D. & Springthorpe, B. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nature Reviews Drug Discovery* **6**, 881–890 (2007).
175. Ferenczy, G. G. & Keserü, G. M. in *Fragment-based Drug Discovery Lessons and Outlook* 75–98 (John Wiley & Sons, Ltd, 2016).
176. Cherry, M. & Mitchell, T. in *Fragment-Based Drug Discovery* 1–13 (John Wiley & Sons, Ltd, 2008).
177. Dennis, S., Kortvelyesi, T. & Vajda, S. Computational mapping identifies the binding sites of organic solvents on proteins. *Proceedings of the National Academy of Sciences* **99**, 4290–4295 (2002).
178. Kozakov, D., Grove, L. E., Hall, D. R., Bohnuud, T., Mottarella, S. E., Luo, L., Xia, B., Beglov, D. & Vajda, S. The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. *Nature Protocols* **10**, 733–755 (2015).
179. Brenke, R., Kozakov, D., Chuang, G.-Y., Beglov, D., Hall, D., Landon, M. R., Mattos, C. & Vajda, S. Fragment-based identification of druggable ‘hot spots’ of proteins using Fourier domain correlation techniques. *Bioinformatics* **25**, 621–627 (Jan. 2009).
180. Guvench, O. & MacKerell, A. D. Computational fragment-based binding site identification by ligand competitive saturation. *PLoS Computational Biology* **5** (2009).
181. Alvarez-Garcia, D. & Barril, X. Molecular simulations with solvent competition quantify water displaceability and provide accurate interaction maps of protein binding sites. *Journal of Medicinal Chemistry* **57**, 8530–8539 (2014).

182. Ung, P. M. U., Ghanakota, P., Graham, S. E., Lexa, K. W. & Carlson, H. A. Identifying binding hot spots on protein surfaces by mixed-solvent molecular dynamics: HIV-1 protease as a test case. *Biopolymers* **105**, 21–34 (2016).
183. Ghanakota, P. & Carlson, H. A. Driving Structure-Based Drug Discovery through Cosolvent Molecular Dynamics. *Journal of Medicinal Chemistry* **59**, 10383–10399 (2016).
184. Defelipe, L. A., Arcon, J. P., Modenutti, C. P., Marti, M. A., Turjanski, A. G. & Barril, X. Solvents to fragments to drugs: MD applications in drug design. *Molecules* **23**, 1–14 (2018).
185. Faller, C. E., Raman, E. P., MacKerell, A. D. & Guvench, O. in *Fragment-Based Methods in Drug Discovery* (Hrsg. Klon, A. E.) 75–87 (Springer New York, New York, NY, 2015).
186. Shuker, S. B., Hajduk, P. J., Meadows, R. P. & Fesik, S. W. Discovering High-Affinity Ligands for Proteins: SAR by NMR. *Science* **274**, 1531–1534 (1996).
187. Claridge, T. D. in *High-Resolution NMR Techniques in Organic Chemistry (Third Edition)* (Hrsg. Claridge, T. D.) Third Edition, 11–59 (Elsevier, Boston, 2016).
188. Claridge, T. D. in *High-Resolution NMR Techniques in Organic Chemistry (Third Edition)* (Hrsg. Claridge, T. D.) Third Edition, 315–380 (Elsevier, Boston, 2016).
189. Eykyn, T. R., Früh, D. & Bodenhausen, G. Identification of Spin Diffusion Pathways in Isotopically Labeled Biomolecules. *Journal of Magnetic Resonance* **138**, 330–333 (1999).
190. Mayer, M. & Meyer, B. Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angewandte Chemie - International Edition* **38**, 1784–1788 (1999).
191. Claridge, T. D. in *High-Resolution NMR Techniques in Organic Chemistry (Third Edition)* (Hrsg. Claridge, T. D.) Third Edition, 421–455 (Elsevier, Boston, 2016).
