

Molekülmechanische Untersuchungen zur Struktur, Funktion
und Dynamik der Hepatitis C viralen RNA-abhängigen RNA-
Polymerase

Dissertation

zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)

vorgelegt der
Naturwissenschaftlichen Fakultät I - Biowissenschaften
der
Martin-Luther-Universität Halle-Wittenberg

von
Sarah Ludwig (geb. Schäfer)
geboren am 25.04.1986 in Staßfurt

Gutachter:

PD Dr. Iris Thondorf

PD Dr. Ralph Golbik

Prof. Dr. Dietmar Schomburg

Datum der öffentlichen Verteidigung: 24.10.2018

INHALTSVERZEICHNIS

1.	Einleitung	1
1.1.	Das Hepatitis C Virus	1
1.2.	Aufbau des Virus	1
1.3.	Aufbau des Genoms	2
1.4.	Der Lebenszyklus von HCV	4
1.5.	Behandlungsmethoden gegen HCV	7
1.6.	Das Nicht-Struktur Protein 5B - NS5B	8
1.7.	Die Nicht-Nukleosid-Inhibitoren von NS5B	11
1.8.	Motivation und Zielstellung	12
2.	Grundlagen	14
2.1.	Superpositionierung von Proteinstrukturen	14
2.2.	Molekülmechanik	14
2.2.1.	Kraftfelder	15
2.2.2.	Energieminimierung	16
2.3.	Moleküldynamik	17
2.3.1.	Integrationsschemata der klassischen Bewegungsgleichung	18
2.3.2.	Beschränkung von Freiheitsgraden	21
2.3.3.	Ensembles	21
2.3.4.	Lösungsmittel	24
2.3.5.	Periodische Randbedingungen und <i>cut-offs</i>	26
2.4.	<i>Root mean square deviation</i> (RMSD)	26
2.5.	Hauptkomponentenanalyse (PCA)	27
2.6.	Clusteranalyse	28
2.6.1.	Bestimmung der optimalen Clusternummer	28
2.7.	<i>Crosslinking</i>	29
2.7.1.	Vorgehensweise und Identifizierung von <i>Crosslinks</i>	30

2.7.2.	<i>Crosslink</i> -Reagenzien	30
2.8.	Massenspektrometrie	31
2.8.1.	Orbitrap Fusion Tribrid Massenspektrometer.....	32
3.	Material und Methoden.....	35
3.1.	Verwendete Programme für die Computerberechnungsmethoden.....	35
3.2.	Primär- und Sekundärstrukturanalyse von NS5B.....	35
3.3.	Analyse der Kristallstrukturen von NS5B	36
3.3.1.	Superpositionierung mit THESEUS	36
3.3.2.	Das Programm R.....	37
3.4.	Moleküldynamik-Simulation.....	37
3.5.	Auswertung der Moleküldynamiken	38
3.6.	Membrananker- und Membranmodellierung	39
3.7.	<i>Crosslinking</i>	40
3.7.1.	Verwendete Chemikalien und Hersteller	40
3.7.2.	Liste der Laborgeräte und Hersteller	40
3.7.3.	Proteinaufreinigung.....	40
3.7.4.	Chemisches Vernetzen (<i>Crosslinking</i>).....	41
3.7.5.	Quantitatives <i>Crosslinking</i>	41
3.7.6.	SDS-PAGE (<i>sodium dodecyl sulfate polyacrylamide gel electrophoresis</i>).....	42
3.7.7.	Gelfärbung <i>nach Neuhoff</i>	42
3.7.8.	In-Gel Proteinspaltung	43
3.7.9.	Orbitrap Fusion Tribrid Massenspektrometer.....	43
3.5.10	Identifizierung der <i>Crosslink</i> -Produkte.....	44
4.	Ergebnisse und Diskussion	45
4.1.	Analyse der Kristallstrukturen.....	45
4.1.1.	Primär- und Sekundärstrukturanalyse der NS5B-Kristallstrukturen	45
4.1.2.	Qualitätsbewertung der NS5B Kristallstrukturen	48
4.1.3.	RMSD- und Hauptkomponentenanalyse der Kristallstrukturen	50
4.1.4.	THESEUS-Analyse der Genotypen	54

4.2.	Moleküldynamik-Simulationen von NS5B	56
4.2.1.	MD-Simulation von NS5B mit Substraten	57
4.2.2.	MD-Simulation einer Doppelstrang-RNA mit NS5B in geöffneter Konformation. 65	
4.3.	Molekülmechanische Untersuchungen der Interaktionen von NS5B mit Nicht-Nukleotid-Inhibitoren	69
4.3.1.	MD-Simulationen mit dem Inhibitor HCV-796	70
4.3.2.	MD-Simulationen mit einem NNI-2-Inhibitor.....	74
4.3.3.	MD-Simulationen mit einem NNI-3-Inhibitor.....	82
4.3.4.	MD-Simulationen mit einem NNI-5-Inhibitor.....	88
4.4.	Vergleich der MD-Simulationen	93
4.4.1.	RMSD-Matrix	93
4.4.2.	Hauptkomponentenanalyse	94
4.5.	Modellierung des Membranankers	96
4.5.1.	Sequenzalignment des C-Terminus	97
4.5.2.	Modellierung des Membranankers	98
4.5.3.	Vergleich mit der NMR-Struktur des Membranankers von NS5B.....	99
4.5.4.	<i>Highly Mobile Membrane-Mimetic (HMMM) -Builder</i>	99
4.5.5.	Die Position des Linkers in Anwesenheit des Membranankers.....	100
4.6.	<i>Crosslink</i> -Experimente	102
5.	Zusammenfassung und Ausblick	115
	Literaturverzeichnis	IV
	Abbildungsverzeichnis	XIV
	Tabellenverzeichnis.....	XVI
	Abkürzungsverzeichnis	XVII
	Anhang XIX	
	Danksagung.....	LIII
	Lebenslauf	LIV
	Eidesstattliche Erklärung	LV

1. EINLEITUNG

1.1. Das Hepatitis C Virus

Das Hepatitis C Virus (HCV) verursacht bei einer chronischen Infektion schwere Leberschäden und kann zu Leberzirrhose oder Leberkrebs führen.¹ Die *World Health Organization* (WHO) schätzt, dass 130-150 Millionen Menschen weltweit an einer chronischen HCV-Infektion leiden.² Eine Impfung gegen HCV ist zurzeit noch nicht möglich.^{3,4} Ein wesentliches Problem bei der Erforschung von Impfstoffen gegen HCV sind die sieben Genotypen, die wiederum in 67 anerkannte und 20 vorläufige Subtypen eingeteilt werden.⁵⁻⁷ Hepatitis C Viren, die zum selben Genotyp gehören, besitzen eine Sequenzidentität von 65 bis 70 % und diejenigen, die zum selben Subtyp gehören von über 85 %.^{8,9} Diese sieben Genotypen zeigen eine unterschiedliche geographische Verteilung, wobei die Typen eins bis drei auf der ganzen Welt auftreten.¹⁰ Insbesondere die Subtypen 1a, 1b, 2a und 3a sind weit verbreitet und machen den Großteil aller HCV-Infektionen aus. Sie werden deshalb auch als *epidemic subtypes* bezeichnet.¹¹ Die Übertragung des Virus erfolgt über kontaminiertes Blut, wobei der Virus über infizierte Makrophagen zur Leber transportiert wird. Da das Hepatitis C Virus selbst nur schwach cytopathogen ist, werden die verursachten Zellschäden vielmehr durch die Immunreaktionen des Körpers hervorgerufen.¹²

Das Hepatitis C Virus gehört zu der Familie der *Flaviviridae*, welche in die drei Gattungen *Flavivirus*, *Pestivirus* und *Hepacivirus* unterteilt wird.¹³ Zu den *Flaviviren* gehören zum Beispiel das Gelbfieber-Virus, das West-Nil Virus oder das Dengue Virus, zu den *Pestiviren* die Schweinepest oder das Bovine-Diarrhoe Virus und zu den *Hepaciviren* gehören das GB-Virus und das Hepatitis C Virus.

1.2. Aufbau des Virus

Das HCV Partikel hat eine durchschnittliche Größe von 55 nm. Es besteht aus einer Membranhülle, auf der die Glykoproteine E1 und E2 verankert sind (siehe Abbildung 1). Weiterhin enthält es das Nukleocapsid, in dem sich die einzelsträngige (+)-RNA befindet.¹⁴

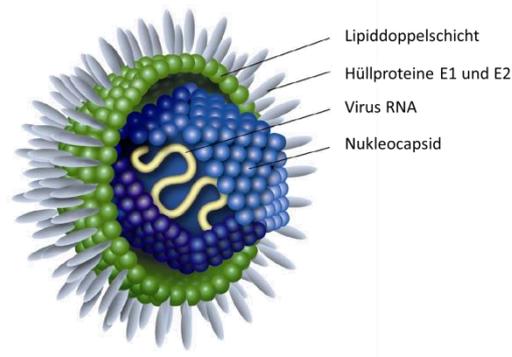


Abbildung 1: Aufbau des HCV Partikels nach Krevulová *et al.*¹⁵
 Das Virion besteht aus einer Lipiddoppelschicht mit den Hüllproteinen E1 und E2. Im Inneren befindet sich das Nukleocapsid mit der einzelsträngigen (+)-RNA.

Das C Protein, auch *Core* oder Capsid Protein genannt, ist ein hoch konserviertes und basisches Protein, welches durch Oligomerisierung das Nukleocapsid bildet. Es kann in 3 Domänen eingeteilt werden. Die N-terminale Domäne 1 (D1) enthält viele basische Aminosäuren und ist für die Bindung des Virusgenoms verantwortlich. Zudem induziert es die Oligomerisierung. Die weniger basische und eher hydrophobe Domäne 2 (D2) besitzt zwei amphipathische Helices und vermittelt die Interaktion zu den *lipid droplets*. Dies könnte den Lipidmetabolismus beeinflussen und zur *Steatosis hepatitis* (Fettleber) führen, die bei einer HCV Infektionen auftritt.¹⁶ Die sehr hydrophobe und C-terminale Domäne 3 (D3) interagiert mit dem Glykoprotein E1.

Die Hüllproteine E1 und E2 sind stark glykosyliert und vermitteln als Komplex das Eindringen des Virus in die Zelle. Die Transmembrandomänen am C-Terminus von E1 und E2 sind an der Heterodimerisierung der Glykoproteine beteiligt. Die Interaktion der Virus-Hüllproteine mit zellulären Rezeptorproteinen (z.B. CD81, LDLR, SR-BI, SR-BII oder CLDN1) veranlasst die Clathrin vermittelte Endocytose.¹⁷

1.3. Aufbau des Genoms

Die 9,6 kb lange (+)-RNA enthält einen offenen Leserahmen (ORF, *open reading frame*), welcher von zwei nicht-translatierten Regionen (NTRs) an dem 5' und dem 3' Ende flankiert wird. Diese NTRs zeigen ein hohes Maß an Strukturiertheit, wie in Abbildung 2 gut zu erkennen ist. Die *internal ribosome entry site* (IRES) befindet sich in der 5'-NTR und initiiert mit dem AUG-Startcodon die Translation des Virusgenoms in den Ribosomen. Die 5'-NTR ist zwischen den HCV Genotypen hoch konserviert und kann in 6 Strukturdomänen eingeteilt

werden, wobei die Domänen II und III essentiell für die RNA Replikation sind. In der 3'-NTR befindet sich das UGA-Stoppocodon, welches die Translation des Polyproteins abbricht. Weiterhin besteht es aus einem poly(U)-Abschnitt mit einer durchschnittlichen Länge von 80 Nukleotiden und einem konservierten Bereich aus 98 Nukleotiden.¹⁸ Das HCV-Genom besitzt zudem im Bereich des Hüllproteins E2 eine hypervariable Region, genannt HVR1.¹⁹

Das im Ribosom translatierte Polyprotein besitzt eine Länge von ca. 3000 Aminosäuren und wird anschließend durch wirtseigene und virale Proteasen (NS2 und NS3) in 10 Proteine prozessiert. Diese beinhalten 3 Strukturproteine (*Core*, Glykoprotein 1 (E1) und Glykoprotein 2 (E2)), welche für die Verpackung des Genoms und den Aufbau des Virus essentiell sind, sowie 7 Nicht-Strukturproteine (p7, NS2, NS3, NS4A, NS4B, NS5A und NS5B), welche für die Replikation benötigt werden (vgl. Abbildung 2).²⁰

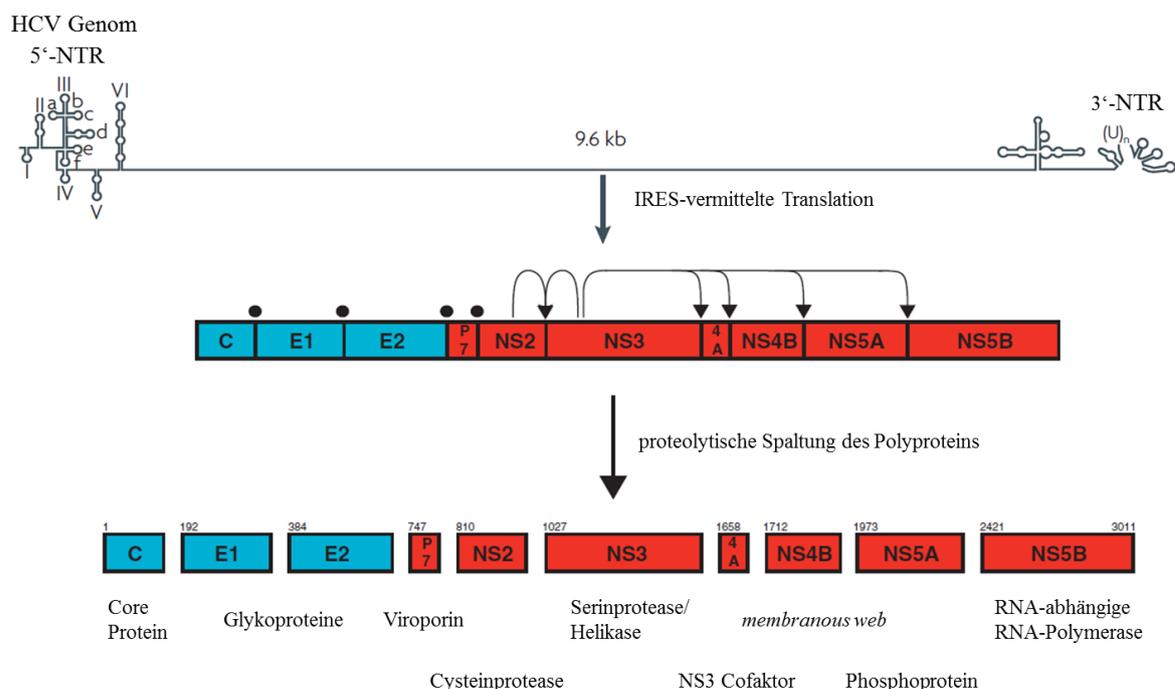


Abbildung 2: Das HCV Genom. Veränderte Darstellung nach Moradpour *et al.*¹⁸ und Weiser *et al.*²⁰ Gezeigt ist der Aufbau der (+)-RNA des translatierten Polyproteins und die Prozessierung durch zelleigene und HCV Proteasen. Die IRES befindet sich in der 5'-NTR. Die Strukturproteine sind blau unterlegt und die Nichtstrukturproteine rot. Die schwarzen Kreise markieren die Signalpeptide der Proteasespaltung. Die schwarzen Pfeile kennzeichnen die Proteaseaktivität der Nicht-Strukturproteine. Die Aminosäurenummern der HCV Proteine entsprechen dem Genotyp Con 1b.

1.4. Der Lebenszyklus von HCV

Der erste Schritt des Lebenszyklus von HCV ist die Interaktion der Virus-Hüllproteine E1 und E2 mit den zellulären Rezeptorproteinen und wird allgemein Adsorption genannt. Nach der Endocytose von HCV entwickeln sich durch das Protein Rab5 frühe Endosomen, deren pH-Wert durch Protonenpumpen reduziert wird.²¹ Dies führt zu einer Fusion der viralen und der endosomalen Membran. Die anschließende Freisetzung des viralen Genoms aus dem Capsid ins Zytoplasma ist jedoch noch nicht genau geklärt und wird allgemein als *Uncoating* bezeichnet. Es ist bekannt, dass das HCV Protein p7 essentiell für eine erfolgreiche Infektion *in vivo* ist. Die genaue Funktion bei diesem Prozess ist jedoch unklar.²² Das p7 ist ein kleines (7 kDa, 63 Aminosäuren) und hydrophobes Protein, das zur Familie der Viroporine gehört.^{23,24} Es ist ein Transmembranprotein und befindet sich im Endoplasmatischen Retikulum (ER). Des Weiteren bildet es Oligomere, die als Kationen-Kanäle fungieren.²⁵

Nachdem das Virusgenom im Zytosol vorliegt, erfolgt dessen Replikation. Die (+)-RNA von HCV besitzt die Polarität einer mRNA und kann direkt durch Ribosomen am ER in Proteine translatiert werden.

Wie die meisten Nicht-Strukturproteine von HCV ist auch das NS2 Protein membranassoziiert. Es besitzt eine hydrophobe Membranbindedomäne (MBD) am N-Terminus und eine Serinprotease-Domäne am C-Terminus. NS2 trennt autokatalytisch die NS2-3 Proteinverbindung in cis und setzt so die NS3 Protease frei, welche anschließend das restliche Polyprotein *downstream* prozessiert.²⁶

Das Protein NS3 ist multifunktional. So zeigt es im N-terminalen Bereich ebenfalls eine Proteaseaktivität, besitzt aber gleichzeitig auch eine NTPase/Helikase-Aktivität im C-terminalen Bereich. NS3 spielt somit eine wichtige Rolle in der RNA Replikation und ist deshalb auch ein primäres Ziel der Wirkstoffentwicklung gegen HCV.^{27,28}

Alle HCV-Proteine sind nach ihrer Prozessierung durch Transmembrandomänen, hydrophobe Membranhelices oder Protein-Protein-Interaktionen (wie im Fall von NS3) in die ER-Membran integriert oder an ihr assoziiert. NS4A wirkt hierbei als Cofaktor auf NS3 und sorgt mit seinem N-terminalen Membrananker für die Lokalisation des NS3/NS4A-Komplexes an der ER-Membran. Aber NS4A dient auch als Cofaktor bei der Phosphorylierung von NS5A und kann ebenfalls mit NS4B interagieren.^{24,29}

Für die produktive Replikation des Hepatitis C Virus bilden sich aus der ER-Membran einzelne Ausstülpungen heraus, die *membranous web* genannt werden (siehe Abbildung 3). Diese Deformierung dient als Grundgerüst für den Replikationskomplex und wird durch das

Protein NS4B hervorgerufen.³⁰ Es ist ein sehr hydrophobes und integrales Membranprotein. Die amphipathische Helix des N-Terminus scheint bei der Deformierung von besonderer Bedeutung zu sein.³¹

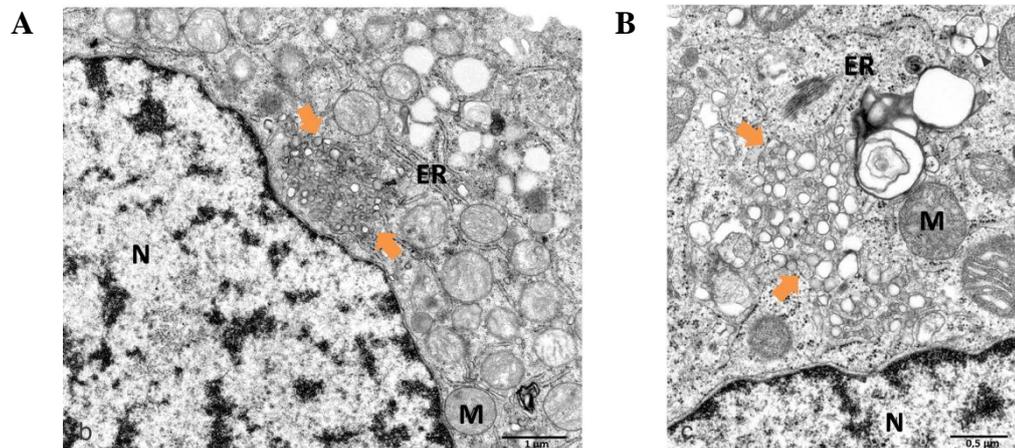


Abbildung 3: Mikroskopische Aufnahme des *membranous web* in Huh-7 Zellen, welche mit dem HCV Replikon infiziert wurden (verändert aus Gosert *et al.*³²).

Die orange gefärbten Pfeile in Abbildung **A**) und **B**) zeigen auf das *membranous web*, das aus Membranveränderungen des ER hervorgerufen wird. Auslöser ist unter anderem das NS4B Protein von HCV. N – Nukleus, M – Mitochondrium, ER – Endoplasmatisches Retikulum.

Für den Replikationskomplex werden neben NS3, NS4A und NS4B auch die HCV Proteine NS5A und NS5B benötigt. NS5A ist ein Phosphoprotein und kann im phosphorylierten oder hyperphosphorylierten Zustand auftreten.³³ NS5A wird in drei Domänen eingeteilt (Domäne I, II und III) und besitzt ein Zink-Binde-Motiv bestehend aus vier konservierten Cysteinen, deren Verlust die Virusreplikation verhindert.³⁴ NS5A besitzt keine Transmembrandomäne, aber eine amphipathische Helix am N-Terminus bzw. in der Domäne I, mit der es an die ER-Membran oder das *membranous web* assoziiert ist.³⁵ Es bindet einzelsträngige RNA mit großer Affinität, vor allem Uridin- und Guanosin-reiche Strukturen, welche sich im Bereich der 3'NTR befinden.³⁶ Obwohl die genaue Funktion von NS5A umstritten ist, ist das Protein essentiell für die RNA Replikation und den Zusammenbau des Virus.^{37,38}

Das NS5B Protein von HCV ist die RNA-abhängige RNA-Polymerase und unabdingbar für die Replikation des Genoms.³⁹ Es ist ein großes (68 kDa), hydrophiles Protein, das über die C-terminale Transmembranhelix an der ER-Membran verankert ist.⁴⁰ Neben NS3 und NS5A ist es eines der Hauptziele bei der Entwicklung von HCV-Wirkstoffen.

Die (+)-RNA des HCV wird durch den Replikationskomplex in eine (-)-RNA überführt. Diese dient wiederum als Template für die Synthese von vielen neuen (+)-RNA-Molekülen, mit deren Hilfe viele neue Virusproteine entstehen. Am Ende der Replikationsphase liegen also die viralen Strukturproteine und das Genom als vielfache Kopien in der Zelle vor. Nun erfolgt der Prozess der Virenmorphogenese. Er beschreibt den Zusammenbau der einzelnen Komponenten zu viralen Partikeln.⁴¹

Die genauen Mechanismen der Morphogenese von HCV sind noch nicht bekannt. Es konnte jedoch nachgewiesen werden, dass NS2, NS3, NS5A und die *lipid droplets* am viralen Zusammenbau beteiligt sind.^{38,42,43,44,45} Des Weiteren scheint die HCV Morphogenese eng mit dem VLDL (*very low density lipoprotein*) -Stoffwechselweg verbunden zu sein, da NS5A auch mit dem Protein ApoE interagiert. In Abbildung 4 wurde versucht, den postulierten Zusammenbau des Virus nachzuvollziehen.