192. Walpole, S., Monaco, S., Nepravishta, R. & Angulo, J. *STD NMR as a Technique for Ligand Screening and Structural Studies* 1. Aufl., 423–451 (Elsevier Inc., 2019).
193. Maveyraud, L. & Mourey, L. Protein X-ray crystallography and drug discovery. *Molecules* **25** (2020).
194. Rupp, B. *Biomolecular Crystallography. Principles, Practice, and Application to Structural Biology* (Garland Science, 2010).

195. Sherwood, D. & Cooper, J. *Crystals, X-rays and Proteins. Comprehensive Protein Crystallography* (Oxford University Press, 2011).
196. Ameh, E. S. A review of basic crystallography and x-ray diffraction applications. *International Journal of Advanced Manufacturing Technology* **105**, 3289–3302 (2019).
197. Massa, W. in *Kristallstrukturbestimmung* 129–148 (Vieweg+Teubner, Wiesbaden, 2009).
198. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallographica Section D* **69**, 1204–1214 (2013).
199. Kaminski, T., Gunnarsson, A. & Geschwindner, S. Harnessing the Versatility of Optical Biosensors for Target-Based Small-Molecule Drug Discovery. *ACS Sensors* **2**, 10–15 (2017).
200. Kumar, A., Voet, A. & Zhang, K. Fragment Based Drug Design: From Experimental to Computational Approaches. *Current Medicinal Chemistry* **19**, 5128–5147 (2012).
201. Petersen, R. L. Strategies using bio-layer interferometry biosensor technology for vaccine research and development. *Biosensors* **7** (2017).
202. Wartchow, C. A., Podlaski, F., Li, S., Rowan, K., Zhang, X., Mark, D. & Huang, K.-S. Biosensor-based small molecule fragment screening with bio-layer interferometry. *Journal of Computer-Aided Molecular Design* **25**, 669–676 (2011).
203. Tab, H., Tan, Y., Witte, K. L., Zuk, R., Carricato, G. & Lockard, S. *US Patent* 7319525B2 (2008).
204. Nolte, D. D. in *Optical Interferometry for Biology and Medicine* 3–48 (Springer New York, New York, NY, 2012).
205. Hariharan, P. in *Optical Interferometry (Second Edition)* (Hrsg. Hariharan, P.) Second Edition, 9–34 (Academic Press, San Diego, 2003).
206. Schmitt, H. M., Brecht, A., Piehler, J. & Gauglitz, G. An integrated system for optical biomolecular interaction analysis. *Biosensors and Bioelectronics* **12**, 809–816 (1997).
207. Gauglitz, G., Brecht, A., Kraus, G. & Mahm, W. Chemical and biochemical sensors based on interferometry at thin (multi-) layers. *Sensors and Actuators: B. Chemical* **11**, 21–27 (1993).
208. Sartorius. *Octet Label-free Detection Systems* <https://www.sartorius.com/en/products/protein-analysis/octet-label-free-detection-systems> (2020).

209. Hall, D. in *Handbook of Surface Plasmon Resonance* 81–122 (The Royal Society of Chemistry, 2008).
210. Rich, R. L. & Myszka, D. G. Extracting kinetic rate constants from binding responses. *Label-Free Biosensors: Techniques and Applications*, 85–109 (2009).
211. Schasfoort, Richard B.M.; McWhirter, A. *SPR Instruments* 35–78 (2008).
212. Rich, R. L., Day, Y. S., Morton, T. A. & Myszka, D. G. High-Resolution and High-Throughput Protocols for Measuring Drug/Human Serum Albumin Interactions Using BIACORE. *Analytical Biochemistry* **296**, 197–207 (2001).
213. De Mol, N. J. & Fischer, M. J. in *Handbook of Surface Plasmon Resonance* 123–172 (The Royal Society of Chemistry, 2008).
214. Rich, R. L. & Myszka, D. G. in *Label-Free Biosensors: Techniques and Applications* (Hrsg. Cooper, M. A.) 48–84 (Cambridge University Press, 2009).
215. Schuck, P. & Minton, A. P. Analysis of mass transport-limited binding kinetics in evanescent wave biosensors. *Analytical Biochemistry* **240**, 262–272 (1996).
216. Case, D., Belfon, K., Ben-Shalom, I., Brozell, S., Cerutti, D., Cheatham, T., III, Cruzeiro, V., Darden, T., Duke, R., Giambasu, G., Gilson, M., Gohlke, H., Goetz, A., Harris, R., Izadi, S., Izmailov, S., Kasavajhala, K., Kovalenko, A., Krasny, R., Kurtzman, T., Lee, T., LeGrand, S., Li, P., Lin, C., Liu, J., Luchko, T., Luo, R., Man, V., Merz, K., Miao, Y., Mikhailovskii, O., Monard, G., Nguyen, H., Onufriev, A., Pan, F., Pantano, S., Qi, R., Roe, D., Roitberg, A., Sagui, C., Schott-Verdugo, S., Shen, J., Simmerling, C., N.R.Skrynnikov, Smith, J., Swails, J., Walker, R., Wang, J., Wilson, L., Wolf, R., Wu, X., Xiong, Y., Xue, Y., York, D. & Kollman, P. *AMBER 2018* (University of California, San Francisco, 2018).
217. Thorn, A. & Sheldrick, G. M. ANODE: anomalous and heavy-atom density calculation. *Journal of Applied Crystallography* **44**, 1285–1287 (2011).
218. Bricogne G., Blanc E., Brandl M., Flensburg C., Keller P., Paciorek W., Roversi P, Sharff A., Smart O.S., Vornrhein C. & Womack T.O. *BUSTER Version* (Cambridge, United Kingdom: Global Phasing Ltd.).
219. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D* **67**, 235–242 (2011).

220. Emsley, P. & Cowtan, K. *Coot*: model-building tools for molecular graphics. *Acta Crystallographica Section D* **60**, 2126–2132 (2004).
221. Sartorius. *ForteBio Data Analysis*
222. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox. *Gaussian 09, Revision C.01* (Gaussian, Inc., Wallingford CT, 2016, 2016).
223. Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B. & Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19–25 (2015).
224. ChemAxon. *Marvin was used for drawing, displaying and characterizing chemical structures, substructures and reactions, Marvin 17.28.0, 2017*
225. Microsoft Corporation. *Microsoft Excel* Version 2019 (16.0). 24. Sep. 2018.
226. Milletti, F., Storchi, L., Sforna, G. & Cruciani, G. New and Original pKa Prediction Method Using Grid Molecular Interaction Fields. *Journal of Chemical Information and Modeling* **47**. PMID: 17910431, 2172–2181 (2007).
227. Milletti, F., Storchi, L., Sforna, G., Cross, S. & Cruciani, G. Tautomer Enumeration and Stability Prediction for Virtual Screening on Large Chemical Databases. *Journal of Chemical Information and Modeling* **49**. PMID: 19123923, 68–75 (2009).
228. Milletti, F. & Vulpetti, A. Tautomer Preference in PDB Complexes and its Impact on Structure-Based Drug Discovery. *Journal of Chemical Information and Modeling* **50**, 1062–1074 (2010).
229. *Origin(Pro)* Version 2019b (OriginLab Corporation, Northampton, MA, USA., 2020).

230. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. *PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallographica Section D* **66**, 213–221 (2010).
231. Schrödinger, LLC. *The PyMOL Molecular Graphics System, Version 2.3* 2020.
232. R Core Team. *R: A Language and Environment for Statistical Computing* R Foundation for Statistical Computing (Vienna, Austria, 2020).
233. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag New York, 2016).
234. Open Source. *RDKit: Open-source cheminformatics* Version 2018.09.1.
235. Sheldrick, G. M. A short history of *SHELX. Acta Crystallographica Section A* **64**, 112–122 (2008).
236. *TopSpin* Version 4.0.8 (Bruker Corporation, 40 Manning Road, Billerica, MA 01821, USA, 2019).