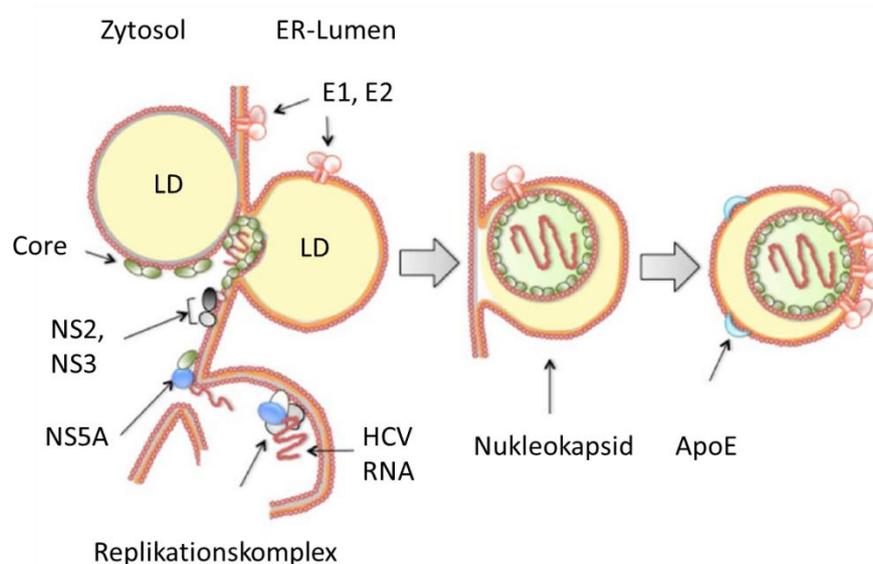


Abbildung 4: Modellvorschlag zur Morphogenese des HCV- Partikels von Suzuki *et al.*⁴¹.

Die Replikationskomplexe sind mit dem VLDL-Syntheseweg verbunden. *Core* Proteine und (+)-RNA werden an die Oberfläche der *lipid droplets* (LD) assoziiert, wobei NS5A zwischen der RNA und dem *Core* vermittelt. Dies fördert die Einkapselung und führt zur Nucleocapsidbildung. Die Einlagerung des Nucleokapsids in das Lumen des *lipid droplets* und die Angliederung der Glykoproteine E1 und E2 führt zur Entstehung des HCV-Lipoprotein-Partikels (HCVcc).

1.5. Behandlungsmethoden gegen HCV

Die Behandlung von Patienten mit HCV ist im großen Maße davon abhängig, in welchem Stadium der Infektion sich der Patient befindet, von welchem Genotyp er infiziert wurde und ob bereits eine Lebertransplantation vorgenommen wurde.⁴⁶ Die Richtlinien der *American Association for the Study of Liver Diseases* (AASLD) und der *European Association for the Study of the Liver* (EASL) geben einen sehr guten Überblick über die aktuellen Behandlungsmethoden von Hepatitis C.^{47,48}

Zu den zurzeit gebräuchlichsten Medikamenten in der HCV-Behandlung zählen NS3/4A-Protease-Inhibitoren, wie Telaprevir, Boceprevir, Simeprevir, Grazoprevir und Paritaprevir, NS5A-Inhibitoren, wie Daclatasvir, Ledipasvir, Elbasvir und Ombitasvir, und NS5B-Polymerase-Inhibitoren, wie Sofosbuvir und Dasabuvir (siehe Abbildung 5).⁴⁶ Diese werden oft in Kombination miteinander und über mehrere Wochen für eine erfolgreiche Therapie gegen HCV verabreicht. So empfiehlt die AASLD z.B. für einen Patienten mit HCV Genotyp 1b, der noch keine Anzeichen einer Leberzirrhose zeigt, folgende Behandlungsmöglichkeiten (aufgelistet nach Erfolgsrate):

- Tägliche Einnahme von Elbasvir (50 mg) in Kombination mit Grazoprevir (100 mg) für 12 Wochen,
- oder Ledipasvir (90 mg) in Kombination mit Sofosbuvir (400 mg) für 12 Wochen,
- oder Paritaprevir (150 mg) in Kombination mit Ritonavir (100 mg) und Ombitasvir (25 mg) plus zweimal täglich Dasabuvir (250 mg) für 12 Wochen,
- oder Simeprevir (150 mg) in Kombination mit Sofosbuvir (400 mg) für 12 Wochen
- oder Daclatasvir (60 mg) in Kombination mit Sofosbuvir (400 mg) für 12 Wochen

Die frühere Standardbehandlung von HCV mit pegyliertem Interferon- α und Ribavirin (Polymerase-Inhibitor, Nukleosid-Analogon) ist dagegen veraltet und wird nicht mehr empfohlen.⁴⁸ Da die Kombination mit Interferon- α (IFN) bei früheren Behandlungen zu starken Nebenwirkungen geführt hat, werden vorzugsweise IFN-freie Zusammenstellungen ausgetestet. Für die Jahre 2017 und 2018 sind weitere Untersuchungen in Kombination mit neuen Wirkstoffen wie ABT-493, MK-3682, Odalasvir, Simeprevir oder ACH-3422 geplant.⁴⁹

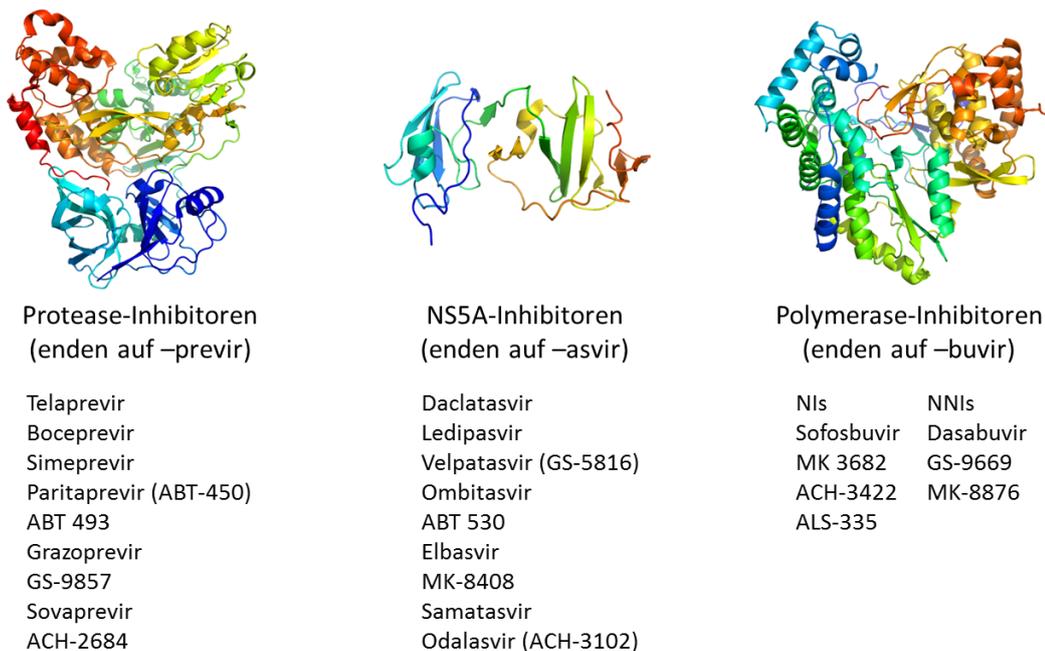


Abbildung 5: Auflistung der zurzeit verfügbaren verschiedenen Medikamente gegen HCV aus Asselah *et al.*⁴⁹.

Die Unterteilung der Wirkstoffe erfolgt anhand des Zielproteins von HCV in Protease-Inhibitoren, NS5A-Inhibitoren und Polymerase-Inhibitoren. Die Abkürzung „NIs“ steht für Nukleosid-Inhibitoren und „NNIs“ für Nicht-Nukleosid-Inhibitoren. Die Zielproteine sind als Kristallstrukturen mit den PDB-IDs 4B6F für NS3, 3FQQ für NS5A und 2XI3 für NS5B dargestellt.

1.6. Das Nicht-Struktur Protein 5B - NS5B

Das NS5B Protein ist für die Replikation des viralen Genoms verantwortlich und somit unabdingbar für die Vermehrung in der Wirtszelle.^{39,50} NS5B produziert eine komplementäre Doppelstrang-RNA mit einer 5' nach 3' Orientierung.

Wie andere RNA-Polymerasen besitzt auch NS5B eine „rechte Hand“-Topologie, die in die Finger-Domäne (Aminosäuren 1-181 und 223-283), die Handflächen- oder Hand-Domäne (Aminosäuren 182-222) und die Daumen-Domäne (Aminosäuren 359-529) eingeteilt werden kann (siehe Abbildung 6).^{51,52,53} Die restlichen Aminosäuren gehören zum Linker und zu der Membranhelix, die NS5B an der ER-Membran verankert. Eine Übersicht der einzelnen Protein-Domänen und Motive ist in Tabelle 1 aufgelistet.

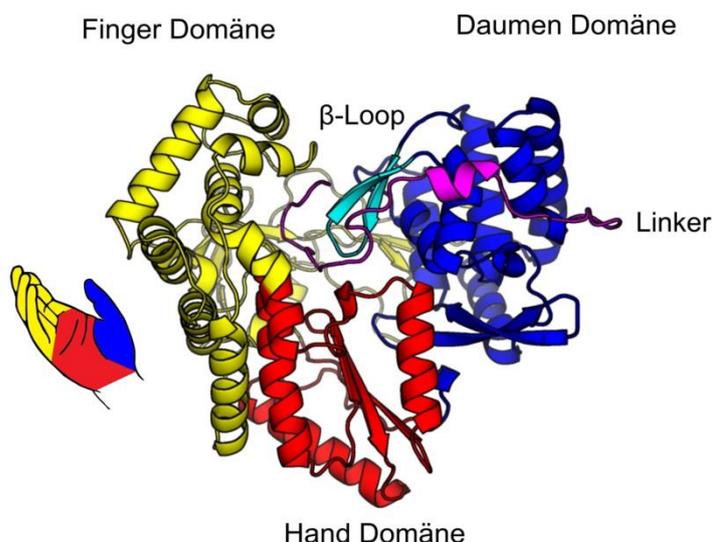


Abbildung 6: Nicht-Struktur Protein 5B des Hepatitis C Virus (PDB Index: 2XI3¹).

Das Protein besitzt eine Rechte-Hand-Topologie, unterteilt in Finger-Domäne (gelb), Hand-Domäne (rot) und Daumen-Domäne (blau). Weitere wichtige Proteinelemente sind der Linker (magenta) und der β -Loop (cyan).

Das aktive Zentrum befindet sich innerhalb der Hand-Domäne, im Zentrum des Proteins. Es enthält das GDD-Motiv (Aminosäuren 317-319, siehe Tabelle 1), auch Motiv C genannt, welches zwei Magnesiumionen koordiniert, die das Triphosphatende der neu einzubauenden Nukleotide komplexieren.⁵⁴ Das Zentrum des Proteins wird von der Finger- und der Daumen-Domäne eng umschlossen.⁵⁵ Für die Replikation gelangt das RNA-Templete über einen Kanal in der Finger-Domäne in das aktive Zentrum.^{56,57} Die Nukleotide gelangen über einen Tunnel zwischen Daumen- und Finger-Domäne zum GDD-Bindemotiv.⁵⁴ Dabei spielt das Motiv F (Aminosäuren 140-160, siehe Tabelle 1) ebenfalls eine wichtige Rolle.⁵⁸

Die dreidimensionale Struktur von NS5B wurde bereits in fast 200 Kristallstrukturen publiziert (Stand Februar 2017), meistens als Komplex mit einem Nicht-Nukleosid-Inhibitor (NNI). In allen Kristallstrukturen nimmt das aktive Zentrum einen geschlossenen Zustand ein, ungeachtet dessen, ob ein Ligand gebunden ist oder nicht. In dieser geschlossenen Konformation können zwar Nukleotide und einzelsträngige RNA (ssRNA) binden, aber durch die Lage des β -Loops und des Linkers kann keine Elongation stattfinden.⁵² Diese zwei Strukturelemente blockieren das katalytische Zentrum und verhindern den Übergang von der Initiation zur prozessiven dsRNA-Synthese.^{51,53,59}

Tabelle 1: Liste der Einteilungen von NS5B nach Domänen und Motiven sowie deren Funktionen.

Proteineinteilung	Aminosäurenummer	Funktion	Referenz
Finger-Domäne	1-181, 223-283		51,53
Hand-Domäne	182-222, 284-358		51,53
Daumen-Domäne	359-529		51,53
Linker	530-570	verhindert die dsRNA Elongation	50
β -Loop	443-454	verhindert die dsRNA Elongation und die Interaktion von RNA mit NTP	52,57
Δ 1-Loop	11-46	definiert die aktive Form und interagiert über Helix A mit der Daumen-Domäne	60,61,62
Helix A	25-35	ist Teil des Δ 1-Loops	
Motiv A	218-226	koordiniert die Metallionen	52,53
Motiv B	282-292	koordiniert die Interaktion zwischen der Template-RNA und den Nukleotiden	52,53,63
Motiv C (=GDD Motiv)	317-319	koordiniert die Metallionen	52,53
Motiv D	345-349	verantwortlich für die Enzymaktivität	52,53
Motiv E	362-367	verantwortlich für die Elongation	52,53
Motiv F	140-160	interagiert mit den Nukleotiden	52,53,57
Motiv G	91-102	bindet Template-RNA	64,65

Die vielen Untersuchungen auf dem Gebiet der NS5B Polymerase haben zu einigen postulierten Mechanismen der RNA-Doppelstrangbildung geführt. Es wird davon ausgegangen, dass eine konformationelle Änderung unerlässlich für die Elongation ist.⁶² Der β -Loop und der Linker müssten sich dafür aus dem aktiven Zentrum wegbewegen, um Platz für den entstehenden Doppelstrang zu schaffen.¹ Dieser Verlagerung folgt eine Rotation der Daumen-Domäne.^{59,66} Des Weiteren wird angenommen, dass die Helix A, die Teil des Δ 1-Loops ist, in die konformationelle Änderung der Polymerase involviert ist. Der genaue Mechanismus ist dabei unklar. Während der Zustandsänderung von NS5B bleiben die Hand-Domäne, welche die Motive A bis D enthält, und die Motive E und F starr.⁵⁹ Mit Hilfe der Spektroskopie des Circular dichroismus (CD-Spektroskopie) konnte bereits gezeigt werden, dass die Bindung von Nukleotiden eine konformationelle Veränderung in einer NS5B Δ 21-

Variante auslöste.⁶⁷ Die $\Delta 21$ -Variante von NS5B entspricht einem rekombinanten Protein ohne die letzten 21 Aminosäuren des C-Terminus bzw. des Membranankers

1.7. Die Nicht-Nukleosid-Inhibitoren von NS5B

Die Suche nach geeigneten Inhibitoren für die HCV RNA-Polymerase führte zur Entwicklung von Nukleosid-Analoga, die als alternative Substrate die Doppelstrang-Replikation unterbrechen, und von Nicht-Nukleosid-Inhibitoren (NNI). Für NS5B existieren vier allosterische NNI-Bindestellen im Enzym (siehe Abbildung 7). Zwei davon befinden sich in der Daumen-Domäne (NNI-1 und NNI-2) und zwei in der Hand-Domäne (NNI-3 und NNI-4).⁶⁸ Die Form und Lage dieser Taschen ist innerhalb des Proteins sehr unterschiedlich. Zusätzlich existieren einige Kristallstrukturen mit NNIs, die eine fünfte Bindestelle aufweisen (NNI-5). Diese wird jedoch häufig zur NNI-4 oder NNI-3 Bindestelle gezählt, da sich die interagierenden Aminosäurereste überschneiden.

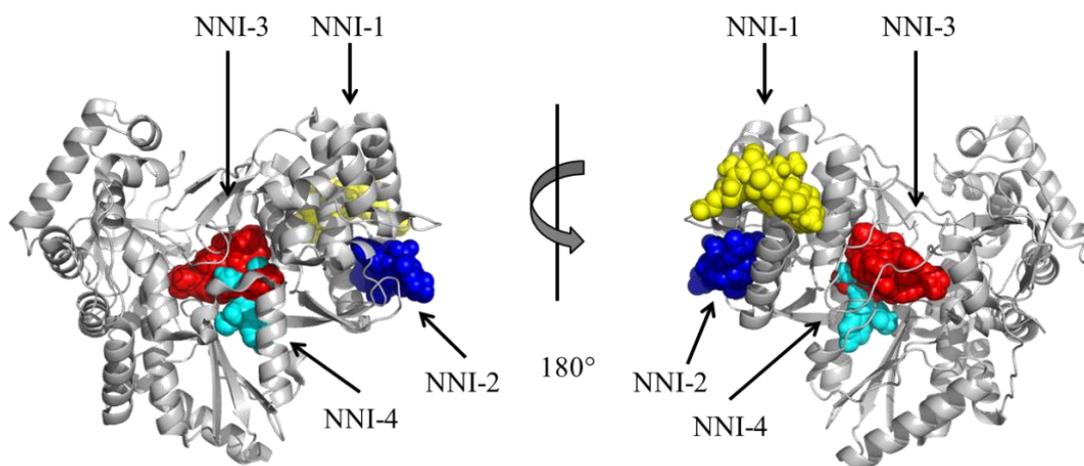


Abbildung 7: Die Bindestellen der Nicht-Nukleosid-Inhibitoren von NS5B.

Die Polymerase ist von vorn und von hinten in grau gezeigt. Die Bindestellen sind farblich markiert mit NNI-1 in gelb, NNI-2 in blau, NNI-3 in rot und NNI-4 in cyan. NNI-1 und NNI-2 befinden sich in der Daumen-Domäne, NNI-3 und NNI-4 in der Hand-Domäne. Die Bindestelle von NNI-5 ist für eine bessere Übersicht nicht mit aufgeführt.

Die NNI-1-Bindestelle liegt zwischen der Helix A des $\Delta 1$ -Loops und dem gegenüberliegenden Daumenabschnitt und ist ca. 30 Å vom aktiven Zentrum entfernt.⁶⁹ Strukturklassen, die mit dieser Bindestelle interagieren, sind Benzimidazole und Indol-6-carboxylate.⁶⁸ Sie inhibieren NS5B während der Initiationsphase der RNA Synthese.⁶⁹ Allerdings zeigen sie keine inhibitorische Aktivität während der Elongationsphase oder bei vor-inkubierten NS5B-RNA-Komplexen.^{70,71} Dies lässt vermuten, dass die Bindetasche bei

der Elongation in einem anderem Zustand vorliegt bzw. eine andere Form hat, oder dass die $\Delta 1$ -Loop-Daumen-Interaktion während dieses Prozesses nicht ausschlaggebend ist.^{57,70,71}

Die Inhibitoren der Klasse NNI-2 binden an eine hydrophobe Bindetasche in der Daumen-Domäne, die sich in der Nähe des Linkers und ca. 35 Å weit weg vom aktiven Zentrum befindet.⁷² Es wird angenommen, dass die Bindung dieser NNIs die Flexibilität der Polymerase so sehr beeinflusst, dass ein Wechsel in die prozessive Konformation nicht mehr möglich ist.⁷³ Der genaue Mechanismus ist jedoch unklar. Strukturklassen, die an diese Tasche binden, sind unter anderem Derivate des Dihydropyrons, Phenylalanins, Thiophens, Thiazolons, Pyridins, 4-Chinolons und Isochinolins.⁷⁴

Die NNI-3-Bindestelle befindet sich zwischen der Hand- und der Daumen-Domäne. Sie ist 15 Å weit und 20 Å tief und nimmt einen Teil des aktiven Zentrums mit ein.⁷⁴ Zu den hier bindenden Inhibitoren gehören eine Vielzahl an unterschiedlichen Stoffklassen, unter anderem Benzothiadiazine, Benzodiazepine, Benzamide, Sulfonamide, Acrylsäuren, Indol-2-carboxylate, Rhodanine, 2-Chinolone, Thiazolidine, Pyrrolidine oder Zimtsäure-Derivate. Diese Inhibitoren sollen sowohl die Initiation der RNA Synthese als auch den Elongationsprozess verhindern.^{75,76}

Die vierte Klasse von NNIs bindet in der Nähe der NNI-3-Bindestelle zwischen Hand- und Daumen-Domäne. Sie ist jedoch sehr viel kleiner und reicht weniger weit in das aktive Zentrum hinein. In diese Tasche konnten ausschließlich Benzofurane wie HCV-796⁷⁷ kristallisiert werden. Der genaue Wirkungsmechanismus der Benzofurane ist noch nicht bekannt. Kinetische Untersuchungen haben jedoch gezeigt, dass bei der Anwesenheit von HCV-796 keine dsRNA gebildet, sondern die ssRNA wieder vom Protein frei gegeben wird.⁶⁷

Die NNI-5-Bindestelle wurde erstmal von Ando *et al.*⁷⁸ beschrieben und konnte durch weitere Kristallstrukturen belegt werden (3LKH und 3QGD-I).^{79,80} Sie befindet sich zwischen allen drei Domänen und verschließt quasi den dsRNA-Ausgang. An diese Bindestelle binden Benzamide, Piperazine und Pyridazine.

1.8. Motivation und Zielstellung

Die RNA-abhängigen RNA-Polymerasen (*RdRPs*) besitzen eine rechte Hand-Topologie, eingeteilt in Finger-, Hand- und Daumen-Domäne. Im Vergleich zu der geöffneten U-Konformation von DNA-Polymerasen sieht die allgemeine Architektur einer RNA-Polymerase eine geschlossene Konformation vor, welche durch den engen Kontakt zwischen

Finger- und Daumen-Domäne hervorgerufen wird.^{61,81} Die Finger-Domäne formt den RNA-Tunnel und den NTP-Eingang. Das Motiv F ist besonders konserviert und einzigartig für RNA-Polymerasen. Die Hand-Domäne beinhaltet das katalytische Zentrum und ist ebenfalls hoch konserviert. Das katalytische Zentrum besteht immer aus einem dreifach-antiparallelen β -Faltblatt-Kern. Die Form und Größe der C-terminalen Daumen-Domäne variiert dagegen stark. *Flaviviridae* besitzen im Vergleich zu anderen Virus-Polymerasen eine große Daumen-Domäne. Diese beinhaltet eine einzigartige β -Faltblatt-Struktur, den β -Loop (auch β -Hairpin oder β -flap genannt), der in allen *RdRPs* von Flaviviren vorkommt, aber nur in der Polymerase von Hepatitis C reicht diese Struktur in das aktive Zentrum des Proteins hinein.^{81,82} Eine weitere strukturelle Besonderheit von HCV NS5B ist der C-terminale Linker, der die Polymerase mit dem Membrananker verbindet. Dieser lagert sich ebenfalls in das aktive Zentrum der Polymerase und verschließt den RNA-Doppelstrang-Ausgang.