237. Humphrey, W., Dalke, A. & Schulten, K. VMD – Visual Molecular Dynamics. *Journal of Molecular Graphics* **14**, 33–38 (1996).
238. Kabsch, W. *XDS. Acta Crystallographica Section D* **66**, 125–132 (2010).
239. *Handbuch-Abschnitt: sdfilter* 2019.01 (Chemical Computing Group ULC, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2021).
240. Halgren, T. A. MMFF VI. MMFF94s option for energy minimization studies. *Journal of Computational Chemistry* **20**, 720–729 (1999).
241. Jinuraj, K. R., Rakhila, M., Dhanalakshmi, M., Sajeev, R., Akshata, G., Jayan, K., Muhammed Iqbal, P., Manuel, A. T. & Abdul Jaleel, U. C. Feature optimization in high dimensional chemical space: Statistical and data mining solutions. *BMC Research Notes* **11**, 1–7 (2018).
242. Akella, L. B. & DeCaprio, D. Cheminformatics approaches to analyze diversity in compound screening libraries. *Current Opinion in Chemical Biology* **14**. Molecular Diversity, 325–330 (2010).
243. Reymond, J. L. The Chemical Space Project. *Accounts of Chemical Research* **48**, 722–730 (2015).

244. Ertl, P., Rohde, B. & Selzer, P. Fast Calculation of Molecular Polar Surface Area as a Sum of Fragment-Based Contributions and Its Application to the Prediction of Drug Transport Properties. *Journal of Medicinal Chemistry* **43**, 3714–3717 (2000).
245. Le, L. T. M., Kang, W., Kim, J.-Y., Le, O. T. T., Lee, S. Y. & Yang, J. K. Structural Details of Ufd1 Binding to p97 and Their Functional Implications in ER-Associated Degradation. *PLOS ONE* **11**, 1–15 (Sep. 2016).
246. Lim, J. J., Lee, Y., Yoon, S. Y., Ly, T. T., Kang, J. Y., Youn, H. S., An, J. Y., Lee, J. G., Park, K. R., Kim, T. G., Yang, J. K., Jun, Y. & Eom, S. H. Structural insights into the interaction of human p97 N-terminal domain and SHP motif in Derlin-1 rhomboid pseudoprotease. *FEBS Letters* **590**, 4402–4413 (2016).
247. *Handbuch-Abschnitt: Conformation Import* 2019.01 (Chemical Computing Group ULC, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2021).
248. *Handbuch-Abschnitt: Protein Ligand Interaction Fingerprints* 2019.01 (Chemical Computing Group ULC, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2021).
249. Feig, M., Onufriev, A., Lee, M. S., Im, W., Case, D. A. & Brooks III, C. L. Performance comparison of generalized born and Poisson methods in the calculation of electrostatic solvation energies for protein structures. *Journal of Computational Chemistry* **25**, 265–284 (2004).
250. Nguyen, C. N., Kurtzman Young, T. & Gilson, M. K. Grid inhomogeneous solvation theory: Hydration structure and thermodynamics of the miniature receptor cucurbit[7]uril. *Journal of Chemical Physics* **137**, 973–980 (2012).
251. Ramsey, S., Nguyen, C., Salomon-Ferrer, R., Walker, R. C., Gilson, M. K. & Kurtzman, T. Solvation thermodynamic mapping of molecular surfaces in ambertools: GIST. *Journal of Computational Chemistry* **37**, 2029–2037 (2016).
252. Martinez-Rosell, G., Harvey, M. J. & De Fabritiis, G. Molecular-Simulation-Driven Fragment Screening for the Discovery of New CXCL12 Inhibitors. *Journal of Chemical Information and Modeling* **58**, 683–691 (2018).
253. Jorgensen, W. L. & Tirado-Rives, J. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proceedings of the National Academy of Sciences* **102**, 6665–6670 (2005).