Die RNA-Polymerase bindet die Substrate, Einzelstrang-RNA (ssRNA) und Nukleotide (NTP), nach einem sequenziellen Zufallsmechanismus (*random bi bi mechanism*). Unabhängig von der Reihenfolge kann zuerst RNA oder NTP gebunden werden, um den Ternärkomplex zu bilden. Ausführliche kinetische Untersuchungen von HCV NS5B erfolgten in den Arbeiten von Reich *et al.*^{67,83}, in denen vor allem die Bindung von Einzelstrang-RNA und von Nukleotiden untersucht wurde. Die Bindung von ssRNA zeigte eine hohe Affinität mit einer vierstufigen Adaptation des Substrates an die Polymerase. Anhand von Circular dichroismus (CD)-Messungen von NS5B konnte eine Konformationsänderung bei Anwesenheit von NTP-Molekülen beobachtet werden. Untersuchungen mit dem Inhibitor HCV-796, der zur Gruppe der NNI-4 gehört, zeigten den Inhibitionsmechanismus durch die Behinderung der Doppelstrang-RNA-Synthese auf.

Aufbauend auf diesen Arbeiten sollte die Konformationsänderung von NS5B molekülmechanistisch untersucht werden. Dazu diente die Analyse der Kristallstrukturen von HCV NS5B als Grundlage für spätere Moleküldynamik-Simulationen. Diese erfolgten mit verschiedenen Substraten und Inhibitoren, wobei ssRNA, NTP und HCV-796 von besonderem Interesse waren. Insbesondere die Tertiärstruktur-Änderungen bei Anwesenheit von Nukleotiden sollten dadurch geklärt werden, aber auch der Zustand der aktiven Polymerase mit einem geöffneten dsRNA-Ausgang war von Bedeutung. Die Untersuchung verschiedener Inhibitor-Klassen auf ihren molekularen Mechanismus ist ein weiteres Ziel dieser Arbeit. Die *in silico*-Daten sollten durch *Crosslink*-Experimente unterstützt werden und die Moleküldynamik-Modelle verifizieren. Um die Quervernetzung des Proteins qualitativ nachzuweisen, wurden ein deuterierter und ein nicht-deuterierter *Crosslinker* verwendet. Der quantitative Nachweis der Konformationsänderung sollte anhand quantitativer *Crosslink*-Experimente erfolgen.

2. GRUNDLAGEN

2.1. Superpositionierung von Proteinstrukturen

Die dreidimensionale Überlagerung (strukturelles *Alignment*) von zwei oder mehreren Proteinstrukturen stellt eine besondere Herausforderung dar, besonders wenn in den Kristallstrukturen Aminosäuren fehlen oder wenn die Sequenzidentität niedrig ist. Im einfachsten Fall erfolgt eine Überlagerung zweier 3D-Strukturen anhand der Methode der kleinsten Quadrate (*least squares*) der C α -Atome oder des Peptidrückgrates. Dabei werden vorhandene Lücken in der Sequenz oder der Proteinstruktur vernachlässigt. Um diesem Datenverlust vorzubeugen, verwendet das frei verfügbare Programm THESEUS⁸⁴ den „*expectation-maximization*“ (EM)-Algorithmus für eine maximale Wahrscheinlichkeit (*Maximum-Likelihood*) der Überlagerung der Strukturen.^{85,86} Die *Maximum-Likelihood*-Methode ist die Schätzung einer Regressionsfunktion, bei der die unbekannt Daten so gewählt werden, dass die neuen Daten eine maximale Plausibilität (*Likelihood*) besitzen.⁸⁷ Dafür wird ein Sequenzalignment der zu positionierenden Strukturen benötigt.

2.2. Molekülmechanik

Im Gegensatz zur Quantenmechanik werden bei der Molekülmechanik die Bewegungen der Elektronen vernachlässigt und die Energie des Systems als eine Funktion der Kernpositionen berechnet. Ohne Berücksichtigung der Elektronen ergeben sich allerdings Einschränkungen in den Molekülmechanik-Rechnungen, was die Beschreibung physikalischer Größen betrifft. So können mit der Molekülmechanik keine chemischen Reaktionen, elektronisch angeregte Zustände oder magnetische Eigenschaften berechnet werden. Bei Kraftfeldrechnungen werden dementsprechend keine Bindungen gebrochen oder gebildet, d.h. Bindungen, die man bei der Eingabe des Moleküls als solche definiert hat, bleiben bestehen. Hierbei müssen für alle denkbaren Bindungszustände im Voraus Parameter definiert werden. Der entscheidende Vorteil der Molekülmechanik gegenüber der Quantenmechanik liegt darin, weitaus größere Systeme in viel kürzerer Zeit berechnen und simulieren zu können. Um eine berechenbare Beschreibung eines Systems zu erhalten, werden Kraftfelder verwendet.^{88,89}

2.2.1. Kraftfelder

Die Grundlage eines Kraftfelds basiert auf dem Kugel-Feder-Modell. Dabei werden Atome als Kugeln und ihre Bindung als Federn betrachtet. Kraftfelder beschreiben Einzelenergien, deren Summe in der Gesamtenergie eines Systems resultiert. Dabei wird davon ausgegangen, dass die potentielle Energie eines Moleküls aus bindenden und nichtbindenden Wechselwirkungen zusammengesetzt ist.

$$E_{\text{total}} = E_{\text{bindende Wechselwirkungen}} + E_{\text{nichtbindende Wechselwirkungen}} \quad \text{Gleichung 1}$$

$$= E_{\text{Bindung}} + E_{\text{Winkel}} + E_{\text{Torsion}} + E_{\text{VanderWaals}} + E_{\text{Coulomb}} \quad \text{Gleichung 2}$$

Die bindenden Wechselwirkungen bestehen aus den Energien für die Deformation von Bindungslängen, Bindungswinkeln und Torsionswinkeln, und die nichtbindenden aus den elektrostatischen (Coulomb) und den Van-der-Waals-Kräften. Damit können die einzelnen Energiebeträge nach Gleichung 3 berechnet werden:

$$E(r^N) = \sum_{\text{Bindungslänge}} \frac{k_i}{2} (l_i - l_{i,0})^2 + \sum_{\text{Bindungswinkel}} \frac{k_i}{2} (\theta_i - \theta_{i,0})^2 + \sum_{\text{Torsion}} \frac{V_n}{2} (1 + \cos(n\omega - \gamma)) \\ + \sum_{i=1}^N \sum_{j=i+1}^N 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \sum_{i=1}^N \sum_{j=i+1}^N \left(\frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right) \quad \text{Gleichung 3}$$

Die potentielle Energie E ist eine Funktion der Position r der Atome N . Der erste Term stellt die Wechselwirkung zweier Atome dar, die durch eine Bindung gekoppelt sind. Er ist umso größer, je mehr die Bindungslänge l_i von der Referenzlänge $l_{i,0}$ abweicht. Der zweite Term beschreibt die Summe der Energien für die Bindungswinkel θ eines Moleküls. Diese beiden Terme werden durch das Hook'sche Gesetz berechnet. Der dritte Term beschreibt die Torsionswinkeldeformationen und enthält die periodische Änderung der Energie bei der Drehung einer Einfachbindung. Die nichtbindenden Wechselwirkungen sind im vierten und fünften Term angegeben. Sie werden zwischen allen Atompaaren i und j berechnet, die sich entweder in verschiedenen Molekülen befinden, oder die im gleichen Molekül mindestens durch drei Bindungen voneinander getrennt sind. Die elektrostatischen Wechselwirkungen werden meist durch die Coulomb-Kraft und die Van-der-Waals-Wechselwirkungen durch das Lennard-Jones-Potential angegeben. Die atomspezifischen Parameter und Konstanten des Kraftfeldes werden durch k , l_0 , θ_0 , V_n und σ beschrieben und wurden empirisch ermittelt.^{88,90} Je nach Spezifizierung des Kraftfeldes werden weitere Terme in die Gleichung integriert, z.B. Energierme für „out of plane“-Winkel oder Wasserstoffbrücken-Wechselwirkungen.

2.2.2. Energieminimierung

Bei der Energieminimierung eines Moleküls wird die Geometrie iterativ verändert, sodass sich die Molekülenergie bei jedem Berechnungsschritt verringert, bis ein lokales Energieminimum erreicht wird. Anders ausgedrückt ist es die Suche nach einem lokalen Minimum entlang der Energiehyperfläche eines Moleküls, welche die potentielle Energie in Abhängigkeit der Ortskoordinaten darstellt, wobei der Gradient Null wird. Für eine Energieminimierung von einem Molekül von N Atomen müssen bei einem kartesischen Koordinatensystem $3N$ unabhängige Variablen berechnet werden, da jedes Atom drei Freiheitsgrade besitzt. Dieses nichtlineare Optimierungsproblem kann durch einige etablierte Algorithmen gelöst werden.

Die *steepest descent*-Methode⁹¹ beruht auf der ersten Ableitung (Gradient) der potentiellen Energie E der Position x_i des Atoms i als Suchkriterium für ein lokales Energieminimum. Das Vorzeichen und der Betrag der ersten Ableitung geben an, in welche Richtung und mit welcher Schrittweite die nächsten Koordinaten auf der Energiehyperfläche berechnet werden. Danach wird die potentielle Energie von neuem bestimmt. Diese Abfolgen sind iterativ und begrenzt durch eine festgelegte Anzahl an Iterationsschritten oder unterliegen einem vorher bestimmten Grenzwert der Steigung der ersten Ableitung. Bei der *steepest descent*-Methode kann es zu einer Oszillation um das Energieoptimum kommen, wenn die Startstruktur bereits nah am lokalen Minimum liegt oder wenn das Minimum keine großen Steigerungen aufweist und sich die Energien nur noch wenig ändern. Die *steepest descent*-Methode eignet sich daher besser für die ersten Schritte einer Energieminimierung.

Ein weiterer gebräuchlicher Algorithmus ist die *conjugate gradient*-Methode⁹², welche eine Weiterentwicklung von *steepest descent* ist. Auch sie bezieht sich auf die erste Ableitung der potentiellen Energie E an der Position x_i des Atoms i . Die Auswahl der Schrittrichtung basiert hier allerdings nicht auf der Information des Gradienten allein, sondern sie bezieht alle vorherigen Gradienten für den nächsten Schritt auf der Energiehyperfläche mit ein. Dadurch wird eine bessere Konvergenz zum Energieminimum erreicht. Andererseits verlängern sich auch die Rechenschritte mit zunehmender Anzahl an Iterationen. Die *conjugate gradient*-Methode eignet sich daher für Optimierungen nahe am Energieminimum.

2.3. Moleküldynamik

Die Moleküldynamik (MD) beschreibt ein Teilgebiet der molekularen Modellierung und basiert auf dem Prinzip der Molekülmechanik und den Kraftfeldern. Dieses Verfahren wird meist zur Suche in Konformationsräumen oder zur Untersuchung von Molekülbewegungen verwendet. Die Bewegungen der einzelnen Atome eines Systems werden durch die Integration der zweiten Newtonschen Bewegungsgleichung ($F = m \cdot a$) auf der Grundlage des vorher ausgewählten Kraftfeldes berechnet (siehe Gleichung 4).⁹³

$$F_{r_i} = m_i \frac{\partial^2 r_i}{\partial t^2}$$

Gleichung 4

Gleichung 4 beschreibt die Bewegung eines Teilchens der Masse m_i entlang einer Koordinate r_i mit der Kraft F_i , die auf das Teilchen einwirkt. Um den zeitlichen Verlauf eines Systems (Trajektorie) zu erhalten muss die Kraft F_i berechnet werden, die sich auf die potentielle Energie (E) des Systems (N) an Position r bezieht.

$$F_i = -\nabla_i E(r^N) = -\frac{\partial E(\vec{r}_1, \dots, \vec{r}_N)}{\partial \vec{r}}$$

Gleichung 5

Die potentiellen Energien der Teilchen des Systems werden durch das vorher ausgewählte Kraftfeld bestimmt. Sind die Anfangskoordinaten, die Beschleunigung und die Geschwindigkeiten der einzelnen Atome bekannt, kann man über die zeitliche Integration den Ort jedes Teilchens zu einem beliebigen Zeitpunkt berechnen. Das ausgewählte Zeitintervall (Δt) muss dabei viel kleiner sein, als die schnellste Bewegung im Molekül. Ein typischer Zeitschritt einer MD entspricht 1 oder 2 fs ($1 \text{ fs} = 10^{-15} \text{ s}$).⁹⁴ Üblicherweise werden diese Zeitintervalle (Δt) immer wieder neu berechnet, bis ein Schrittlimit erreicht ist. Um also eine Trajektorie von mehreren Nanosekunden ($1 \text{ ns} = 10^{-9} \text{ s}$) mit einem Zeitintervall (Δt) von 1 oder 2 fs zu erhalten, müssen mehrere Millionen Berechnungsschritte ausgeführt werden.

Zu Beginn einer MD müssen die Startkoordinaten, die Starttemperatur und das Zeitintervall des zu simulierenden Systems definiert sein. Die Startkoordinaten stammen häufig aus Kristall- oder NMR-Strukturen. Den einzelnen Atomen werden Starttemperaturen zugeordnet. Dies erfolgt mit Hilfe der Maxwell-Boltzmann'schen Geschwindigkeitsverteilung, deren Resultat eine Verteilung der Geschwindigkeiten gemäß einer Gauß-Kurve darstellt.^{95,96} Im weiteren Verlauf erfolgt die iterative Berechnung des Potentials und der Kräfte der Atome, deren Ortswechsel zum Zeitschritt Δt sowie das Speichern der neuen Ortskoordinaten und

Energien in einer Trajektorie. Der Ablauf einer MD ist im Fließdiagramm in Abbildung 8 dargestellt.

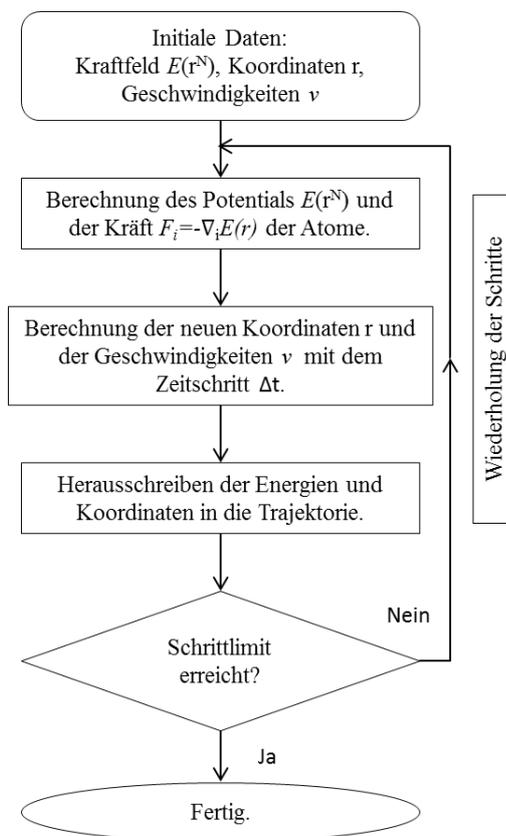


Abbildung 8: Vereinfachende Darstellung einer MD als Fließdiagramm nach A.Kukol.⁹⁴

2.3.1. Integrationsschemata der klassischen Bewegungsgleichung

Da es sich bei der klassischen MD Simulation um ein System mit N nichtlinearen Differentialgleichungen zweiter Ordnung handelt, können die einzelnen Schritte der MD nicht analytisch, sondern nur numerisch gelöst werden. Für diese Integration der Newtonschen Bewegungsgleichung steht eine Vielzahl an numerischen Algorithmen zur Verfügung (*Verlet*,⁹⁷ *Velocity-Verlet*,⁹⁸ *Leapfrog* oder *Beeman*⁹⁹), welche alle der „*Finite Difference*“ Methode (FDM) angehören und meist auf dem *Verlet*-Algorithmus beruhen.⁸⁸ All diese Integrationsalgorithmen beruhen auf der Annahme, dass die Position, die Geschwindigkeit und die Beschleunigung durch *Taylor*-Reihen ausgedrückt werden können. Da *Taylor*-Reihen instabil und ungenau sind, wurde daraufhin der *Verlet*-Algorithmus entwickelt.⁹⁵ Er summiert die *Taylor*-Reihen für $(t+\delta t)$ (siehe Gleichung 6) und $(t-\delta t)$ (siehe Gleichung 7) und resultiert in Gleichung 8, wobei sich die Geschwindigkeiten der Atome herauskürzen.

$$r(t + \delta t) = r(t) + \delta t v(t) + \frac{1}{2} \delta t^2 a(t)$$

Gleichung 6

$$r(t - \delta t) = r(t) - \delta t v(t) + \frac{1}{2} \delta t^2 a(t)$$

Gleichung 7

$$r(t + \delta t) = 2r(t) - r(t - \delta t) + \delta t^2 a(t)$$

Gleichung 8

Der *Verlet*-Algorithmus⁹⁷ gebraucht die Atomkoordinaten und Beschleunigungen (a) zum Zeitpunkt t sowie die Positionen vom vorherigen Zeitpunkt $r(t-\delta t)$, um die neuen Koordinaten des darauffolgenden Zeitintervalls $r(t+\delta t)$ zu bestimmen. Die Geschwindigkeiten können anschließend berechnet werden, indem die Differenz der Positionen zu den Zeitpunkten $(t+\delta t)$ und $(t-\delta t)$ durch das entsprechende Zeitintervall ($2\delta t$) geteilt werden (Gleichung 9).

$$v(t) = \frac{[r(t + \delta t) - r(t - \delta t)]}{2\delta t}$$

Gleichung 9

Alternativ kann die Geschwindigkeit zum halben Zeitschritt $(t+\frac{1}{2}\delta t)$ bestimmt werden, wie in Gleichung 10.

$$v\left(t + \frac{1}{2}\delta t\right) = \frac{[r(t + \delta t) - r(t)]}{\delta t}$$

Gleichung 10

Der *Verlet*-Algorithmus berechnet die Geschwindigkeiten über zwei Zeitschritte. Sie sind zur Bestimmung der kinetischen Energien und damit der Gesamtenergien notwendig. Ein weiterer Nachteil ist, dass er kein „*self-starting*“-Algorithmus ist, da immer die Positionen vom vorherigen Zeitschritt $r(t-\delta t)$ benötigt werden. Gerade zu Beginn der Simulation ($t=0$) müssen die vorherigen Positionen z.B. durch eine *Taylor*-Reihe bestimmt werden.^{88,93}

Am gebräuchlichsten sind bei einer MD-Simulation der *Velocity-Verlet*- und der *Leapfrog*-Algorithmus. Der *Leapfrog*-Algorithmus ist eine Adaption des *Verlet*-Algorithmus und bezieht sich auf die Atomkoordinaten zum nächsten Zeitintervall $r(t+\delta t)$ (Gleichung 11) und den Geschwindigkeiten zum nächsten halben Zeitschritt $v(t+\frac{1}{2}\delta t)$ (Gleichung 12):

$$v\left(t + \frac{1}{2}\delta t\right) = v\left(t - \frac{1}{2}\delta t\right) + a(t)\delta t$$

Gleichung 11

$$r(t + \delta t) = r(t) + \delta t v\left(t + \frac{1}{2}\delta t\right)$$

Gleichung 12

Zu Beginn werden die Geschwindigkeiten zum nächsten halben Zeitschritt $v(t+\frac{1}{2}\delta t)$ berechnet. Auf diese Weise können mit Gleichung 12 die neuen Koordinaten $r(t+\delta t)$ berechnet werden. Die Momentangeschwindigkeit zum Zeitpunkt t kann aus folgender Gleichung gemittelt werden:

$$v(t) = \frac{1}{2} \left[v \left(t + \frac{1}{2} \delta t \right) + v \left(t - \frac{1}{2} \delta t \right) \right]$$

Gleichung 13

Der *Leapfrog*-Algorithmus enthält also schon die Berechnung der Geschwindigkeiten. Allerdings können bei dieser Berechnungsmethode die Geschwindigkeiten und Koordinaten nicht synchron dokumentiert werden, sondern immer um einen halben Zeitschritt versetzt. Dies bedeutet, dass die berechnete kinetische Energie gemittelt werden muss, um in die Gesamtenergie des Systems mit einfließen zu können, da die Positionen und die Geschwindigkeiten zu unterschiedlichen Zeitpunkten berechnet werden.

Eine andere Weiterentwicklung des *Verlet*-Algorithmus ist der *Velocity-Verlet*-Algorithmus, der die Koordinaten, Geschwindigkeiten und Beschleunigungen zu einem Zeitpunkt berechnen kann und keine Atomkoordinaten von dem Zeitpunkt t benötigt. Er erfolgt in 3 Schritten. Die Position zum neuen Zeitpunkt $r(t+\delta t)$ wird aus der aktuellen Position, sowie aus der Momentangeschwindigkeit und Beschleunigung berechnet (siehe Gleichung 15). Die Beschleunigung zum neuen Zeitpunkt $a(t+\delta t)$ resultiert aus der potentiellen Energie der neuen Atomkoordinaten (siehe Gleichung 16). Die neue Geschwindigkeit zum nächsten Zeitpunkt $v(t+\delta t)$ wird im letzten Algorithmus-Schritt zur Hälfte aus der aktuellen Beschleunigung $\frac{1}{2}a(t)$ und zur Hälfte aus der Beschleunigung zum nächsten Zeitpunkt $\frac{1}{2}a(t+\delta t)$ berechnet.

$$r(t + \delta) = r(t) + \delta t v(t) + \frac{1}{2} \delta t^2 a(t)$$

Gleichung 14

$$a(t + \delta t) = \frac{F}{m} = - \frac{1}{m} \frac{\partial E(r_N(t + \delta t))}{\partial r_N}$$

Gleichung 15

$$v(t + \delta t) = v \left(t + \frac{1}{2} \delta t \right) + \frac{1}{2} \delta t [a(t) + a(t + \delta t)]$$

Gleichung 16

$$\text{mit } v \left(t + \frac{1}{2} \delta t \right) = v(t) + \frac{1}{2} \delta t a(t)$$

Gleichung 17

In Abbildung 9 sind die Vorgehensweisen der oben vorgestellten Integrationsalgorithmen graphisch dargestellt.

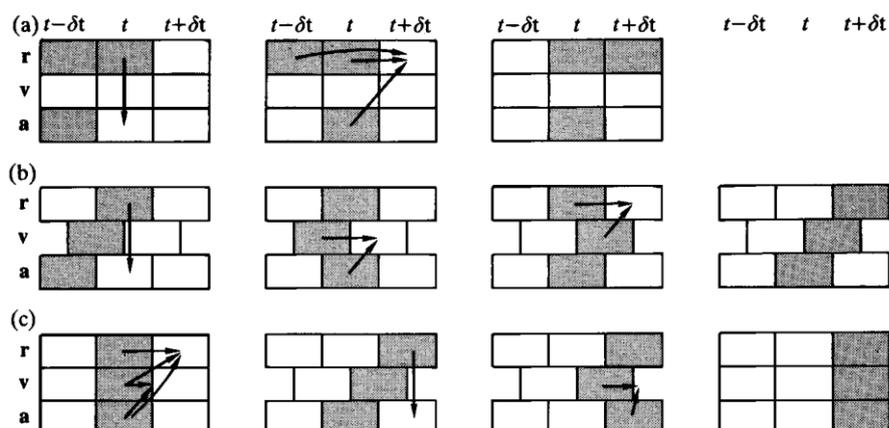


Abbildung 9: Schematische Darstellung der Vorgehensweise der Integrationsmethoden.⁹⁶

a) *Verlet*-Algorithmus, **b)** *Leapfrog*-Algorithmus und **c)** *Velocity-Verlet*-Algorithmus. Die Pfeile deuten an, welche physikalischen Größen von welchen Variablen als nächstes berechnet werden. Am Ende der Abbildung sind alle Eigenschaften des Algorithmus für den nächsten Zeitpunkt ($t+\delta t$) berechnet und der Ablauf könnte von neuem starten.