254. Dodda, L. S., Vilseck, J. Z., Tirado-Rives, J. & Jorgensen, W. L. 1.14*CM1A-LBCC: Localized Bond-Charge Corrected CM1A Charges for Condensed-Phase Simulations. *The Journal of Physical Chemistry B* **121**. PMID: 28224794, 3864–3870 (2017).
255. Dodda, L. S., Cabeza de Vaca, I., Tirado-Rives, J. & Jorgensen, W. L. LigPar-Gen web server: an automatic OPLS-AA parameter generator for organic ligands. *Nucleic Acids Research* **45**, W331–W336 (Apr. 2017).
256. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* **52**, 7182–7190 (1981).
257. Kaunitz, J. D. The Discovery of PCR : ProCuRement of Divine Power. *Digestive Diseases and Sciences* **60**, 2230–2231 (2015).
258. Van Den Ent, F. & Löwe, J. RF cloning: A restriction-free method for inserting target genes into plasmids. *Journal of Biochemical and Biophysical Methods* **67**, 67–74 (2006).
259. Fairhead, M. & Howarth, M. in *Site-Specific Protein Labeling: Methods and Protocols* (Hrsg. Gautier, A. & Hinner, M. J.) 171–184 (Springer New York, New York, NY, 2015).
260. Hänzelmann, P., Buchberger, A. & Schindelin, H. Hierarchical binding of cofactors to the AAA ATPase p97. *Structure* **19**, 833–843 (2011).
261. Tang, W. K. & Xia, D. Role of the D1-D2 Linker of Human VCP/p97 in the Asymmetry and ATPase Activity of the D1-domain. *Scientific Reports* **6**, 20037 (2016).
262. Laemmli, U. K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* **227**, 680–685 (1970).
263. Studier, F. & Moffatt, B. A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology* **189**, 113–130 (1986).
264. Dubendorf, J. W. & Studier, F. Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor. *Journal of Molecular Biology* **219**, 45–59 (1991).
265. Porath, J., Carlsson, J., Olsson, I. & Belfrage, G. Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* **258**, 598–599 (1975).
266. Porath, J. & Flodin, P. Gel filtration: A method for desalting and group separation. *Nature* **183**, 1657–1659 (1959).

267. *Purification of His-tag proteins, User manual Rev.08* (Macherey-Nagel GmbH & Co. KG, Neumann-Neander-Str. 6–8, 52355 Düren, Germany, 2019).
268. Rupp, B. in *Biomolecular Crystallography. Principles, Practice, and Application to Structural Biology* 78–91 (Garland Science, 2010).
269. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. Phaser crystallographic software. *Journal of Applied Crystallography* **40**, 658–674 (2007).
270. Painter, J. & Merritt, E. A. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallographica Section D* **62**, 439–450 (2006).
271. Myszka, D. G. Improving biosensor analysis. *Journal of Molecular Recognition* **12**, 279–284 (1999).
272. ForteBio. Super Streptavidin (SSA) Biosensors. *ForteBio Data Sheet*, 1–2 (2019).
273. Martin, E., Wang, J., Zaror, I., Yu, J., Yan, K., Doyle, M., Feucht, P., Shoemaker, K., Warne, B., Chin, M., Sy, B., Leder, L., Meyerhofer, M., Wartchow, C. & Yao, D. in *Label-Free Technologies for Drug Discovery* 223–240 (John Wiley & Sons, Ltd, 2011).
274. ForteBio. Small Molecule Binding Kinetics. *ForteBio Technical Note* 16 (2010).
275. *A Curve* <https://www.sprpages.nl/sensorgram-tutorial/a-curve/#ref4> (2019).
276. Zhang, J.-H., Chung, T. D. Y. & Oldenburg, K. R. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *Journal of Biomolecular Screening* **4**. PMID: 10838414, 67–73 (1999).