2.3.2. Beschränkung von Freiheitsgraden

Die Genauigkeit der Integrationsmethoden ist allerdings immer von der gewählten Schrittgröße anhängig. Bei zu großen Zeitintervallen wird die Berechnung zu ungenau und resultiert in großen Fluktuationen, welche die Simulation instabil machen.⁹⁵ Bei einer Simulation sollte δt nicht größer sein als die durchschnittliche Zeit der schnellstmöglichen Bewegung innerhalb des Systems. Dies entspricht bei einer Simulation von organischen Molekülen und Proteinen der Schwingung einer C-H Bindung, welche 0,5 bis 1 fs dauert. Diese Schwingung ist bei der Untersuchung von Biomolekülen meist von geringer Bedeutung. Um die Effizienz der Simulation zu steigern, verwendet man daher häufig eine feste Beschränkung dieser Freiheitsgrade der C-H Bindung durch Algorithmen wie SHAKE^{100,101}, RATTLE¹⁰² oder LINCS¹⁰³. Dies ermöglicht den Gebrauch von Zeitintervallen von 1 bis 2 fs und reduziert zudem die Anzahl der zu berücksichtigenden Freiheitsgrade des Systems, woraus eine schnellere MD-Simulation resultiert.

2.3.3. Ensembles

In einem kanonischen Ensemble bleibt die Anzahl der Teilchen N , das Volumen V und die Temperatur T des zu simulierenden Systems konstant. Man spricht in diesem Fall auch von einem NVT Ensemble. Das isotherm-isobare Ensemble besitzt neben der konstanten Temperatur T auch einen konstanten Druck P (NPT). Diese Ensembles ermöglichen einen

Vergleich zu experimentellen Untersuchungen, welche ebenfalls bei konstanter Temperatur und Volumen oder bei konstanter Temperatur und Druck durchgeführt werden.

Bei einer Moleküldynamik-Simulation ist die Temperatur des Systems direkt abhängig von der kinetischen Energie der Atome und lässt sich daher über die Geschwindigkeit der Teilchen regulieren*. Der einfachste Weg wäre, zu jedem Zeitschritt die Geschwindigkeiten mit einem Faktor λ zu multiplizieren, welcher wie folgt definiert ist (Gleichung 18):

$$\lambda = \sqrt{T_{neu}/T_t}$$

Gleichung 18

Dabei ist T_{neu} die aus der kinetischen Energie berechnete Temperatur des Systems und T_t die gewünschte Temperatur des Systems. Diese Art der Temperaturregulierung kann kein richtiges kanonisches Ensemble wiedergeben und macht die Dynamik des Systems instabil.⁹⁵

Eine weitere Methode, die Temperatur einer MD-Simulation zu regulieren, ist die Kopplung an ein externes Wärmebad, entwickelt von Berendsen *et al.*¹⁰⁴. Dazu wird die Bewegungsgleichung nach Gleichung 19 verändert:

$$a_i = \frac{F_i}{m_i} + \frac{1}{2\tau_T} \left(\frac{T_{neu}}{T_t} - 1 \right) v_i$$

Gleichung 19

Die Kopplungskonstante τ_T bestimmt die Stärke der Kopplung zwischen System und Wärmebad. Ist diese Konstante unendlich groß ($\tau_T \rightarrow \infty$), erhält man wieder das NVE-Ensemble. Im Gegensatz dazu entspricht die Temperaturkontrolle bei einer Kopplungskonstanten, welche dem Berechnungsschritt entspricht ($\tau_T \rightarrow \delta t$), einer Temperaturkontrolle über die Geschwindigkeiten der Teilchen.¹⁰⁴

Weitere Thermostate auf Grundlage der Temperaturkopplung an ein Wärmebad sind der Andersen-Thermostat¹⁰⁵ oder der Nosé-Hoover-Thermostat¹⁰⁶.

Eine Moleküldynamik unter Verwendung des Langevin-Thermostats wird auch als Langevin-Dynamik bezeichnet. Bei dieser Art der Dynamik sind alle Atome eines Systems in einem Meer aus noch kleineren Teilchen eingebettet. Die kleineren Partikel dämpfen dabei die Bewegungen der Atome oder können ihnen auch Stöße verpassen, da sie ebenfalls eine kinetische Energie besitzen. Die Dynamik eines solchen Systems wird durch eine

* Die mittlere kinetische Energie (E_{kin}) eines Moleküls beträgt $\frac{3}{2}k_B T$ (k_B bezeichnet die Boltzmann-Konstante). Molekülmechanisch betrachtet lässt sich E_{kin} ebenfalls aus $\frac{m}{2} v^2$ berechnen.

stochastische Differenzialgleichung, der Langevin-Gleichung, berechnet. Hierbei wird zu der Newtonschen Bewegungsgleichung eines Teilchens i eine zufällige Kraft R_i zum Zeitpunkt t addiert.

$$F_i - \gamma_i v_i + \sigma_i R_i(t) = m_i a_i$$

Gleichung 20

In Gleichung 20 ist v_i die Geschwindigkeit und γ_i der Kollisionskoeffizient des Partikels i .¹⁰⁷ Bei der zufälligen Kraft R_i handelt es sich um einen zentrierten Gaußprozess mit einem konstanten Erwartungswert von 0 und einem Verteilungswert von σ_i . Dieser Verteilungswert berechnet sich wie in Gleichung 21 dargestellt und ist abhängig von dem Kollisionskoeffizienten γ_i , der Boltzmannkonstanten k_B , der Masse des Partikels m_i und der gewählten Temperatur des Systems.

$$\sigma_i = \sqrt{\frac{2\gamma_i k_B T m_i}{\delta t}}$$

Gleichung 21

Die Temperatur des Systems wird folglich durch die Beziehung zwischen $R_i(t)$ und γ_i geregelt, also durch die stochastischen Zusammenstöße der Teilchen.^{93,107}

Genau wie es in der Moleküldynamik Thermostate zur Regulierung der Temperatur gibt, verhält es sich mit Barostaten zur Regelung des Drucks in NPT-Ensembles. Eine Veränderung des Drucks zieht eine Veränderung des Volumens mit sich, und bei einer Veränderung des Volumens müssen auch die Atompositionen dieser Veränderung angepasst werden.

Der Grad der Volumenfluktuation ist abhängig von der isothermalen Kompressibilität κ des umgebenden Mediums. Mit steigender Konstante κ steigt auch die Fluktuation des Volumens eines Stoffes (vgl. Gleichung 22).

$$\kappa = -\frac{1}{V} \left(\frac{\partial V}{\partial p} \right)_T$$

Gleichung 22

Analog zum Thermostat kann man eine Simulation auch an ein „Druck-Bad“ koppeln, um den Druck nicht über das Volumen zu regulieren. Die Druckkopplung lässt sich wie folgt formulieren:

$$\frac{\delta P(t)}{\delta t} = \frac{1}{\tau_p} (p_{Bad} - p(t))$$

Gleichung 23

In Gleichung 23 ist τ_p die Kopplungskonstante, p_{Bad} der Druck des Kopplungsbads und $p(t)$ der Druck des Systems zum Zeitpunkt t . Das Volumen der Simulationsbox wird mit dem Faktor λ entsprechend Gleichung 24 angepasst.

$$\lambda = 1 - \kappa \frac{\delta t}{\tau_p} (p - p_{Bad})$$

Gleichung 24

Die neuen Atompositionen resultieren anschließend aus der dritten Wurzel von λ , multipliziert mit den alten Koordinaten (siehe Gleichung 27).⁸⁸

$$r'_i = \sqrt[3]{\lambda} r_i$$

Gleichung 25

2.3.4. Lösungsmittel

Eine wichtige Komponente einer MD-Simulation ist die Solvation des zu simulierenden Moleküls. Alternativ kann auch eine Simulation im Vakuum erfolgen, was aber wenig vorteilhaft ist, wenn man biologische Systeme untersuchen möchte. Prinzipiell unterscheidet man zwischen expliziten und impliziten Wassermodellen.

Fällt die Wahl auf ein explizites Lösungsmittel, so existieren viele etablierte Wassermodelle, die verwendet werden können. TIP3P¹⁰⁸, TIP4P¹⁰⁸, TIP5P¹⁰⁹ (*transferable intermolecular potential 3,4,5 point*), SPC¹¹⁰ (*simple point charge*) und SPC/E¹¹¹ (*simple point charge extended*) werden dabei am häufigsten verwendet. Dabei handelt es sich um 3- bis 5-Seiten Modelle, die die nichtbindenden Interaktionen von Wasser widerspiegeln sollen. Je nach Modell variieren die geometrischen und physikochemischen Parameter der Wassermoleküle wie Dipol, Dichte, dielektrische Konstante oder Wärmekapazität. Beispielsweise kann der Winkel zwischen H-O-H dem idealen Winkel eines Tetraeders entsprechen (109,47°), oder dem realen Winkel von Wasser (104,45°) angepasst sein. Wird zusätzlich der SHAKE-Algorithmus^{100,101} während der Simulation angewandt, erhält man ein komplett starres Wassermolekülmodell.

Eine weitere Methode, die Lösungsmittelumgebung zu simulieren, ist die Verwendung von implizitem Lösungsmittel. Es ist das Modell eines kontinuierlich isotropischen Mediums, welches eine lösungsmittelähnliche Umgebung simulieren soll. Es eignet sich zur

Verwendung bei der Berechnung der freien Energie von Solvat-Solvens-Interaktionen oder bei Konformationsänderungen von Proteinen, DNA, RNA oder Polysacchariden. Die freie Lösungsenthalpie ($\Delta G_{\text{solvation}}$) beschreibt den Prozess, bei dem ein Stoff aus dem Vakuum in Lösung gebracht wird. Sie wird aus den unterschiedlichen Termen in Gleichung 26 berechnet.

$$\Delta G_{\text{solvation}}^0 = \Delta G_{\text{elek}} + \Delta G_{\text{Hohlraum}} + \Delta G_{\text{vdW}}$$

Gleichung 26

Der elektrostatische Teil der Gleichung ΔG_{elek} (Gleichung 27) beschreibt die elektrostatischen Wechselwirkungen zwischen Solvat und Solvens. Der Stoff muss im Lösungsmittel einen Hohlraum formen, der seinem Volumen entspricht. Die freie Energie, die dazu benötigt wird, ist $\Delta G_{\text{Hohlraum}}$. Dieser Term ist positiv und setzt sich aus dem entropischen Nachteil bei der Reorganisation der Lösungsmittelmoleküle um das Solvat und aus der Verdrängungsarbeit, welche bei der Entstehung des Hohlrums benötigt wird, zusammen. Die Van-der-Waals-Interaktionen ΔG_{vdW} können auch als Summe der London-Kräfte und der Abstoßungskräfte zwischen dem Stoff und dem Lösungsmittel beschrieben werden.

In der Molekülmechanik wird oft das generalisierte Born-Modell (*generalized born*, GB) verwendet, um die elektrostatischen Energien ΔG_{elek} einer Lösung zu beschreiben. Dieses Modell beschreibt die Atome eines Moleküls als eine Kugel mit einem Radius a_i , welcher auch Bornradius genannt wird, und einer Ladung q_i in ihrem Mittelpunkt. Der Innenbereich des Atoms ist einer dielektrischen Konstante (ϵ) von 1 zugeordnet, während die Umgebung des Atoms einer hohen Konstante ϵ von 80 entspricht. Die elektrostatische Energie einer Lösung ($G_{\text{elek}}(\text{Lösung})$) mit einer relativen dielektrischen Konstante ϵ entspricht der Summe der Coulomb-Kraft und der Born-Gleichung (vgl. Gleichung 27).

$$\Delta G_{\text{elek}} = \sum_{i=1}^N \sum_{j=i+1}^N \frac{q_i q_j}{r_{ij}} - \left(1 - \frac{1}{\epsilon}\right) \sum_{i=1}^N \sum_{j=i+1}^N \frac{q_i q_j}{r_{ij}} - \frac{1}{2} \left(1 - \frac{1}{\epsilon}\right) \sum_{i=1}^N \frac{q_i^2}{a_i}$$

Gleichung 27

Die Variablen q_i und q_j sind die Ladungen der Atome i und j und a_i ist der Bornradius von i . Der Abstand zwischen i und j ist r_{ij} . Der erste Term dieser Gleichung (rot markiert) entspricht dabei der Coulomb-Wechselwirkung zweier Ladungen im Vakuum. Um die freie elektrostatische Energie (ΔG_{elek}) zu berechnen, die bei der Lösung eines Stoffes aus dem Vakuum aufgewandt wird, wird die Differenz aus $G_{\text{elek}}(\text{Lösung}) - G_{\text{elek}}(\text{Vakuum})$ berechnet und entspricht der generalisierten Born-Gleichung (Gleichung 28). Diese Gleichung wurde von Still *et al.*¹¹² und Qiu *et al.*¹¹³ weiterentwickelt.⁸⁸

$$\Delta G_{elek} = -\left(1 - \frac{1}{\epsilon}\right) \sum_{i=1}^N \sum_{j=i+1}^N \frac{q_i q_j}{r_{ij}} - \frac{1}{2} \left(1 - \frac{1}{\epsilon}\right) \sum_{i=1}^N \frac{q_i^2}{a_i}$$

Gleichung 28

2.3.5. Periodische Randbedingungen und *cut-offs*

Die Etablierung von periodischen Randbedingungen in einer Simulation dient zur Vermeidung von Randeffekten. Durch die Vervielfachung der Simulations-Box kann ein Teilchen, welches an einem Rand der Box austritt, an dem gegenüberliegenden Rand wieder eintreten. Die Kantenlänge der Box muss ausreichend groß gewählt werden, damit ein zentrales Molekül nicht gleichzeitig mit einem anderen Molekül und dessen Kopie wechselwirkt. Die Simulations-Box kann verschiedene Formen besitzen, z.B. kubisch oder als abgestumpfter Oktaeder (Oktaederstumpf) in AMBER¹¹⁴. Weitere Box-Formen in GROMACS¹¹⁵ sind triklinisch oder haben die Form eines Rhombendodekaeders.

Um zusätzlich Rechenzeit zu sparen, wird für die nicht-bindenden Wechselwirkungen eines Systems ein *cut-off*-Radius r_c bestimmt, der z.B. die Hälfte der Box-Kantenlänge L beträgt ($r_c = L/2$). In einem System mit N Atomen müssen für jedes Atom i die Distanzen zu den anderen Atomen bestimmt werden, damit die Kraft F_i , welche auf i einwirkt, berechnet werden kann. Daraus resultieren $\frac{1}{2}N(N-1)$ Rechenoperationen, welche durch einen *cut-off*-Radius verringert werden können. Dazu verwendet man Nachbar- und Zellen-Listen der Atome.

In den Nachbarlisten werden alle Atome aufgelistet, deren Abstand geringer ist als $(r_c + \Delta_{list})$. Der Betrag von Δ_{list} kann frei gewählt werden und liegt meist zwischen 1 bis 2 Å.⁹⁵ So müssen für jeden zu berechnenden Zeitschritt nur die Atome aus der Nachbarliste berücksichtigt werden, deren Distanz $< r_c$ ist. Diese Liste wird aktualisiert, wenn ein Atom eine größere Distanz als Δ_{list} zurücklegt.

Ähnlich verhält es sich mit der Zellenliste. Hierzu teilt man den Simulationsraum in Felder der Kantenlänge r_c ein. Ein Atom der Zelle i darf nur mit Atomen der eigenen oder mit direkt benachbarten Zellen interagieren. Die Liste wird aktualisiert, wenn ein Atom in eine andere Zelle wechselt.

2.4. Root mean square deviation (RMSD)

Der Unterschied zwischen zwei dreidimensionalen Strukturen wird anhand des *root mean square deviation* (RMSD)-Wertes bemessen (vgl. Gleichung 29).

$$RMSD = \sqrt{\frac{\sum_{i=1}^{N_{\text{Atome}}} d_i^2}{N_{\text{Atome}}}}$$

Gleichung 29

N_{Atome} ist die Anzahl der berechneten Atome und d_i ist die Distanz zwischen den Koordinaten der Atome i . Das Ziel dabei ist es, die relative Orientierung der Moleküle zu finden, indem die RMSD-Funktion minimiert wird.⁸⁸ Bei großen Biomolekülen wie Proteinen werden häufig nur das Proteinerückgrat (N, C $_{\alpha}$, C) oder die C $_{\alpha}$ -Atome zur Berechnung verwendet. Bei der Analyse von Moleküldynamik-Simulationen erhält man so für jeden Zeitschritt einen RMSD-Wert.

2.5. Hauptkomponentenanalyse (PCA)

Die Verwendung der Hauptkomponentenanalyse (engl.: *principal component analysis*, PCA) kann mehrere Ziele verfolgen, z.B. um große Datenmengen zu reduzieren und zu vereinfachen oder gemeinsame, wesentliche Eigenschaften zu identifizieren. Vor allem dient sie dazu, systematische Datenstrukturen vom Datenrauschen zu unterscheiden. Bei der PCA werden mehrdimensionale Daten statistisch in einen niederdimensionalen Raum umgewandelt.¹¹⁶

Für die Hauptkomponentenanalyse einer Moleküldynamiksimulation müssen die translationalen und die rotationalen Bewegungen des Moleküls durch das Überlagern der Koordinaten der Trajektorie herausgefiltert werden. Für N Atome des Moleküls x kann nun die massengewichtete Kovarianzmatrix C erstellt werden. Für die Koordinaten i und j ergibt sich die Matrix C_{ij} wie in Gleichung 30 dargestellt.¹¹⁷

$$C_{ij} = \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle$$

Gleichung 30

Dabei sind die in eckigen Klammern geschriebenen Koordinaten der Durchschnitt aller Datenpunkte, die über die MD erhalten wurden. Diese symmetrische Matrix C wird mit einer orthonormalen Transformationsmatrix R diagonalisiert, und man erhält Gleichung 31:

$$R^T C R = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N)$$

Gleichung 31

Die i -te Spalte der Matrix R ist der Eigenvektor, der zum Eigenwert λ_i gehört. Der Eigenwert ist die mittlere quadratische Fluktuation in Richtung des Eigenvektors. Die Projektion der Eigenvektoren auf die Atomkoordinaten bilden die Hauptkomponenten (*principal components*, PC _{i}) mit $i=1$ bis $i=N$ (Gleichung 32).¹¹⁷⁻¹²¹

$$PC = R^T(x - \langle x \rangle)$$

Gleichung 32

Da die Eigenwerte der Größe nach geordnet werden ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$), entspricht die größte Fluktuation im Protein der ersten Hauptkomponente (PC1).¹¹⁷

2.6. Clusteranalyse

Die Clusteranalyse teilt Daten anhand von Algorithmen in Datengruppen (Cluster) auf. Die Daten eines Clusters verfügen über ähnliche Eigenschaften und unterscheiden sich darin von anderen Clustergruppen.^{87,122} Das hierarchische Clusterverfahren kann zudem innerhalb einer Gruppe weitere Untergruppen bestimmen und besitzt daher eine verbreitete Anwendung. Im ersten Schritt der Analyse werden in einem distanzbasiertes Verfahren die Ähnlichkeiten bzw. Distanzen (Proximitätsmaß) des Datensatzes bestimmt. Anschließend erfolgt die Gruppierung anhand des Clusteralgorithmus, wie z.B. „*single-linkage*“, „*complete-linkage*“, „*average-linkage*“ oder „*ward*“. Im letzten Schritt muss eine optimale Anzahl an Gruppen festgelegt werden.⁸⁷

2.6.1. Bestimmung der optimalen Clusternummer

Die Identifizierung einer geeigneten Clusternummer ist genauso wichtig, wie die Clusteranalyse an sich. Zwei Analysemöglichkeiten sollen im Folgenden kurz erläutert werden, der Silhouettenkoeffizient¹²³ und der Dunn-Index¹²⁴.

Der Silhouettenkoeffizient $S(i)$ vergleicht den durchschnittlichen Abstand der Daten innerhalb eines Clusters (A) mit den durchschnittlichen Abständen der Daten aus anderen Clustern (B). Dies erfolgt nach Gleichung 33:

$$S(i) = \frac{b_i - a_i}{\max(b_i, a_i)}$$

Gleichung 33

Für den Datenpunkt i ist a_i die durchschnittliche Distanz zwischen i und allen anderen Datenpunkten einer Gruppe, und b_i die gemittelte Distanz zwischen i und allen Datenpunkten der nächsten Gruppe. Je höher der Wert des Silhouettenkoeffizienten $S(i)$ an 1 reicht, desto besser ist die Gruppeneinteilung, denn die Distanzen von a_i sind kleiner als die von b_i . Schlechte Einteilungen nähern sich dem Wert -1.^{122,123}

Der Dunn-Index (DI) beschreibt das Verhältnis zwischen der kleinsten Distanz ($dist$) der Datenpunkte aus unterschiedlichen Gruppen (A und B) mit dem längsten Abstand der Daten innerhalb eines Clusters, also dem Durchmesser ($diam$) der Gruppe (siehe Gleichung 34 und Gleichung 35).¹²²

$$diam A \triangleq \max_{x,y \in A} d(x,y)$$

Gleichung 34

$$dist(A,B) \triangleq \min_{x \in A, y \in B} d(x,y)$$

Gleichung 35

Dabei sind A und B Teilmengen von X ($A, B \subset X$) und dürfen keine leeren Mengen sein ($A, B \neq \emptyset$). Die k -te Einteilung von X ist die Gruppe $P = \{X_1, \dots, X_k\}$. Daraus wird der Dunn-Index wie folgt definiert¹²⁴:

$$DI_k = \frac{\min_{1 \leq q \leq k} \min_{1 \leq r \leq k; r \neq q} dist(X_q, X_r)}{\max_{1 \leq p \leq k} diam(X_p)}$$

Gleichung 36

Der Dunn-Index beschreibt einen Wert zwischen 0 und ∞ und sollte möglichst groß sein.¹²⁴ Der maximale Wert einer Analyse mit dem Dunn-Index beschreibt also die optimale Clusternummer.