277. Giannetti, A. M., Gilbert, H. N., Huddler, D. P., Reiter, M., Strande, C., Pitts, K. E. & Bravo, B. J. in *Fragment-Based Drug Discovery* 19–48 (The Royal Society of Chemistry, 2015).
278. Viegas, A., Manso, J., Nobrega, F. L. & Cabrita, E. J. Saturation-transfer difference (STD) NMR: A simple and fast method for ligand screening and characterization of protein binding. *Journal of Chemical Education* **88**, 990–994 (2011).
279. Lea, W. A. & Simeonov, A. Fluorescence polarization assays in small molecule screening. *Expert Opinion on Drug Discovery* **6**. PMID: 22328899, 17–32 (2011).
280. Böhler, S. Strukturbasierte Entwicklung kovalenter p97-Inhibitoren. *Masterarbeit* (2020).

281. Hänzelmann, P. & Schindelin, H. Characterization of an Additional Binding Surface on the p97 N-Terminal Domain Involved in Bipartite Cofactor Interactions. *Structure* **24**, 140–147 (2016).
282. Vagenende, V. & Trout, B. L. Quantitative characterization of local protein solvation to predict solvent effects on protein structure. *Biophysical Journal* **103**, 1354–1362 (2012).
283. Giordanetto, F., Jin, C., Willmore, L., Feher, M. & Shaw, D. E. Fragment Hits: What do They Look Like and How do They Bind? *Journal of Medicinal Chemistry* **62**, 3381–3394 (2019).
284. Sperandio, O., Miteva, M. A. & Villoutreix, B. O. in *Virtual Screening* 435–465 (John Wiley & Sons, Ltd, 2011).
285. Janin, J. in *Computational Protein-Protein Interactions* 1–20 (CRC Press, 2009).
286. Feuston, B., Holloway, M. K., McGaughey, G. & Culberson, J. C. in *Fragment-Based Drug Discovery* 223–244 (John Wiley & Sons, Ltd, 2008).
287. Jacquemard, C., Drwal, M. N., Desaphy, J. & Kellenberger, E. Binding mode information improves fragment docking. *Journal of Cheminformatics*, 1–15 (2019).
288. Huang, D. & Caflisch, A. in *Virtual Screening* 467–489 (John Wiley & Sons, Ltd, 2011).
289. Kawatkar, S., Wang, H., Czerminski, R. & Joseph-McCarthy, D. Virtual fragment screening: an exploration of various docking and scoring protocols for fragments using Glide. *Journal of computer-aided molecular design* **23**, 527–539 (2009).
290. Amaro, R. E., Baudry, J., Chodera, J., Demir, Ö., McCammon, J. A., Miao, Y. & Smith, J. C. Ensemble Docking in Drug Discovery. *Biophysical Journal* **114**, 2271–2278 (2018).
291. Chou, T.-F., Bulfer, S. L., Weihl, C. C., Li, K., Lis, L. G., Walters, M. A., Schoenen, F. J., Lin, H. J., Deshaies, R. J. & Arkin, M. R. Specific Inhibition of p97/VCP ATPase and Kinetic Analysis Demonstrate Interaction between D1 and D2 ATPase domains. *J Mol Biol* **426**, 2886–2899 (2014).
292. Rich, R. L. & Myszka, D. G. in *Label-Free Technologies for Drug Discovery* 1–25 (John Wiley & Sons, Ltd, 2011).
293. Stenberg, E., Persson, B., Roos, H. & Urbaniczky, C. Quantitative determination of surface concentration of protein with surface plasmon resonance using radiolabeled proteins. *Journal of Colloid and Interface Science* **143**, 513–526 (1991).

294. Shah, N. Bio-layer interferometry for measuring kinetics of protein-protein interactions and allosteric ligand effects. *Journal of Visualized Experiments*, 1–7 (2014).