2.7. Crosslinking

Die Konformationsänderungen von HCV NS5B sollten nicht nur durch computergestützte Methoden untersucht, sondern ebenfalls mittels experimenteller Methoden verifiziert werden. Eine geeignete Methode hierfür ist das *Crosslinking*. Der Grundgedanke ist, nahstehende Strukturelemente eines Proteins durch geeignete Reagenzien mit definierter Länge kovalent miteinander zu verknüpfen und anschließend massenspektrometrisch zu analysieren. Die

dabei entdeckten *Crosslinks* entsprechen je nach Spacerlänge einem bestimmten Abstandsintervall und geben zusätzliche Informationen zum Aufbau des Proteins. Dabei kann ein Protein intramolekular oder mehrere Proteine intermolekular vernetzt werden. Intramolekulare Vernetzungen geben einen Hinweis über die Konformation eines Proteins. Im Gegensatz dazu dient das intermolekulare *Crosslinking* der Aufklärung der Interaktionsflächen mehrerer Proteine.¹²⁵

2.7.1. Vorgehensweise und Identifizierung von *Crosslinks*

Bei dem Verfahren des *Crosslinkings* wird zwischen zwei unterschiedlichen Arbeitsabläufen unterschieden: „*Bottom-up*“ und „*Top-down*“. Die meisten *Crosslink*-Untersuchungen verwenden vorwiegend die „*Bottom-up*“-Vorgehensweise. In diesem Fall wird das verlinkte Protein durch eine SDS-PAGE (*sodium dodecyl sulfate polyacrylamide gel electrophoresis*) aufgetrennt, anschließend proteolytisch aufgespalten und erst danach mit Hilfe der Massenspektrometrie analysiert. Bei der „*Top-down*“-Methode wird das verlinkte und intakte Protein direkt im Massenspektrometer fragmentiert und analysiert.¹²⁶ Die Identifizierung der quervernetzten Peptide aus der MS-Analyse erfolgt mittels geeigneter Software (z.B. StavroX¹²⁷). Aus der bekannten Sequenz des Proteins werden die theoretischen Massen der Peptidfragmente nach dem Verdau mit den eventuellen *Crosslinks* berechnet und mit den real gemessenen Daten verglichen. Die Verifizierung eines möglichen *Crosslinks* erfolgt anschließend anhand der MS/MS-Fragmentspektren. Eine zusätzliche Erleichterung bezüglich der *Crosslink*-Identifizierung ermöglicht die Verwendung von isotope markierten *Crosslinkern*, was zu einer Massenverschiebung im Spektrum führt. Existiert in dem MS-Spektrum die theoretisch vorhergesagte Masse sowohl isotope markiert als auch nicht-markiert, kann von einem verlinkten Proteinfragment ausgegangen werden. Hierfür müssen der isotope markierte und der normale *Crosslinker* im Verhältnis 1:1 zum Reaktionsansatz zugegeben werden. Weitere Möglichkeiten der *Crosslink*-Identifizierung bieten fluorogene *Crosslinker*, welche nach der *Crosslink*-Reaktion fluoreszieren, oder spaltbare *Crosslinker*, welche chemisch spaltbar sind, oder nach der MS-Fragmentierung spezifische Protein-*Crosslink*-Fragmente bilden.¹²⁸

2.7.2. *Crosslink*-Reagenzien

Für die Verknüpfung von Proteinen mit *Crosslinkern* sind unterschiedliche reaktive Gruppen verfügbar. Häufig werden aminreaktive *Crosslinker* eingesetzt, wie *N*-Hydroxysuccinimidester (NHS-Ester), welche primäre Amine (Lysin-Seitenketten und *N*-

Terminus) modifizieren. In Abbildung 10 ist exemplarisch der Reaktionsmechanismus des aminreaktiven Crosslinkers BS²G dargestellt. Zudem reagieren NHS-Ester pH-abhängig mit den Hydroxygruppen der Seitenketten von Serin, Threonin und Tyrosin.^{128–130} Eine weitere Gruppe von *Crosslinkern* sind die sulfhydrylreaktiven Reagenzien, wie Maleimid. Sie reagieren mit der freien SH-Gruppe von Cysteinen. Neben diesen spezifischen *Crosslinkern* gibt es noch die Möglichkeit der Quervernetzung durch photoreaktive Reagenzien, wie Arylazide, Diazirine oder Benzophenone. Diese bilden bei UV-Bestrahlung ein Radikal, welches anschließend mit dem Protein reagiert.¹³¹ Eine *Crosslink*-Reagenz, welche zwei unterschiedliche funktionelle Gruppen besitzen, die mit dem Protein interagieren, heißt heterobifunktionaler *Crosslinker*. Neben den Crosslinkern mit einer definierten Abstandslänge gibt es auch „zero-length“-Crosslinker, die zwei Peptide ohne einen Spacer verknüpfen, wie die Carbodiimide. Ein Vertreter dieser Verbindung ist EDC (1-Ethyl-3-(3-dimethylaminopropyl)carbodiimid). Es aktiviert eine Carboxylgruppe zur Reaktion mit einer Aminogruppe unter Ausbildung einer Amidbindung. Meist erfolgt die Zugabe von Sulfo-NHS zu EDC, um die Stabilität des reagierenden Intermediats zu gewährleisten.

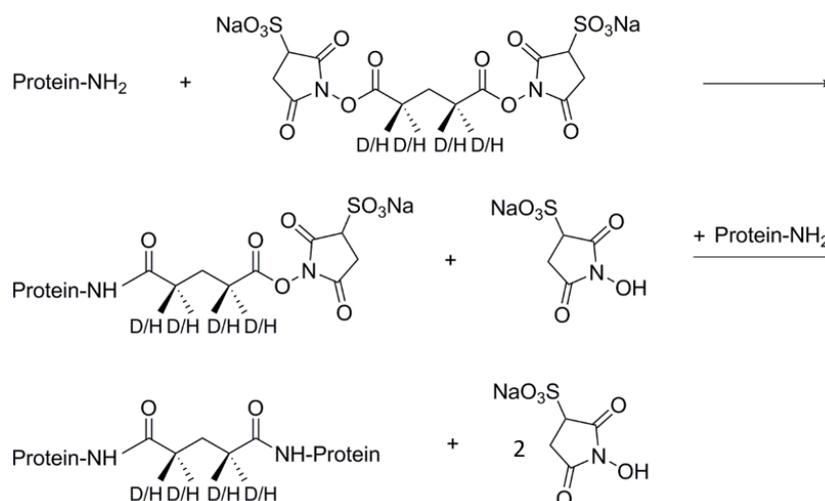


Abbildung 10: Reaktionsmechanismus des aminreaktiven *Crosslinkers* BS²G-d⁰/d⁴.¹³²

2.8. Massenspektrometrie

Die Massenspektrometrie dient der Bestimmung des Masse-zu-Ladungs-Verhältnisses (m/z) von Ionen einer chemischen Verbindung im Hochvakuum. Der Aufbau eines Massenspektrometers besteht im Wesentlichen aus einer Ionenquelle, welche gasförmige

Ionen aus der Probensubstanz erzeugt, einem Massenanalysator, der die Ionen gemäß ihrem Masse-zu-Ladungs-Verhältnisses auf trennt, und einem Detektor, der den Ionenstrom misst. Dabei sind für jedes dieser Module unterschiedliche Komponenten verfügbar, die unterschiedlich kombiniert werden können.¹³³⁻¹³⁵ Da die *Crosslink*-Untersuchungen an einem Orbitrap Fusion Tribrid Massenspektrometer durchgeführt wurden, sollen dessen Komponenten weiter beschrieben werden (siehe Abbildung 11).

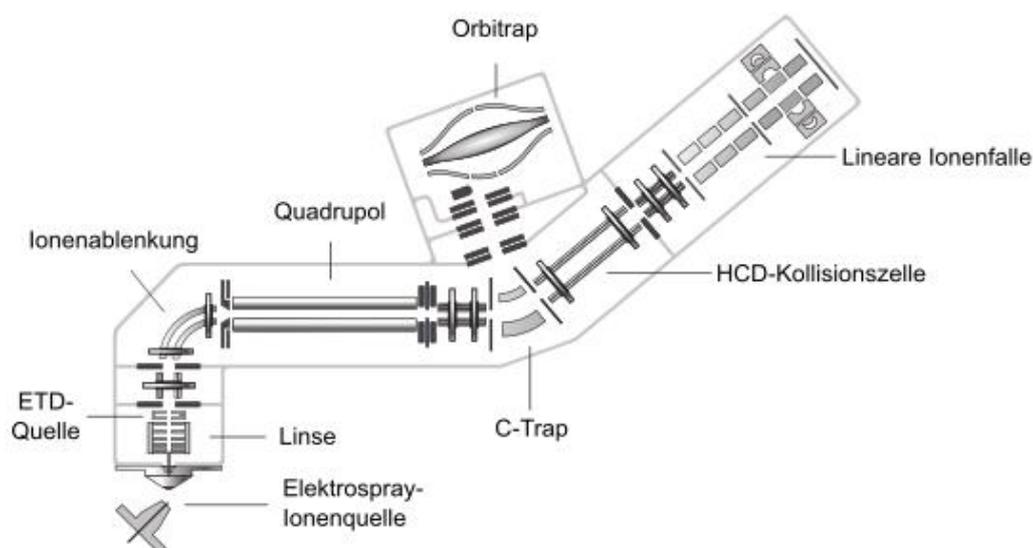


Abbildung 11: Schematische Darstellung des Orbitrap Fusion Tribrid Massenspektrometers (Thermo Fisher Scientific).¹³⁶

2.8.1. Orbitrap Fusion Tribrid Massenspektrometer

Vor der Ionisierung der Probelösung erfolgt die Überführung durch das Einlasssystem des Massenspektrometers. Ein Flüssigkeitschromatograph (Nano-HPLC, UltiMate 3000) ist mit dem Massenspektrometer gekoppelt, um das komplexe Gemisch aus vernetzten und unvernetzten Peptiden der Größe nach zu fraktionieren. Dabei wird die Analyt-Lösung durch eine Kapillare geleitet, an deren Ende eine Spannung angelegt ist und ein elektrisches Feld erzeugt. Dieses elektrische Feld ist für die elektroforetische Trennung und die Bildung kleiner geladener Tröpfchen aus der Lösung verantwortlich. Mit zunehmendem Lösungsmittelverlust durch die Begasung mit einem N_2 -Trockengas (*curtain*-Gas) nimmt die Ladungsdichte an der Tröpfchenoberfläche zu und damit auch die abstoßende Coulomb-Kraft der Teilchen untereinander. Dies führt zu einem spontanen Zerfall, der Coulomb-Explosion, in viele kleinere Nano-Tröpfchen und trägt zur Desolvatisierung des Analyten bei.¹³⁴ Über die Ionenablenkung (*beam guide*) gelangen die geladenen Partikel zum Quadrupol. Dieser besteht aus vier parallel angeordneten Metallelektroden, welche ein elektrisches Feld im inneren

Hohlraum erzeugen, und zwei Endkappen-Elektroden. An den gegenüberliegenden Stabelektroden sind jeweils eine Gleich- und eine Hochfrequenzspannung angelegt mit der gleichen Polarität bzw. Phase. Je nach Stärke des elektrischen Feldes können die Ionen mit dem gewünschten Masse-zu-Ladungs-Verhältnis auf stabilen oszillierenden Flugbahnen den Quadrupol durchlaufen. Die Ionen mit einem abweichendem m/z Verhältnis bewegen sich auf instabilen Bahnen und werden durch die Kollision mit den Metallstäben gestoppt. Nach dieser Selektion der Ionen erfolgt die genaue Detektion im Orbitrap-Analysator. Bevor die Partikel dorthin gelangen, müssen sie die *C-Trap* passieren. Die *C-Trap* ist ein gekrümmter Quadrupol und dient der Akkumulation und der Speicherung der Ionen. In der *C-Trap* kommt es durch Kollisionskühlung (*collisional cooling*) mit Stickstoffgas zum Verlust der kinetischen Energie der Ionen. Dabei ist dieser Gasstrom gering genug um Fragmentierungen zu vermeiden.^{137–140} Durch Hochspannungspulse verlassen die Ionen die *C-Trap* und gelangen in den Orbitrap-Analysator. Dieser besitzt eine fassförmige zweiteilige äußere Elektrode und eine spindelförmige Zentralelektrode. Die äußeren Elektroden sind geerdet und von der inneren isoliert. Wie auch bei anderen Ionenfallen besitzt der Orbitrap-Analysator zwei Endkappenelektroden, welche eine Potentialbarriere erzeugen, sodass die Ionen den Analysator nicht seitlich verlassen können.^{134,140} Durch die elektrostatische Anziehung der Zentralelektrode und durch die Tangentialgeschwindigkeit, mit der die Ionen aus der C-Trap geschleudert wurden, ergibt sich eine spiralförmige Kreisbewegung, mit der sich die Ionen um die Zentralelektrode bewegen. Zusätzlich treten seitliche Oszillationen auf, deren Frequenz (ω_ϕ) sich umgekehrt proportional zur Quadratwurzel des m/z Verhältnisses der Teilchen verhält (vgl. Gleichung 37).¹⁴¹

$$\omega_\phi = \sqrt{k} \left(\frac{q}{m_i} \right)$$

Gleichung 37

Zusätzlich zur Ionenladung (q) und der Ionenmasse (m_i) hängt die Frequenz (ω_ϕ) noch von der Krümmung (k) des elektrischen Feldes in der Orbitrap ab.¹⁴² Die Oszillationen werden von der äußeren Elektrode als induzierter Strom gemessen. Mittels schneller Fourier-Transformation (*Fast Fourier Transformation*, FFT) wird die Oszillation in ein Massenspektrum umgewandelt.^{137,139}

Aus diesem ersten Übersichtsspektrum können ausgewählte Vorläuferionen (*precursor ion*) weiter fragmentiert und analysiert werden. Die Fragmentierung bei der Tandem-Massenspektrometrie, auch MS/MS oder MSⁿ genannt, erfolgt durch eine kollisionsinduzierte Dissoziation (*higher-energy collisional dissociation*, HCD) in der HCD-Kollisionszelle. Ein Kollisionsgas (z.B. N₂) versetzt die Bindungen entlang des Peptidrückgrats durch die

kinetische Energie des Zusammenstoßes in Schwingungen. Diese Vibrationsenergie verursacht anschließend die Fragmentierung des Peptids. Dabei entstehen meist y- und b-Ionen durch Spaltung der Peptidbindungen (siehe Abbildung 12)¹⁴³. Anschließend werden die Peptidfragmente im Orbitrap-Analysator untersucht.¹⁴⁴ Eine Detektion der Partikel in der linearen Ionenfalle wurde in dieser Arbeit nicht vorgenommen.

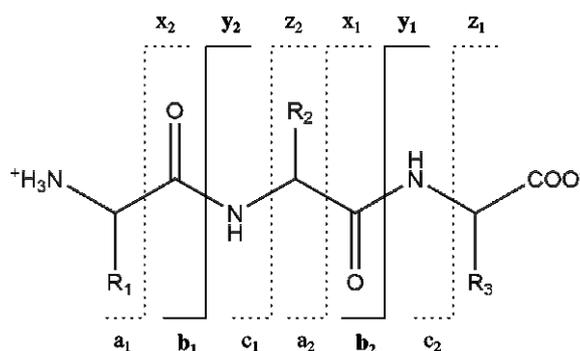


Abbildung 12: Nomenklatur der Peptidfragment-Ionen.¹⁴³

Die Ionen a, b und c enthalten das N-terminalen Peptidfragment und die x-, y- und z- Ionen den C-terminale Teil. Bei der HCD-Fragmentierung entstehen meist b- und y- Ionen. Diese wurden fett markiert.

3. MATERIAL UND METHODEN

3.1. Verwendete Programme für die Computerberechnungsmethoden

Programm	Version	Hersteller	Referenz
AMBER	12	University of California, San Francisco und Rutgers University	114
Ambertools	14	University of California, San Francisco und Rutgers University	114
Bio3D	2.2	freies Softwarepaket, University of Michigan	146
gnuplot	4.4	freie Software (Autoren: Thomas Williams, Colin Kelley, Russell Lang, Dave Kotz, John Campbell, Gershon Elber, Alexander Woo)	152
GROMACS	4.5.4- 4.6.2	freie Software (Projektleiter: Erik Lindahl, David van der Spoel, Berk Hess)	115
Jalview	2.8	The Barton Group, University of Dundee, Scotland, UK	153
MOE	2010/214	Chemical Computing Group	150
MS Office	2010	Microsoft	
ProDy	1.5.1	freie Software, University of Pittsburgh	149
PyMol	1.5.0.4	Schrödinger, LLC	147
Python	3.5	freie Software	
R	3.1	freie Software (Entwickelt von Bell Labs)	145
StavroX	3.4.5	freie Software (Autor: Michael Götze)	127
SYBYL	8.1	Tripos International	151
THESEUS	2.0.1	freie Software, Brandeis University	84
VMD	1.9.1	freie Software, University of Illinois	148
XWalk	Server	freier Server	154

3.2. Primär- und Sekundärstrukturanalyse von NS5B

Die Analyse der Primärstruktur von HCV NS5B erfolgte in Jalview mit den Sequenzen der UniProt-Datenbank (P27958, P26664, O92972, P26663, Q9WMX2, Q99IB8, P26660 und P26661). Das Sequenz-Alignment und die Sequenzeinfärbung wurden unter Verwendung der BLOSUM62-Matrix durchgeführt. Die Einfärbung der Sequenzen nach der Hydrophobizität erfolgte in Jalview nach Kyte *et al.*¹⁵⁵. Die Sekundärstrukturmerkmale wurden mittels R und dem Bio3D-Paket aus den PDB-Einträgen von NS5B entnommen und graphisch mit gnuplot dargestellt.

3.3. Analyse der Kristallstrukturen von NS5B

Die Analyse der Kristallstrukturen von HCV NS5B erfolgte anhand des jeweiligen Genotyps und seines entsprechenden UniProt¹⁵⁶-Eintrags, der Auflösung, der Auflistung der gebundenen Substrate oder Inhibitoren (falls vorhanden), der tatsächlichen Aminosäureanzahl der Kette A im Kristall, dem Qmean6-Wert der Kette A, dem prozentualen Anteil der Aminosäuren im begünstigten Bereich des Ramachandran-Diagramms von Kette A und der Auflistung der fehlenden Aminosäureseitenketten im Kristall, wie sie in der jeweiligen PDB-Datei vermerkt wurden. Diese Auflistung ist im Anhang in Tabelle 7 zu finden. Um die Daten der Kette A zu erhalten, wurden die PDB-Dateien mittels einer Schleife bearbeitet:

```
for file in *.pdb; do
  grep '^ATOM .\{14\}A' $file > ${file:0:4}_A.pdb
done
```

3.3.1. Superpositionierung mit THESEUS

Die strukturelle Analyse erfolgte mit dem Programm THESEUS^{84,84}. Dabei wurden wie bei der allgemeinen Kristallstrukturanalyse nur die A-Ketten der NS5B-Kristallstrukturen herausgeschrieben und für die THESEUS-Analyse zu Grunde gelegt. THESEUS verwendet zur Übereinanderlagerung der Strukturen mit unterschiedlichen Sequenzen den *Maximum-Likelihood* (ML)-Algorithmus in Kombination mit dem Sequenzalignment der zu untersuchenden Strukturen.⁸⁴ Für das Sequenzalignment wurden FASTA-Dateien (mit der Endung .fst) verwendet, die von Theseus mittels der folgenden Eingabe aus der jeweiligen PDB-Datei erstellt wurden:

```
theseus -F *_A.pdb
```

Die zusammengeführten FASTA-Dateien (cat *.fst > muscle_input) wurden anschließend mit dem Server „MUSCLE“ (*Multiple Sequence Comparison by Log-Expectation*) aligniert.^{157,158} Die MUSCLE-Ausgabe wurde im ClustalW-Format gespeichert und mit dem Namen muscle_output.clw versehen. Die anschließende Eingabe für die Alignierung der Strukturen in Theseus lautete:

```
theseus -A muscle_output.clw *_A.pdb > results_theseus
```

3.3.2. Das Programm R

Die Berechnung der RMSD-Werte (*root mean square deviation*) der Atompositionen und die Hauptkomponentenanalyse (*Principal Component Analysis*, PCA) der Kristallstrukturen erfolgte mit dem Programm R¹⁴⁵ unter der Verwendung der Pakete Bio3D¹⁴⁶, cIValid¹²² und muscle¹⁵⁷. Im Skript I des Anhangs ist die Abfolge der Befehle dargestellt. Zuerst erfolgte eine BLAST-Suche mit einer beliebigen NS5B-Kristallstruktur (in diesem Fall 2XXD). Anschließend wurden die passenden Strukturen von der PDB-Datenbank heruntergeladen. Um die verschiedenen Strukturen übereinander zu lagern, wurden die gemeinsamen Kernregionen aller Strukturen ermittelt. Aus dem Alignment wurden die RMSD-Werte berechnet, geclustert und als Dendrogramm dargestellt. Die Auswahl der Clusternummer erfolgte mittels Dunn-Index und Silhouettenkoeffizienten (vgl. Kapitel 2.6.1).

Anschließend erfolgte die eigentliche PCA. Dazu wurden die Lücken im Alignment identifiziert und für die Berechnung der Komponenten ausgelassen. Die Komponentenanalyse der Kristallstrukturen wurde geclustert und als Punktediagramm dargestellt. Die Auswahl der Clusternummer erfolgte ebenfalls mittels Dunn-Index und Silhouettenkoeffizienten.

3.4. Moleküldynamik-Simulation

Die Moleküldynamik-Simulationen wurden mit der Programm-Implementation pmemd aus AMBER12¹¹⁴ durchgeführt. Die AMBER-Kraftfelder ff99, ff99SB, ff12, ff12SB und ff03 wurden zuerst für die Testläufe der Dynamiken und das Kraftfeld ff03 für die eigentlichen Simulationen verwendet. Die Berechnungen wurden im NVT Ensemble mit implizitem Lösungsmittel durchgeführt. Daher wurden dem System keine Wassermoleküle oder Gegenionen zugefügt. Die Salzkonzentration des Systems wurde in Anlehnung an die Experimente von Reich *et al.*⁶⁷ bei 50 mM festgelegt. Vor der Moleküldynamik-Simulation wurde die Minimierung des Systems vorgenommen. Hierzu wurden 500 Schritte mit dem *steepest descent* und 4500 Schritte mit dem *conjugate gradient*-Algorithmus durchgeführt. Zu Beginn der Dynamik wurde das Protein schrittweise auf 300 K erhitzt. Dabei wurde die Temperatur alle 10 ps um 50 K erhöht. Die Schrittlänge während der Aufheizphase betrug 0,5 fs. Um die Proteinstruktur zu stabilisieren, wurde eine Kraftkonstante von 5 kcal/mol auf das Rückgrat von NS5B gelegt. Nach Erreichen der Temperatur von 300 K erfolgte die Äquilibration des Proteins und die schrittweise Rücknahme der Kraftkonstanten auf 2, 0,5 und 0,1 kcal/mol pro 100 ps. Zusätzlich erfolgte eine MD-Simulation von 400 ps Länge ohne Kraft-Beschränkungen des Proteins zur vollständigen Äquilibration des Systems. Die

Schrittlänge bei dieser Äquilibration betrug 1 fs. Bis zu diesem Zeitpunkt betrug die Dynamikzeit 760 ps. Für die Produktionsphase wurde eine Schrittlänge von 2 fs verwendet. In den Simulationen wurde der SHAKE-Algorithmus^{100,101} auf alle C-H-Bindungen angewandt. Die Temperaturregulierung erfolgte mit dem Berendsen-Thermostat. Der Grenzwert der nichtbindenden Wechselwirkungen für Berechnungen mit einem generalisiertem Born-Model sollte möglichst hoch sein (effektiv unendlich¹⁵⁹) und wurde auf 1000 Å gesetzt. Die Berechnungsweite der effektiven Bornradii wurde auf 16 Å eingestellt. Für alle weiteren Bedingungen wurden die Standardeinstellungen von AMBER gewählt. Eine detaillierte Auflistung der Einstellungen der einzelnen Dynamikabschnitte ist im Anhang in Tabelle 1 zu finden.