295. Ewens, C. A., Panico, S., Kloppsteck, P., McKeown, C., Ebong, I. O., Robinson, C., Zhang, X. & Freemont, P. S. The p97-faf1 protein complex reveals a common mode of p97 adaptor binding. *Journal of Biological Chemistry* **289**, 12077–12084 (2014).
296. Vogel, S. Synthese prenylierter Chalkone aus Hopfen und Bestimmung ihrer cytotoxischen und antioxidativen Aktivität. *Doktorarbeit* (2008).
297. Silverman, R. B. & Holladay, M. W. *The organic chemistry of drug design and drug action* (Academic Press, 20014).
298. Khupse, R. S. & Erhardt, P. W. Total synthesis of xanthohumol. *Journal of Natural Products* **70**, 1507–1509 (2007).
299. Rich, R. L. & Myszka, D. G. Survey of the year 2007 commercial optical biosensor literature. *Journal of Molecular Recognition* **21**, 355–400 (2008).
300. Marquart, J. A. in *Handbook of Surface Plasmon Resonance* (2) 106–148 (The Royal Society of Chemistry, 2017).
301. Cunningham, B. T. in *Label-Free Biosensors: Techniques and Applications* (Hrsg. Cooper, M. A.) 1–28 (Cambridge University Press, 2009).
302. Schiebel, J., Radeva, N., Krimmer, S. G., Wang, X., Stieler, M., Ehrmann, F. R., Fu, K., Metz, A., Huschmann, F. U., Weiss, M. S., Mueller, U., Heine, A. & Klebe, G. Six Biophysical Screening Methods Miss a Large Proportion of Crystallographically Discovered Fragment Hits: A Case Study. *ACS Chemical Biology* **11**, 1693–1701 (2016).
303. Banfi, D. & Patiny, L. www.nmrdb.org: Resurrecting and Processing NMR Spectra On-line, Chima Zürich. *Chima* **62**, 280–281 (2008).
304. Castillo, A. M., Patiny, L. & Wist, J. Fast and accurate algorithm for the simulation of NMR spectra of large spin systems. *Journal of Magnetic Resonance* **209**, 123–130 (2011).
305. Aires-de Sousa, J., Hemmer, M. C. & Gasteiger, J. Prediction of ¹H NMR Chemical Shifts Using Neural Networks. *Analytical Chemistry* **74**. PMID: 11795822, 80–90 (2002).
306. Guvench, O. & MacKerell Jr, A. D. Computational Fragment-Based Binding Site Identification by Ligand Competitive Saturation. *PLOS Computational Biology* **5**, 1–10 (Juli 2009).

307. Kumar, S. & Nussinov, R. Close-range electrostatic interactions in proteins. *ChemBioChem* **3**, 604–617 (2002).
308. Herschlag, D. & Pinney, M. M. Hydrogen Bonds: Simple after All? *Biochemistry* **57**, 3338–3352 (2018).
309. Pearce, N. M., Krojer, T., Bradley, A. R., Collins, P., Nowak, R. P., Talon, R., Marsden, B. D., Kelm, S., Shi, J., Deane, C. M. & Von Delft, F. A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density. *Nature Communications* **8**, 24–29 (2017).
310. Bemis, G. W. & Murcko, M. A. The properties of known drugs. 1. Molecular frameworks. *Journal of Medicinal Chemistry* **39**, 2887–2893 (1996).
311. Ferenczy, G. G. & Keseru, G. M. How are fragments optimized? A retrospective analysis of 145 fragment optimizations. *Journal of Medicinal Chemistry* **56**, 2478–2486 (2013).
312. Young, R. J. & Leeson, P. D. Mapping the Efficiency and Physicochemical Trajectories of Successful Optimizations. *Journal of Medicinal Chemistry* **61**, 6421–6467 (2018).
313. Olsson, T. S., Williams, M. A., Pitt, W. R. & Ladbury, J. E. The Thermodynamics of Protein-Ligand Interaction and Solvation: Insights for Ligand Design. *Journal of Molecular Biology* **384**, 1002–1017 (2008).