Die Liganden eines Proteins (vor allem Inhibitoren) sind meist kleine organische Moleküle, die andere Atomtypen besitzen, welche in den Kraftfeldern für Proteine nicht vorgesehen sind. Daher wurden sie mit dem *general AMBER force field* (GAFF)¹⁶⁰ berechnet, welches für kleine organische Moleküle entwickelt wurde. Die Parametrisierung dieser Liganden erfolgte mit Antechamber.¹⁶¹ Die Parameter für GTP wurden von Meagher *et al.*¹⁶² übernommen.

3.5. Auswertung der Moleküldynamiken

Die Auswertung der MDs, die in AMBER durchgeführt wurden, erfolgte mit Ambertools bzw. cpptraj. Die Diagramme wurden mit gnuplot und die Abbildungen der Proteine mit PyMol erstellt. Die PCAs der MDs wurden in VMD mit dem ProDy Zusatzpaket angefertigt.

Die Berechnung der RMSD- und RMSF-Werte erfolgte anhand des Proteinrückgrats der Atome N, C_α und C, wobei für die RMSF-Berechnung die letzten 100 ns der MD verwendet wurden. Dabei wurde angenommen, dass sich das System während dieser Zeitspanne in einem thermodynamischen Gleichgewicht der MD befindet. Die PCA wurde für die C_α-Atome der jeweiligen MD durchgeführt. Die Berechnung der Abstände und Winkel erfolgte mit cpptraj anhand der Massenschwerpunkte oder C_α-Atome. Für die Ermittlung des Interdomänen-Winkels und für die Abstandsmessung des RNA-Tunnels wurden die Vorgaben aus Davis *et al.*¹⁶³ entnommen. Die Domäneneinteilungen entsprachen für die Finger-Domäne den Aminosäuren 1-188 und 227-287, für die Hand-Domäne 189-226 und 288-370 und für die Daumen-Domäne 371-530. Der Durchmesser des RNA-Tunnels wurde anhand des Abstandes zwischen Met139 und Val405 bemessen (vgl. Abbildung 13).

Die Durchschnittsstrukturen wurden mit cpptraj aus den letzten 100 ns der jeweiligen MD-Simulation berechnet. Anschließend erfolgte eine Energieminimierung. Der Vergleich der

Durchschnittsstrukturen und die Berechnung des RMSD-Wertes pro Aminosäure wurden in MOE durchgeführt.

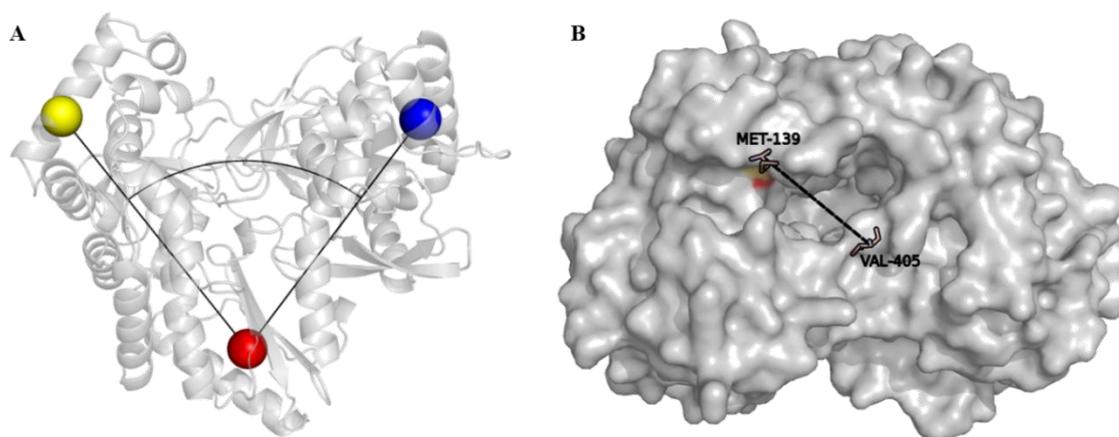


Abbildung 13: Berechnung des Interdomänen-Winkels und des RNA-Tunnel-Durchmessers.

A) Die Massenzentren der Domänen sind als Kugeln dargestellt und für die Finger-Domäne gelb, für die Hand-Domäne rot und für die Daumen-Domäne blau eingefärbt. B) Darstellung des Abstands zwischen Met139 und Val405 mit Blick auf den RNA-Tunnel.

3.6. Membrananker- und Membranmodellierung

Die Sekundärstrukturanalyse des Membranankers erfolgte mit den Servern Porter¹⁶⁴, Robetta¹⁶⁵, I-TASSER¹⁶⁶, LOMETS¹⁶⁷, QUARK¹⁶⁸, HMMTOP¹⁶⁹ und InterPro¹⁷⁰. Die Vorhersage des Startpunkts der Membranhelix schwankte dabei zwischen den Aminosäuren 564 und 570. Für die Modellierung des Membranankers wurden verschiedene Modelling-Server verwendet, wie SwissModel¹⁷¹, I-TASSER, QUARK, LOMETS, Robetta und MUSTER¹⁷². Die Modelle des Muster-Servers (MUlti-Source ThreadER) stellten sich dabei als am geeignetsten heraus. Die Modellierung der gesamten Polymerase (mit Membrananker) erfolgte per Hand in SYBYL. Eine anschließende kurze Minimierung des Systems (1000 Schritte mit *conjugate gradient*) erfolgte mit AMBER und dem AMBER-Kraftfeld 03.

Die Erstellung des Membrananker-Membran-Systems erfolgte mit dem HMMM Builder¹⁷³. Dabei wurden eine rechteckige Box und 37 Wassermoleküle pro Lipidmolekül verwendet. Die Zusammensetzung der Lipiddoppelschicht bestand aus 53 % Phosphatidylcholin, 20 % Phosphatidylethanolamin, 4 % Sphingomyelin und 13 % Phosphatidylinositol und Phosphatidylserin, was in etwa der ER-Membran einer Rattenleber entspricht¹⁷⁴. Die Ionenkonzentration wurde auf 0,1 M MgCl eingestellt und die entsprechenden Ionen zufällig im System verteilt.

3.7. Crosslinking

3.7.1. Verwendete Chemikalien und Hersteller

Acetonitril (CH ₃ CN)	VWR
Ammoniumhydrogencarbonat (NH ₄ HCO ₃)	Sigma-Aldrich
Coomassie-Brilliant-Blue G250	Sigma-Aldrich
d ₀ -Bis(sulfosuccinimidyl)glutarat (d ₀ -BS ² G)	Thermo Fisher Scientific
d ₄ -Bis(sulfosuccinimidyl)glutarat (d ₄ -BS ² G)	Thermo Fisher Scientific
DL-Dithiothreitol (DTT)	AppliChem
Dimethylsulfoxid (DMSO) (CH ₃ SOCH ₃)	Thermo Fisher Scientific
Essigsäure (CH ₃ COOH)	Roth
Inhibitor HCV-796	Zur Verfügung gestellt durch die
[5-Cyclopropyl-2-(4-fluorphenyl)-6-[(2- hydroxyethyl)(methylsulfonyl)amino]- N-methyl-1-benzofuran-3- carboxamid]	Abteilung Mikrobielle
2-(4-(2-Hydroxyethyl)-1-piperazinyl)-ethansulfonsäure (HEPES)	Biotechnologie der Universität
Iodacetamid (ICH ₂ CONH ₂)	Halle-Wittenberg
Ladepuffer (Leammi)	AppliChem
Marker PageRuler™ Prestained Protein Ladder	AppliChem
Methanol (CH ₃ OH)	Bio-Rad
N,N,N',N'-Tetramethylethyldiamin (TEMED)	Thermo Fisher Scientific
Nukleosidtriphosphat (NTP)	Roth
ssRNA (5`-CUAAGAUGCUCGCUGC-3`)	Roth
Trifluoressigsäure (TFA) (CF ₃ COOH)	Fermentas
Trypsin	IBA GmbH, Göttingen
	Merk
	Promega

3.7.2. Liste der Laborgeräte und Hersteller

Fusion-Tribrid-MS	Thermo Fisher Scientific
Nano-HPLC (UltiMate 3000)	Dionex
Geldokumentationssystem Gel Doc XR	Bio-Rad
Vakuumkonzentrator miVac	GeneVac
Wasseraufbereitungsanlage DirectQ5	Millipore

3.7.3. Proteinaufreinigung

Die RNA-abhängige RNA-Polymerase NS5B von HCV wurde von Dr. Ralph Golbik aus der Abteilung Mikrobielle Biotechnologie der Universität Halle-Wittenberg zur Verfügung

gestellt. Die Aufreinigungsprozedur erfolgte wie in der Dissertation von Stephan Reich⁸³ beschrieben.

3.7.4. Chemisches Vernetzen (*Crosslinking*)

Für die *Crosslink*-Experimente wurden pro Ansatz 4 μM NS5B eingesetzt. Das Protein wurde im Apo-Zustand und nach Zugabe von NTP (ATP, GTP, CTP, 500 μM), ssRNA (4 μM) oder Inhibitor HCV-796 (500 μM) vernetzt. Vor der Quervernetzung wurden die Reaktionsansätze für 15 Minuten bei Raumtemperatur (RT) inkubiert, um die Ausbildung des Protein-Substrat-Komplexes zu gewährleisten. Um eine ausreichende Reaktion zu erhalten, wurde der *Crosslinker* BS²G (d_0/d_4 1:1 (v/v)) in 200-fachem Überschuss (800 μM) zum Reaktionsansatz hinzugegeben. Für die Vernetzung wurde der *Crosslinker* frisch in DMSO gelöst und als 25 mM Stammlösung hergestellt. Nach der Zugabe von BS²G wurden die Reaktionsansätze für 30 Minuten bei RT inkubiert. Um die Quervernetzung anschließend abzustoppen, erfolgte die Zugabe von 20 mM Ammoniumhydrogencarbonat (NH_4HCO_3). Die einzelnen *Crosslink*-Ansätze wurden auf eine SDS-PAGE aufgetragen und elektrophoretisch getrennt.

3.7.5. Quantitatives *Crosslinking*

Die quantitativen *Crosslink*-Experimente erfolgten in Anlehnung an das normale *Crosslinking* mit dem Unterschied, dass hier der nicht-deuterierte *Crosslinker* (d_0) nur zum Apoprotein hinzugegeben wurde (entspricht d_0 -BS²G + NS5B) und der deuterierte *Crosslinker* (d_4) zu einem vorinkubiertem Protein-Substrat-Komplex (entspricht d_4 -BS²G + NS5B + NTP/ssRNA/HCV-796) oder nur zu NS5B (entspricht d_4 -BS²G + NS5B). Nach der *Crosslinker*-Inkubationszeit von 30 Minuten bei RT wurden die einzelnen Ansätze ebenfalls wieder mit Ammoniumhydrogencarbonat (NH_4HCO_3) abgestoppt. Nun erfolgte die Mischung des Apo-Reaktionsansatzes und nicht-deuterierten *Crosslinker* mit jeweils einem Komplexansatz bzw. dem Proteinansatz mit deuteriertem BS²G im Verhältnis 1:1 (v/v), wie in Abbildung 14 gezeigt.

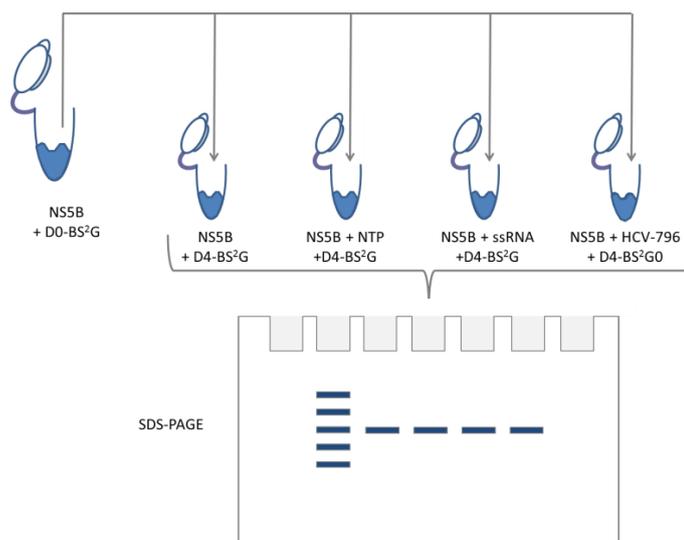


Abbildung 14: Vereinfachter Überblick zum quantitativen *Crosslink*-Experiment. Dargestellt ist die Mischung von nicht-deuterierter (d_0) *Crosslinker*-Protein-Lösung mit deuterierter (d_4) *Crosslinker*-Protein-Substrat-Lösung, welches der wesentliche Unterschied zum normalen *Crosslink*-Experiment ist. Anschließend werden die vereinigten Lösungen auf eine SDS-PAGE aufgetragen und elektrophoretisch getrennt.

3.7.6. SDS-PAGE (*sodium dodecyl sulfate polyacrylamide gel electrophoresis*)

Die 10 %-ige SDS-PAGE für die Auftrennung der *Crosslink*-Protein-Mischungen wurde bereits fertig von der Abteilung Mikrobielle Biotechnologie zur Verfügung gestellt. Für die elektrophoretische Auftrennung der Proteinproben wurden diese mit Ladepuffer (2x) vermischt, 10 Minuten bei 95 °C erhitzt und anschließend abzentrifugiert. Maximal 50 μ l jeder Mischung wurden in eine Probentasche geladen. Die gelelektrophoretische Trennung erfolgte für zwei Gele bei 200 V und 70 mA über 60 Minuten.

3.7.7. Gelfärbung nach Neuhoff

Nach der Proteinauftrennung wurden die SDS-PAGE in der Fixierlösung für eine Stunde inkubiert, anschließend mit MilliQ-Wasser gewaschen und mit der Färbelösung über Nacht inkubiert. Die Entfärbung erfolgte durch Waschen mit bidest. Wasser (MilliQ-Wasser). Die Färbelösung musste 4 Stunden vor der Verwendung hergestellt werden.

Fixierlösung: 40 % (v/v) Methanol, 10 % (v/v) Essigsäure

Färbelösung: 2 % (v/v) Lösung A, 98 % (v/v) Lösung B

Lösung A: 5 % (w/v) Coomassie-Brilliant-Blue G250 in MilliQ-Wasser

Lösung B: 2 % (w/v) Phosphorsäure, 20 % (w/v) Ammoniumsulfat-Lösung ($(\text{NH}_4)_2\text{SO}_4$ in MilliQ-Wasser)

3.7.8. In-Gel Proteinspaltung

Um das vernetzte Protein optimal spalten zu können, wurden die entsprechenden Gelbanden ausgeschnitten und in ca. 1 mm³ große Würfel zerteilt. Diese Gelwürfel wurden in einem 1,5 ml Eppendorfgefäß mit 500 µl Acetonitril für 10 Minuten bei RT gewaschen. Der Überstand wurde verworfen. Anschließend erfolgte die Reduktion der Disulfidbrücken des Proteins mit 100 µl einer Reduktions-Lösung bestehend aus 10 mM DTT, gelöst in einer frisch hergestellten 100 mM Ammoniumhydrogencarbonat-Lösung. Die Gelwürfel-Lösung wurde für 30 Minuten bei 56 °C geschüttelt. Danach erfolgte die Zugabe von 1 ml Acetonitril, 10 Minuten schütteln bei RT und die Verwerfung des Überstands. Die Alkylierung der freien Cysteinseitenketten mittels 100 µl einer Iodacetamid-Lösung (55 mM Iodacetamid in 100 mM Ammoniumhydrogencarbonat-Lösung) erfolgte für 30 Minuten bei RT im Dunkeln. Anschließend wurden die Gelstücke wieder mit 1 ml Acetonitril für 10 min bei RT gewaschen und der Überstand verworfen. Um die Gelwürfel weiter zu entfärben, wurden 100 µl einer 100 mM Ammoniumhydrogencarbonat- und Acetonitril-Lösung im Verhältnis 1:1 (v/v) hinzugegeben und für 30 Minuten bei RT geschüttelt. Danach erfolgten das Waschen mit 500 µl Acetonitril für 10 Minuten bei RT und das Entfernen des Überstandes.

Die Gelwürfel wurden mit 40 µl Trypsin-Lösung für 2 Stunden im Kühlschrank aufbewahrt, um eine ausreichende Diffusion der Lösung in das Gel zu gewährleisten. Um die Würfel vor dem Austrocknen zu schützen, wurden 40 µl einer 25 mM Ammoniumhydrogencarbonat-Lösung zugegeben. Anschließend wurde die Gellösung bei 37 °C über Nacht inkubiert.

Die Extraktion der gespaltenen Proteine erfolgte mit 80 µl einer Lösung bestehend aus einer 5 %-igen TFA- und Acetonitril-Lösung im Verhältnis 1:2 (v/v). Dieses Gemisch wurde für 15 Minuten bei 37 °C geschüttelt und der Überstand in ein neues 1,5 ml Eppendorfgefäß überführt. Die Extraktion wurde noch zweimal mit den Gelstücken wiederholt. Der gesammelte Überstand wurde im Vakuumkonzentrator auf 10-20 µl eingengt und anschließend für die Massenspektrometrie verwendet.

3.7.9. Orbitrap Fusion Tribrid Massenspektrometer

Die *Crosslink*-Ansätze wurden mit einer 0,1%igen TFA-Lösung auf 40 µl aufgefüllt. Je 38 µl wurden für die Messungen im Massenspektrometer verwendet. Die Peptidfragmente wurden über die Nano-HPLC fraktioniert.

Auf der Vorsäule des HPLC-Systems (C18, 300 μm x 5 mm, 3 μm Partikelgröße) erfolgte die Entsalzung des Peptidgemisches für 15 Minuten bei 60 °C mit 2 % ACN und 0,1 % TFA und einer Fließgeschwindigkeit von 30 $\mu\text{l}/\text{min}$. Die Fraktionierung der Peptid-Lösung erfolgte auf einer Trennsäule (C18, 20 cm x 75 mm, 1,9 μm Partikelgröße) mittels eines Lösungsmittelgradienten für 60 Minuten bei 50 °C. Die Zusammensetzung des Gradienten erfolgte mittels zweier Fließmittel, A: 0,1 % TFA in H_2O und B: 0,08 % TFA in ACN, und betrug für die Messzeit 3 %-35 % an Fließmittel B. Die Nano-HPLC ist direkt an die Nano-ESI-Quelle des Massenspektrometers gekoppelt. Die Massenspektren wurden in einem Massenbereich von m/z 300-1500 aufgenommen. Die fünf höchsten Signale wurden mittels HCD fragmentiert (Fragmentierungsenergie 27 % - 33 %) und in der Orbitrap analysiert. Das Isolationsfenster betrug dabei 2 Th. Gemessene Vorläuferionen wurden für die nächsten 90 Sekunden aus der Analyse ausgeschlossen (*dynamic exclusion*).

3.5.10 Identifizierung der *Crosslink*-Produkte

Die Auswertung der Massenspektren erfolgte durch die Abteilung Pharmazeutische Chemie & Bioanalytik der Universität Halle-Wittenberg. Dabei wurden die MS-Spektren mittels Xcalibur 2.07 und StavroX 3.4.5 analysiert. Die Massenabweichung wurde auf 4 ppm gesetzt. Es durften maximal 3 verpasste Spaltungen durch Trypsin nach den Aminosäuren Arginin und Lysin auftreten. Das Verhältnis zwischen Signal und Rauschen musste höher als 2 sein. Die Seitenketten von Lysin, Serin, Threonin und Tyrosin wurden als potentielle Reaktionspartner von BS²G angesehen.¹³⁰ Die mit der Software identifizierten *Crosslinks* wurden manuell überprüft sowie deren d_0/d_4 -Verhältniss berechnet. Die erhaltenen *Crosslink*-Daten wurden anschließend anhand der C_α - C_α -Abstände in den Kristallstrukturen und den MD-Simulationen von NS5B miteinander verglichen.

4. ERGEBNISSE UND DISKUSSION

4.1. Analyse der Kristallstrukturen

In der Protein-Datenbank (*Protein Data Bank*, PDB) sind derzeit (Stand Dezember 2016) 189 Kristallstrukturen von HCV NS5B der Genotypen (GT) 1 und 2 hinterlegt. Um die Unterschiede dieser Kristallstrukturen zu analysieren, wurde eine Sequenzanalyse durchgeführt und die Sekundär- und Tertiärstrukturen untersucht. Die PDB-Einträge zu HCV NS5B wurden zunächst manuell analysiert. Anschließend erfolgte die Bewertung der Güte der Kristallstrukturen anhand einer Hauptkomponentenanalyse (*principal component analysis*, PCA) mit dem Programm R¹⁴⁵ und einer strukturellen Überlagerung mit dem Programm THESEUS⁸⁴. Eine vergleichbare Analyse wurde bereits von Caillet-Saguy *et al.*⁶⁶ mit dem Programm Escet^{175,176} durchgeführt. Der Vorteil der Analyse mit THESEUS und der Hauptkomponentenanalyse liegt in der Möglichkeit, alle Genotypen miteinander zu vergleichen, wohingegen mit Escet nur Strukturen mit genau derselben Sequenz analysiert werden konnten.

4.1.1. Primär- und Sekundärstrukturanalyse der NS5B-Kristallstrukturen

Die RNA-abhängige RNA-Polymerase von Hepatitis C besitzt keine *proofreading*-Funktion. Dies resultiert in einer *in vitro* Mutationsrate von $1,4 \times 10^{-6}$ bis $8,7 \times 10^{-3}$ falsch eingebauten Nukleotiden pro Position, wie *in vitro* gemessen wurde.^{9,10,177-179} Die Mutationsrate, welche aus *in vivo* Experimenten erhalten wurde, ist mit durchschnittlich $2,5 \times 10^{-5}$ pro Nukleotid pro repliziertem Genom vergleichbar.¹⁸⁰ Daher besitzt HCV eine große genetische Variabilität, welche sich ebenfalls in der Aminosäure-Sequenz widerspiegelt.

Die Kristallstrukturen von NS5B enthalten die Subtypen 1a, 1b, 2a und 2b. Insgesamt wurden bereits 8 Varianten von NS5B kristallisiert. Der Sequenzvergleich dieser Genotypvarianten von NS5B ist in Tabelle 2 dargestellt. Vor dem jeweiligen Subtyp ist die UniProt-ID vermerkt. Gleiche Subtypen (z.B. 1a-H und 1a-iso1) zeigen eine Aminosäure-Sequenzidentität von über 95 %, bei unterschiedlichen Subtypen eines Genotyps (z.B. zwischen GT-1a und GT-1b) liegen die Identitäten noch zwischen 85 % und 88 %. Die Sequenzidentität der Genotypen 1 und 2 liegt zwischen 73 % und 76 %.

Tabelle 2: Vergleich der Sequenzidentitäten der Genotypen und Subtypen der NS5B-Kristallstrukturen. Die UniProt-ID ist vor dem NS5B-Subtyp vermerkt. Die Aminosäureidentitäten sind in Prozent angegeben und wurden entsprechend ihrer Übereinstimmung farblich markiert. Daher entsprechen die grün markierten Zellen einer Sequenzidentität von über 90 %, blau markiert einer Sequenzidentität zwischen 80 % und 90 % und die orange gefärbten weisen eine Sequenzidentität von weniger als 80 % auf.

UniProt-ID Subtyp	P27958 1a, H77	P26664 1a-iso1	O92972 1b, HCJ4	P26663 1b, BK	Q9WMX2 1b-CON1	Q99IB8 2a, JFH1	P26660 2a, HCJ6	P26661 2b-HCJ8
P27958 1a, H77	100	97,63	85,45	87,14	86,80	73,27	73,27	74,45
P26664 1a-iso1	97,63	100	86,63	87,99	87,99	73,77	74,11	75,13
O92972 1b, HCJ4	85,45	86,63	100	95,94	95,43	74,45	75,13	74,96
P26663 1b, BK	87,14	87,99	95,94	100	96,28	74,11	74,96	74,62
Q9WMX2 1b-CON1	86,80	87,99	95,43	96,28	100	74,96	75,63	74,45
Q99IB8 2a, JFH1	73,27	73,77	74,45	74,11	74,96	100	95,60	87,31
P26660 2a, HCJ6	73,27	74,11	75,13	74,96	75,63	95,60	100	87,99
P26661 2b-HCJ8	74,45	75,13	74,96	74,62	74,45	87,31	87,99	100

In Abbildung 15 ist das Alignment der Primärsequenzen abgebildet, für welche Kristallstrukturen verfügbar sind. Die Motive A, C, D und E sind bei diesen Genotypen konserviert. Die Motive B, F und G sind weitestgehend konserviert, unterscheiden sich aber zwischen den Genotypen. Die für HCV NS5B typischen Strukturmerkmale wie der Δ 1-Loop, β -Loop und Linker sind in ihrer Aminosäure-Zusammensetzung variabler. Darüber hinaus existieren sehr konservierte Regionen zwischen den Aminosäureresten 6-10, 50-56, 118-129, 161-172, 190-201, 255-265, 274-282, 422-433 und 508-527. Im Anhang in Abbildung 67 ist das Sequenzalignment aller anerkannten (*reviewed*) HCV NS5B Sequenzen zum Vergleich abgebildet, welches diese Regionen ebenfalls als konserviert aufzeigt.

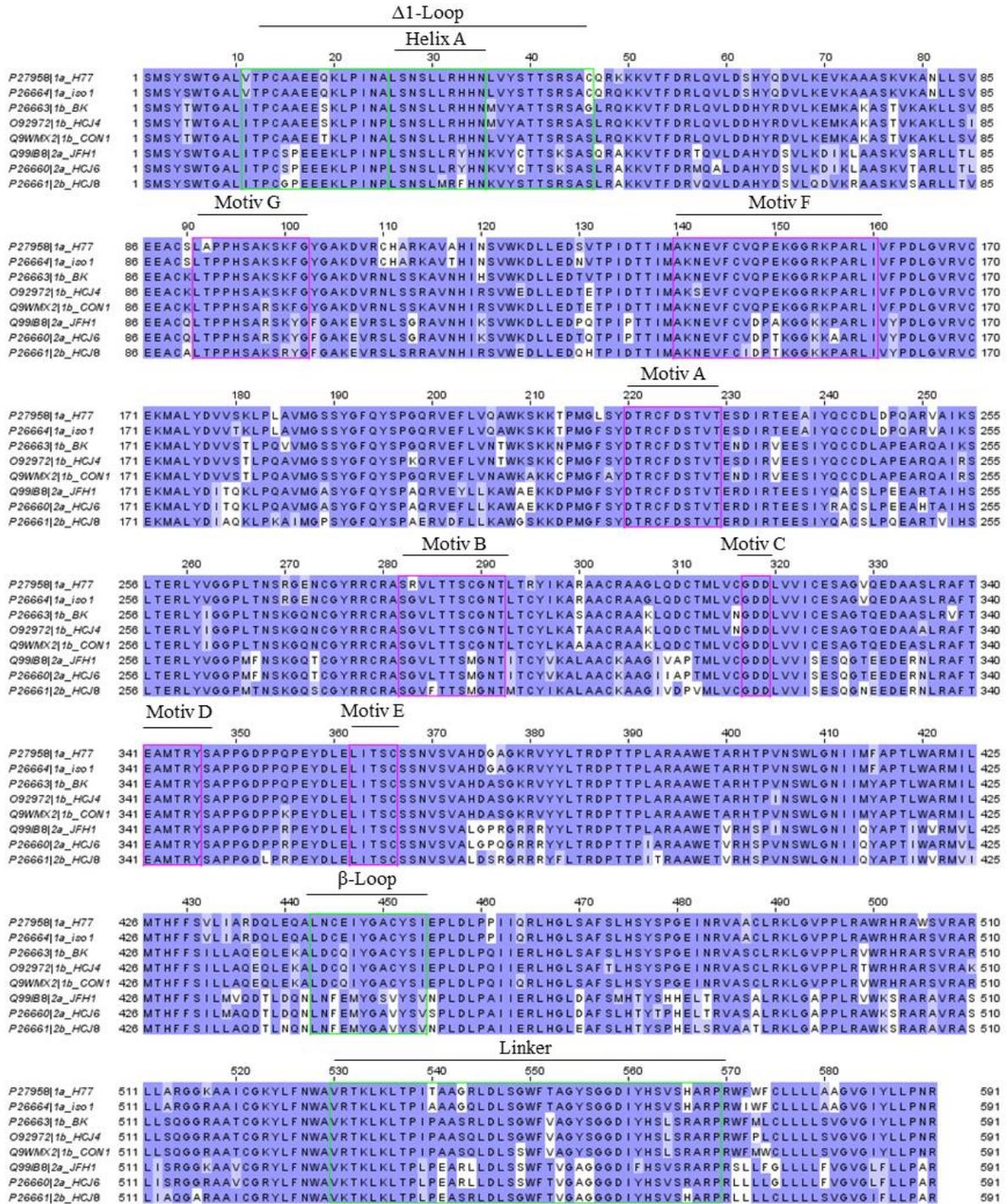


Abbildung 15: Sequenzalignment der Genotypen von NS5B, von denen eine Kristallstruktur existiert. Die Einfärbung erfolgte anhand der BLOSUM62-Matrix. Das Alignment wurde in Jalview erstellt.

Im Vergleich zum Sequenzalignment sind in Abbildung 16 die Sekundärstrukturelemente (SSE) der NS5B Kristallstrukturen prozentual dargestellt. Die Angaben dafür wurden aus den jeweiligen PDB-Einträgen entnommen. Die rote Linie in Abbildung 16 zeigt den Anteil der fehlenden Koordinaten in den Kristallstrukturen.

Die meisten SSE sind zu 100 % in den NS5B Kristallstrukturen vorhanden. Der Bereich der Helix A (AS 23-35) besitzt allerdings nur in 75 % aller Proteinstrukturen eine helikale Struktur. Teilweise war die Elektronendichte zu gering für eine Strukturaufklärung (10 %), teilweise existieren diese Regionen als *coil*. Im Bereich des Motivs F konnten zu 63 % die Koordinaten des Proteins nicht ermittelt werden. Dies spricht für einen sehr flexiblen Bereich in der Polymerase. Weiterhin scheint der Linker sehr unstrukturiert zu sein. Zu 70 % existiert eine α -Helix zwischen den Aminosäuren 540 und 545, andererseits fehlen fast 20 % der Daten in diesem Bereich aufgrund geringer Elektronendichten.

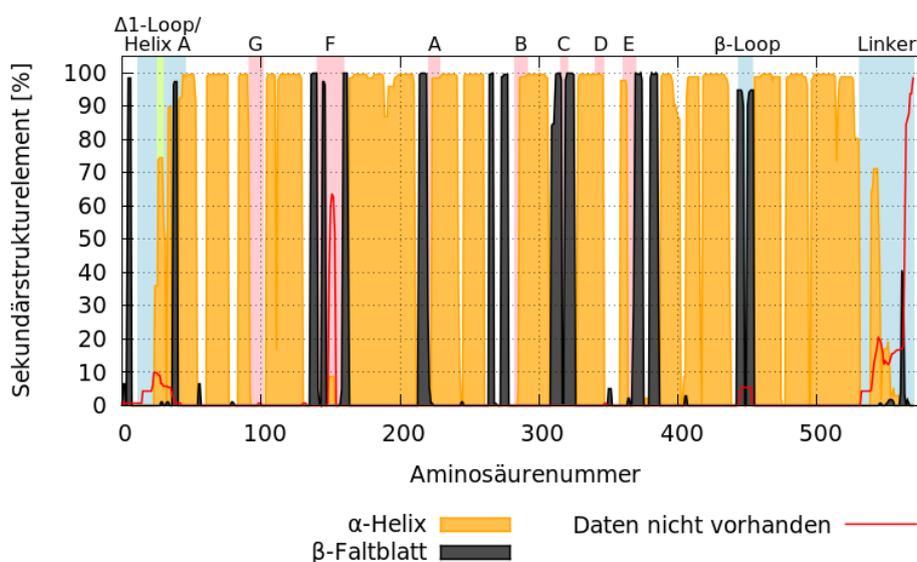


Abbildung 16: Sekundärstrukturelemente von NS5B in den 189 Kristallstrukturen. Die Häufigkeiten der α -Helices und der β -Faltblätter sind als Flächen aufgetragen. Der prozentuale Anteil der nicht vorhandenen Daten ist als rote Linie eingezeichnet.

4.1.2. Qualitätsbewertung der NS5B Kristallstrukturen

Um für nachfolgende Untersuchungen geeignete Startstrukturen zu identifizieren, wurden die verfügbaren Kristallstrukturen nach folgenden Kriterien bearbeitet:

- Es wurde die Kette A für die Analyse verwendet.
- Eventuell enthaltene Wassermoleküle, Liganden oder Ionen wurden gelöscht.
- Die Bewertung der 3D-Strukturen wurde mit Hilfe des QMEAN Servers und PROCHECK durchgeführt.
- Die Kristallstruktur sollte möglichst vollständig sein.
- Die Einteilung erfolgte anhand der Substrate, wie NTPs und RNAs, und anhand der Inhibitoren.

Die Analyse der PDB Strukturen ist im Anhang in Tabelle 7 aufgelistet. Zusätzlich wurden die fehlenden Aminosäurereste der Proteinstrukturen aus der PDB-Datei angegeben. Bei der Betrachtung der aufgelisteten Strukturen fällt auf, dass bei den meisten Kristallstrukturen die Aminosäuren 148 bis 152 fehlen, welche sich im Motiv F befinden. Dieses Motiv interagiert mit dem Triphosphatende des Nukleotids am NTP-Eingang der Polymerase.⁵⁷ Die fehlende Elektronendichte kann hierbei durch eine intrinsische Flexibilität des Proteinsegments erklärt werden.

Die Expressierung und anschließende Kristallisation von NS5B erfolgt meist mit zwei unterschiedlichen Deletionsvarianten, bei der entweder 21 ($\Delta 21$) oder 55 ($\Delta 55$) Aminosäuren des C-Terminus fehlen. Bei einer Gesamtlänge von 591 Aminosäuren resultiert damit entweder ein Protein mit der Länge von 570 oder 536 Bausteinen. Bei der $\Delta 55$ Variante ist jedoch zu beachten, dass dadurch auch die Linker-Domäne von NS5B verkürzt wird, die die Polymerase mit dem Membrananker verbindet und welche in den Kristallstrukturen in das aktive Zentrum der Kristallstrukturen hineinreicht. Da der Position des Linkers eine zentrale Rolle bei der Untersuchung des dynamischen Verhaltens von NS5B zukommt, wurde die $\Delta 55$ Variante für spätere MD-Simulationen nicht berücksichtigt.

Einen Ausreißer innerhalb der untersuchten Kristallstrukturen stellt der Eintrag 4NLD¹⁸¹ dar. Er ist mit 63 fehlenden Aminosäureresten (der N-Terminus beginnt erst mit Aminosäure Phe54) die am wenigsten aufgelöste Struktur und zeigt auch eine völlig andere Positionierung des Motivs F und des Linkers, der nicht in das aktive Zentrum hineinreicht. Des Weiteren zeigt die Proteinstruktur im Bereich von Motiv F viele Kontakte mit dem Symmetriemolekül, sodass z.B. in der Elektronendichte eine Disulfidbrücke zwischen Cys146 aus der Kette A und Cys146 aus dem Nachbarmolekül ersichtlich ist (siehe Abbildung 17). Diese Disulfidbrücke ist zwar nicht in der PDB-Datei verzeichnet, aber eindeutig anhand der Differenzelektronendichte zu identifizieren. Des Weiteren besitzt 4NLD die Raumgruppe I222, während die Mehrzahl der PDB-Einträge (145) die Raumgruppe $P2_12_12_1$ besitzt. Unterschiedliche Raumgruppen können zu unterschiedlichen Interaktionsflächen führen, welche in diesem Fall die Disulfidbrücke ermöglicht. Unter physiologischen Bedingungen ist diese Disulfidbrücke unwahrscheinlich, da NS5B als Monomer und auf das Zytosol ausgerichtet in der ER-Membran vorliegt. Aus diesen Gründen wurde der PDB-Eintrag 4NLD in allen weiteren Untersuchungen nicht berücksichtigt. Grundsätzlich kann jedoch 4NLD nicht als „falsche“ Kristallstruktur bezeichnet werden, da diese lediglich eine weitere mögliche Konformation aufweist, welche aber nicht biologisch aktiv sein muss.

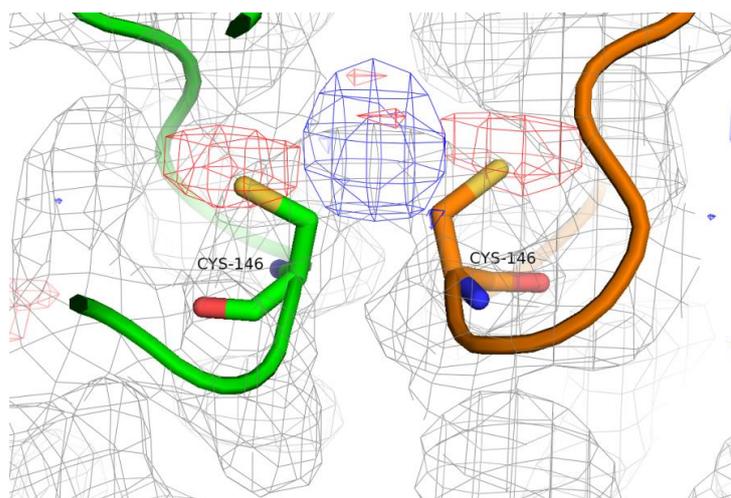


Abbildung 17: Darstellung der Struktur und der Elektronendichte von 4NLD mit dem Fokus auf Cys146. Die Kette A der PDB-Struktur ist grün gefärbt, das symmetrie-verwandte Nachbarmolekül in orange. Die $2F_o-F_c$ -Elektronendichte (1σ) der Proteinstrukturen ist als graues Netz abgebildet, die F_o-F_c -Differenzelektronendichte ($3\sigma/-3\sigma$) ist für die positive Differenz blau und für die negative Differenz rot dargestellt.

4.1.3. RMSD- und Hauptkomponentenanalyse der Kristallstrukturen

Nach der manuellen Qualitätsbewertung der Kristallstrukturen erfolgten eine RMSD- und eine Hauptkomponentenanalyse, um die Proteinmodelle anhand ihrer dreidimensionalen Struktur zu validieren. Dies erfolgte mit Hilfe des Programms R und dem „Bio3D“ Programmpaket¹⁴⁶. Für die Analysen wurden nur die A-Ketten der Kristallstrukturen verwendet.

Bei der RMSD-Analyse wurden die Kristallstrukturen anhand der Methode der kleinsten Fehlerquadrate überlagert und mit einer hierarchischen (distanzbasierten) Methode und dem *average linkage*-Algorithmus¹⁸² in 4 Gruppen geteilt. Die Auswahl der Gruppennummer erfolgte anhand des Dunn-Index und des Silhouettenkoeffizienten. Diese Gruppeneinteilung kann als Dendrogramm dargestellt werden (Abbildung 18). Hierbei verfügen Gruppierungen mit kleinen RMSD-Unterschieden über kurze Verzweigungen und Gruppen mit großen RMSD-Werten sind weiter verzweigt. Mit den gewählten Auswahlkriterien ergeben sich 4 Cluster, wobei die Mehrzahl der untersuchten Proteinmodelle die größte Gruppe (Gruppe 1) bildet (vgl. Abbildung 18). Die Gruppen 2 bis 4 enthalten 12 und zweimal 8 Kristallstrukturen. Die ersten beiden Gruppeneinteilungen sind unabhängig vom Genotyp. In Gruppe 3 und 4 sind jedoch nur Kristallstrukturen der Variante 2a, JFH1 enthalten. Andererseits wurden die Strukturen dieser Gruppen ohne den β -Loop oder mit einer dsRNA kristallisiert. Hier scheint die Gruppeneinteilung vielmehr durch konformative Unterschiede bedingt.

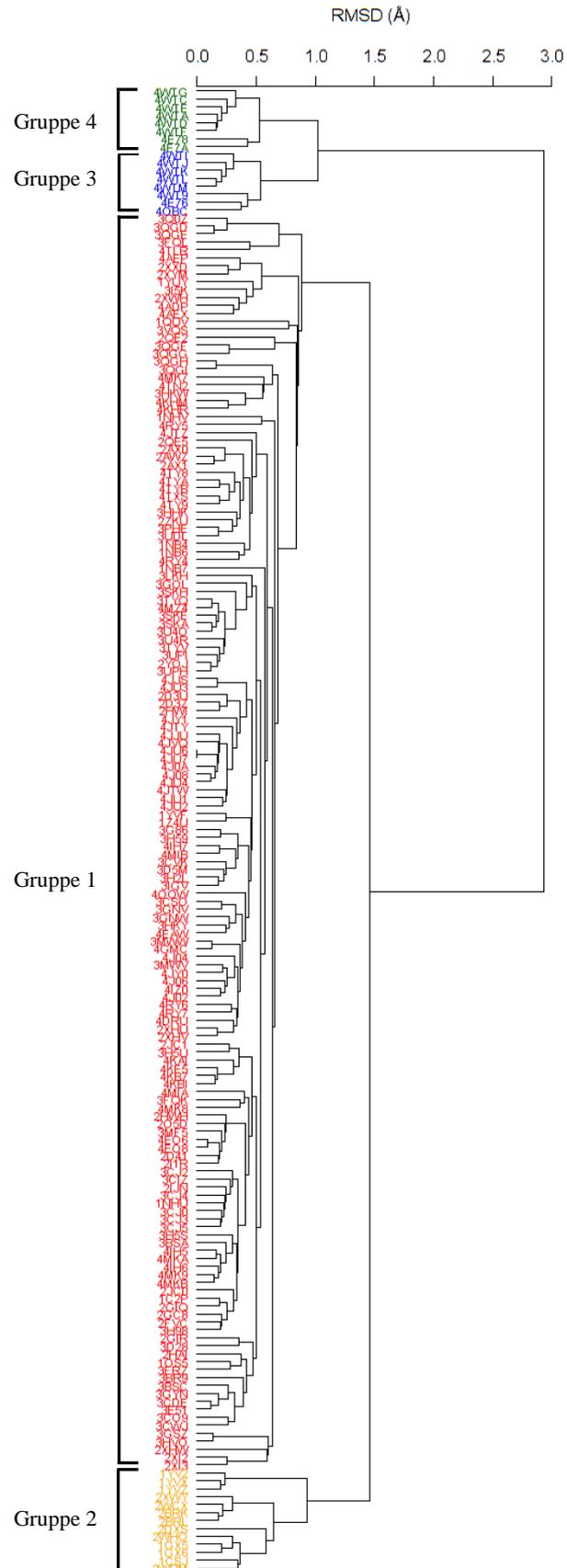


Abbildung 18: Dendrogramm der RMSD-Analyse der Kristallstrukturen von NS5B. Die 4 Clustergruppen sind farblich in rot (Gr. 1), orange (Gr. 2), blau (Gr. 3) und grün (Gr. 4) eingeteilt.

Die Strukturunterschiede der berechneten Gruppen sind in Abbildung 19 dargestellt. Besonders starke Positionsunterschiede existieren innerhalb der Daumen-Domäne mit dem β -Loop und im Bereich des Motivs F. Die Kristallstrukturen der Gruppe 4 zeigen einen etwas weiter geöffneten Zustand als die Strukturen der Gruppe 1 (siehe Abbildung 19A). Die Kristallstrukturen stimmen in den Bereichen der Finger- und der Hand-Domäne nahezu perfekt überein. Deutliche Unterschiede ergeben sich in der relativen Position der Daumendomäne, insbesondere in der Region der Daumenspitze zwischen den Aminosäuren 398-408, 430-440 und 510-520, wie in Abbildung 19B dargestellt. Dabei variieren die Abstände an den einzelnen Messpunkten (Aminosäuren 400, 404, 437, 514, 515 und 517) zwischen der Struktur aus Gruppe 1 und 4 unterschiedlich stark zwischen 6 Å und 11 Å.

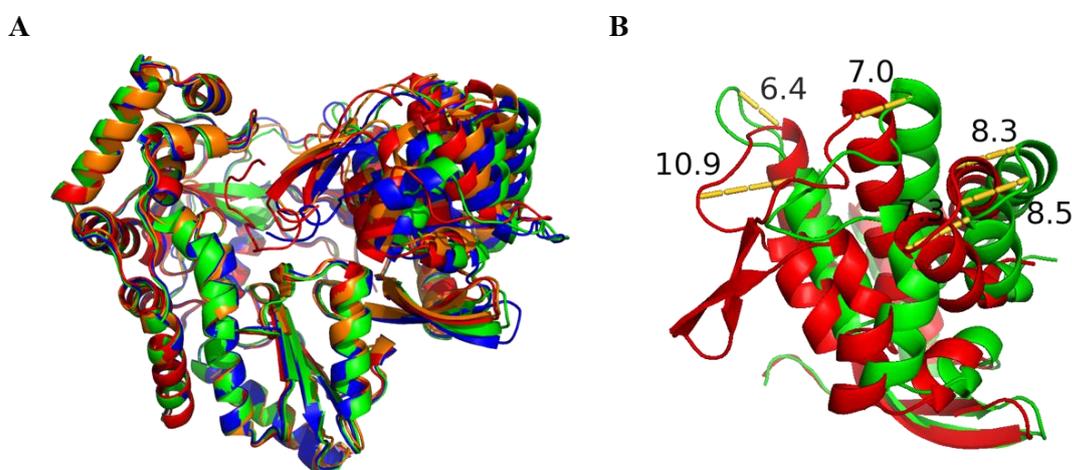


Abbildung 19: *Cartoon*-Darstellung repräsentativer Kristallstrukturen der RMSD-Einteilung in 4 Gruppen. **A)** Die PDB-Einträge 2QE2, 1GX5, 4OBC und 4E7A repräsentieren die Gruppen: Gr.1 – rot, Gr.2 – orange, Gr.3 – blau, Gr.4 – grün. **B)** Detailaufnahme der überlagerten Daumen-Domäne von 2QE2 (Gr. 1, rot) und 4E7A (Gr. 4, grün) und deren Abstände in Å. Die Messungen wurden an den C α -Atomen der Aminosäuren 400, 404, 437, 514, 515 und 517 vorgenommen.

Die Ergebnisse der Hauptkomponentenanalyse sind in Abbildung 20 zu sehen. Im Diagramm A sind die einzelnen Komponenten und deren Anteile an der Strukturvarianz dargestellt. Allein die erste Komponente (PC1) beschreibt 79,6 % aller Strukturunterschiede der Kristallstrukturen. In Diagramm B wurden die Komponenten 1 und 2 gegeneinander aufgetragen und hierarchisch in 4 Gruppen eingeteilt. Die Wahl der Clusteranzahl erfolgte ebenfalls mittels Dunn-Index und Silhouettenkoeffizienten. Die resultierende Einteilung der einzelnen PDB-Strukturen ist mit der RMSD-Einteilung identisch. Dabei zeigte sich eine große Clustergruppe von Strukturen um den Koordinatenursprung (Gruppe 1 in rot) und drei kleinere Cluster davon entfernt. Klar zu identifizierende Ausreißer sind bei dieser Auftragung

jedoch nicht zu erkennen. Nur die Kristallstruktur 2DXS liegt etwas abseits der zweiten Gruppe.

Die Kristallstruktur 2DXS wurde mit dem Genotyp 1b (*isolate BK*) und einem tetrazyklischem NNI-1 Inhibitor co-kristallisiert. Dieser Struktur fehlen viele Aminosäuren im Bereich des Δ 1-Loops (15-36) und des Motivs F (148-152). Die Linker-Domäne reicht nur bis Aminosäure 544 und zeigt eine etwas veränderte Position im Vergleich zu anderen Kristallstrukturen.

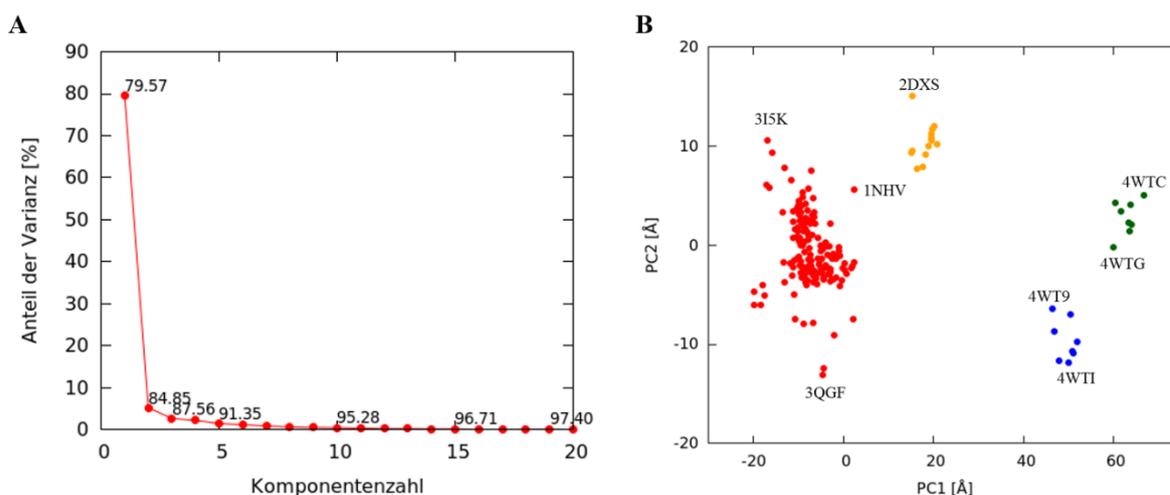


Abbildung 20: PCA der Kristallstrukturen mit Bio3D in R.

A) Die Eigenwerte der einzelnen Komponenten sind in dem Diagramm gegen ihren Varianzanteil der Struktur aufgetragen. Die erste Komponente (PC1) beinhaltet demnach 79,6 % der Strukturunterschiede der Kristallstrukturen. **B)** Auftragung von PC1 gegen PC2. Die Komponenten wurden hierarchisch in 4 Gruppen geteilt und entsprechend gefärbt: Gr.1 – rot, Gr.2 – orange, Gr.3 – blau, Gr.4 – grün.

Ein interessantes Ergebnis von RMSD- und Hauptkomponenten-Analyse ist die Zuordnung aller 9 Kristallstrukturen der Δ 55-Variante von NS5B (530 bis 532 AS) in Gruppe 2 (orange). Die restlichen 4 Strukturen der Gruppe 2 sind 544 bis 548 Aminosäuren lang. Die Primärsequenzen der Kristallstrukturen der vierten Gruppe (grün) sind ebenfalls kurz und enden nach 540 bis 545 Aminosäuren. Zudem wurden alle Strukturen der Gruppe 4 mit einer dsRNA co-kristallisiert und besitzen eine Δ 8-Deletion im Bereich des β -Loops. Über Gruppe 3 (blau) kann jedoch keine einheitliche Aussage bezüglich besonderer Strukturmerkmale des Proteins getroffen werden.

Die RMSD- und die Hauptkomponentenanalyse der NS5B-Kristallstrukturen können zwischen geschlossenen und weiter geöffneten Strukturen unterscheiden. Anscheinend spielt die Deletion des β -Loops sowie die Länge und Position des Linkers eine entscheidende Rolle für die Konformation der kristallisierten Polymerase. Schon geringe Abweichungen innerhalb der Rotation der Daumen-Domäne können so identifiziert werden. Diese Resultate stellen die

korrekte Anordnung des Linkers in den Kristallstrukturen in Frage, wenn man bedenkt, dass bei Fehlen des Linkers die geöffnete und damit prozessive Form der Polymerase bevorzugt wird. Somit könnte der Linker im nativen Zustand des Proteins eine andere Position aufweisen, als es aus der Kristallstruktur ersichtlich ist.

4.1.4. THESEUS-Analyse der Genotypen

Das Programm THESEUS überlagert die Kristallstrukturen anhand des ML-Algorithmus, der im Gegensatz zur Methode der kleinsten Fehlerquadrate gut übereinstimmende Strukturabschnitte stärker wichtet als variable Bereiche. Aus der Überlagerung wird anschließend der RMSD-Wert berechnet. Bei dieser Analyse wurden Kristallstrukturen von HCV NS5B mit Substraten oder Inhibitoren nicht berücksichtigt, um mögliche konformationelle Beeinflussungen auszuschließen. Die Apoprotein-Kristallstrukturen 4E76 und 4WTR weisen eine Deletion des β -Loops auf und wurden daher weggelassen. Daraus ergaben sich insgesamt 23 Apoprotein-Strukturen von NS5B, von denen eine Kristallstruktur den Genotyp 1a besitzt, 12 den Genotyp 1b, 9 den Genotyp 2a und eine den Genotyp 2b (vgl. die Protein-Strukturen im Anhang in Tabelle 7).

Die Superpositionierung der Apoprotein-Strukturen mit dem ML-Algorithmus zeigt größere Strukturvarianten in den Bereichen des $\Delta 1$ -Loops und der Helix A (11-46 bzw. 25-35), sowie im Bereich von Motiv F (140-160), der Daumen-Domäne sowie des Linkers (359-529 bzw. 530-570) (vgl. Abbildung 21 und Abbildung 22). Diese Strukturvarianzen sind sehr gut mit den Ergebnissen der RMSD- und der Hauptkomponenten-Analyse vergleichbar, welche ebenfalls im Bereich der Daumen-Domäne und des Motivs F große strukturelle Unterschiede aufgezeigt haben. Die größten Varianzen zeigt der Genotyp 2a insbesondere im Bereich des Daumens. Dies deutet auf eine große Flexibilität innerhalb dieses Bereichs. Im Gegensatz dazu zeigt der Genotyp 1b eine geringe Varianz zwischen seinen Proteinstrukturen.

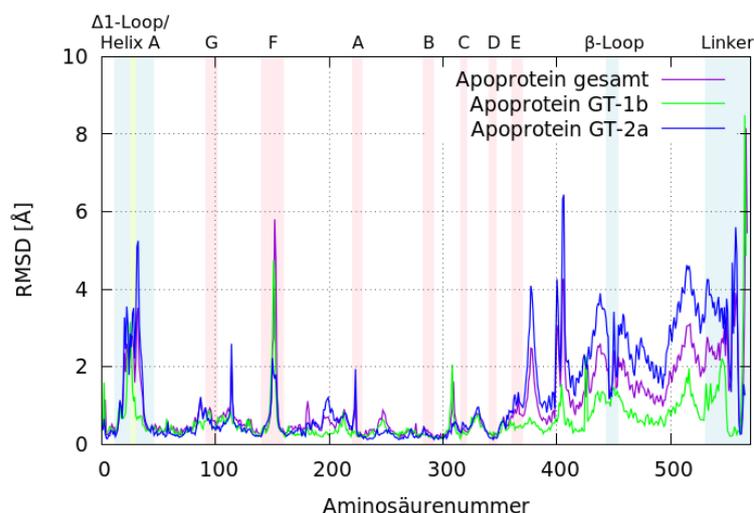


Abbildung 21: Darstellung der THESEUS-Analyse der 23 Apoprotein-Strukturen von NS5B aus der PDB. Diese wurden entsprechend der jeweiligen Aminosäurenummer aufgetragen (violett). Die Einteilung der Kristallstrukturen erfolgte in die jeweiligen Genotypen (GT-1b in grün und GT-2a in blau).

Die mit dem ML-Algorithmus überlagerten Kristallstrukturen sind in Abbildung 22 gezeigt. Strukturelle Abweichungen sind in blau für geringe Unterschiede, über grün und gelb nach rot für große Abweichungen gefärbt. Deutlich ist die farbliche Abhebung der Daumen-Domäne und des Linkers in grün und hellblau zu erkennen. Das C-terminale Ende des Linkers ist rot gefärbt. Betrachtet man das Protein vom NTP-Eingang aus, so erkennt man das Motiv F (140-160) in grün, gelb und rot und einen hellblau eingefärbten Δ 1-Loop mit der Helix A. Vergleicht man die Superpositionierung der Kristallstrukturen nach der Methode der kleinsten Quadrate (Abbildung 19) und nach der ML-Methode (Abbildung 22), so ist die generelle Übereinstimmung offensichtlich.

Diese Ergebnisse sind vergleichbar mit den Untersuchungen von Caillet-Saguy *et al.*⁶⁶. Diese beschrieben die unterschiedlich starke Rotation der Spitze der Daumen-Domäne zwischen den Kristallstrukturen mit dem Linker (Variante Δ 21) und ohne den Linker (Variante Δ 55) mit einem maximalen Unterschied von 4,2 Å in diesem Bereich (genaue Messpunkte wurden nicht beschrieben). Des Weiteren zeigten Caillet-Saguy *et al.* eine Verschiebung der äußeren Finger-Domäne („*outer fingertips*“), wobei die Region um Helix A gemeint ist, zwischen den Aminosäurenummern 20 und 30, und innerhalb der Fingerspitzen („*top of fingers*“) zwischen den Aminosäuren 83 und 111 (beinhalten Motiv G). Diese Ergebnisse sind analog zu den hier mit dem ML-Algorithmus und aus der RMSD-Analyse erhaltenen Daten. Zusätzlich konnten in dieser Arbeit weitere Unterschiede innerhalb des Linkers, innerhalb des gesamten Δ 1-Loops, sowie innerhalb des Motivs F aufgezeigt werden. Unterschiede zwischen den Proteinstrukturen im Bereich des Motivs G konnten jedoch nicht identifiziert werden.

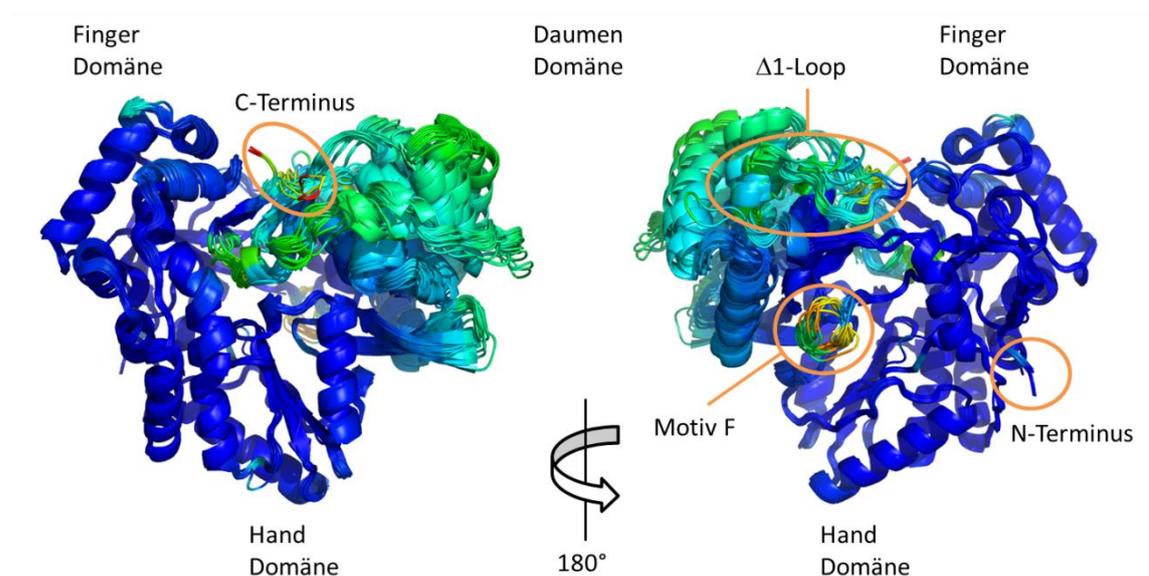


Abbildung 22: Überlagerung der 23 Apoprotein-Kristallstrukturen aus der THESEUS-Analyse. Die Proteine sind entsprechend ihrer Varianzen zueinander eingefärbt von blau für eine geringe Varianz bis rot für eine hohe Abweichung. Linke Seite: Die um 180° gedrehte Ansicht zeigt auf den verschlossenen dsRNA-Ausgang. Auf der rechten Seite der Abbildung ist das Protein mit Sicht auf den NTP-Eingang gezeigt.

4.2. Moleküldynamik-Simulationen von NS5B

In den Kristallstrukturen von NS5B liegt die Polymerase in einem geschlossenen Zustand vor. In dieser Konformation ist zwar die Initiation der RNA-Replikation möglich, die prozessive Phase wird durch den verschlossenen dsRNA-Ausgang jedoch ausgeschlossen.¹ Für die Replikation des Virusgenoms muss die Polymerase mit ssRNA und Nukleotiden interagieren. Dabei bindet NS5B die ssRNA mit einer höheren Affinität als NTP-Moleküle. CD-Messungen an NS5B zeigten eine Veränderung der Tertiärstruktur während der NTP-Bindung.⁶⁷ Die bisherigen Untersuchungen an der Polymerase und ihren Kristallstrukturen geben einen Hinweis auf die notwendige Konformationsänderung des Proteins beim Übergang in den aktiven Zustand.^{1,52,54,59,62,66} Die Kristallisation der prozessiven Polymerase konnte jedoch nur durch Deletionsvarianten des β -Loops und des Linkers erreicht werden. Weiterhin wurde der Wirkungsmechanismus des Inhibitors HCV-796 von Reich *et al.* kinetisch untersucht. Dabei stellte sich heraus, dass HCV-796 die dsRNA-Bildung stört, nicht aber die ssRNA-Bindung. Im CD-Spektrum konnte ebenfalls bei Anwesenheit von HCV-796 eine Veränderung der Tertiärstruktur detektiert werden.⁶⁷

Um die Untersuchungen von Reich *et al.* molekularmechanistisch zu analysieren, wurden MD-Simulationen mit NTP, ssRNA und dem Inhibitor HCV-796 durchgeführt. Die in der

PDB vorhandenen Kristallstrukturen wiesen, wie bereits in Kapitel 4.1.2 diskutiert, fehlende Koordinaten für einzelne Aminosäuren oder kurze Abschnitte der Proteinkette auf. Diese wurden in MOE mit Hilfe der „*structure preparation*“-Anwendung modelliert. Die MD-Berechnungen der Proteinstrukturen erfolgten mit AMBER, dem AMBER Kraftfeld FF03 und den in Abschnitt 3.4 beschriebenen Einstellungen. Die Dauer jeder Simulation betrug 200 ns. Dies war die minimale Zeitspanne, um jedes Ensemble mindestens 100 ns im thermodynamischen Gleichgewicht, gekennzeichnet durch einen auf der MD-Zeitskala konstanten RMSD-Wert, analysieren zu können.

4.2.1. MD-Simulation von NS5B mit Substraten

Um die Konformationsänderungen bei der Bindung von ssRNA oder Nukleotiden zu untersuchen, wurden MD-Simulationen von NS5B mit ssRNA und GTP als Ligand durchgeführt. Für NS5B existiert nur eine Kristallstruktur mit gebundener Einzelstrang-RNA (1NB7⁵⁷). Hierbei diente ein Uracil-Tetramer (3'-U₁-U₂-U₃-U₄-5') als ssRNA-Modell. Zur Untersuchung der Konformationsänderung von NS5B in Anwesenheit von Nukleotiden wurde die Kristallstruktur 2XI3¹ verwendet. Diese Struktur enthält keine fehlenden Aminosäuren, ein vollständig aufgelöstes GTP-Molekül im aktiven Zentrum des Proteins und eine sehr gute Auflösung von 1,77 Å. Zum Vergleich wurden die Kristallstrukturen der Apoproteine 1NB4⁵⁷ und 2XI2¹ simuliert, die von denselben Arbeitsgruppen und unter den gleichen Bedingungen wie 1NB7 bzw. 2XI3 kristallisiert wurden.

In Abbildung 23 sind die RMSD-Werte der MD-Simulationen dargestellt. Die MD-Simulation von 2XI3 zeigt die größten Strukturveränderungen mit einem RMSD von 8 Å auf, 2XI2 liegt bei 5 Å, 1NB4 bei 4,4 Å und 1NB7 bei 4,3 Å. In der MD-Simulation von 1NB7 war der N-Terminus bis zur Aminosäure Ala9 instabil, die restliche Proteinstruktur blieb aber erhalten, sodass dies keine weiteren Auswirkungen auf den Verlauf der Simulation hatte. Im Anhang sind die RMSF-Werte der jeweiligen MD-Simulationen in einem Multiplot dargestellt (Abbildung 68).

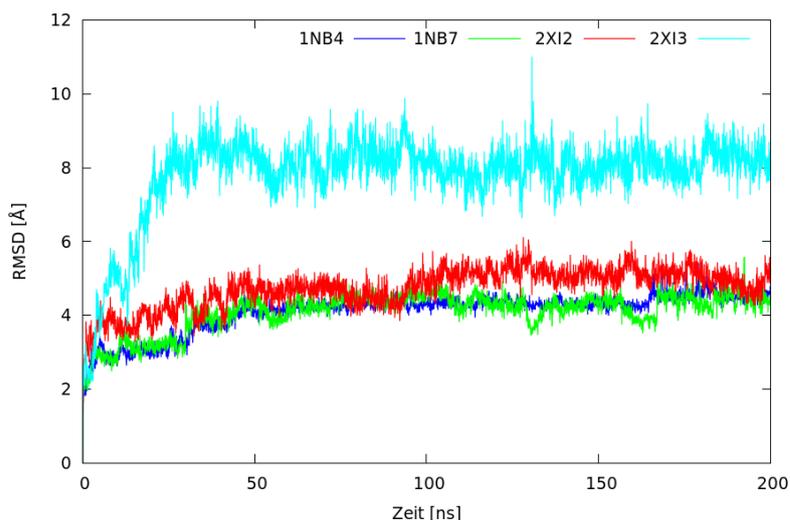


Abbildung 23: *Root mean square deviation* (RMSD)-Werte der *backbone*-Atome N, C $_{\alpha}$ und C der MD-Simulationen.

Die Kristallstruktur 1NB4 ist blau gefärbt, 1NB7 grün, 2XI2 rot und 2XI3 cyan.

Die RMSD-Werte beschreiben die Veränderung der Proteinstruktur über die Simulationszeit. Um diese besser zu analysieren wurden nach Davis *et al.*¹⁶³ der Interdomänen-Winkel und der Durchmesser des RNA-Tunnels berechnet (vgl. Abschnitt 3.5) und diese in Abbildung 24 gegeneinander aufgetragen. Ein Winkel größer als 70° und ein Durchmesser größer als 20 Å deuten auf eine geöffnete Konformation der Polymerase hin, kleinere Werte dagegen auf eine geschlossene. Die MD-Simulationen von 1NB4 und 1NB7 besitzen geringe Werte unter 70° und 20 Å und zeigen eine geschlossene Konformation, wobei 1NB7 die kleineren Winkel und Distanzen aufweist. In den Simulationen von 2XI2 und 2XI3 besitzen beide Proteine einen großen Interdomänen-Winkel, wobei 2XI3 deutlich über der 70°-Grenze liegt und einen großen RNA-Tunnel-Durchmesser bis zu 30 Å besitzt. 2XI2 besitzt demzufolge eine geschlossene Konformation während der MD-Simulation, wohingegen 2XI3 eine geöffnete Form der Polymerase darstellt.

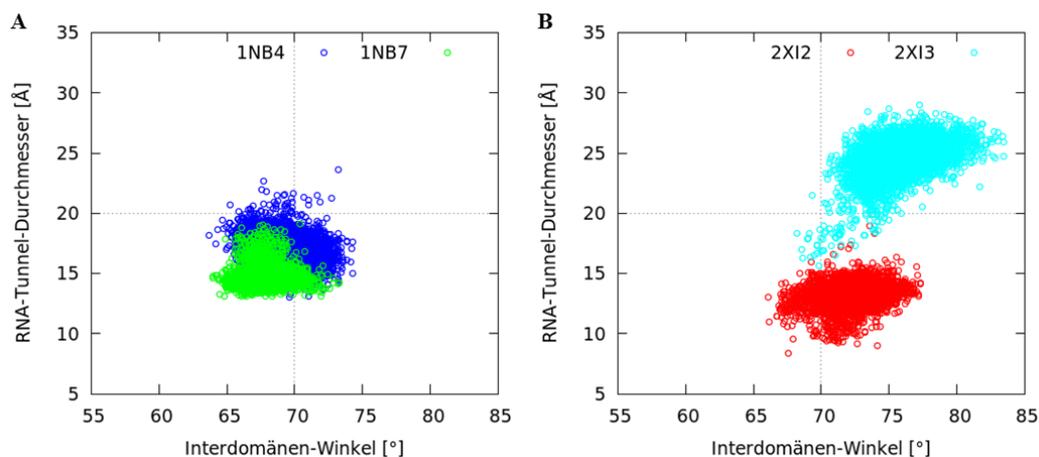


Abbildung 24: Darstellung des RNA-Tunnel-Durchmessers aufgetragen gegen den Interdomänen-Winkel. **A)** Berechnete Werte der Simulation der Kristallstrukturen 1NB4 gefärbt in blau und 1NB7 in grün. **B)** Ermittelte Werte der Simulation der Kristallstrukturen 2XI2 in rot und 2XI3 in cyan.

Der RMSD-Wert und der Interdomänen-Winkel geben jedoch nur Hinweise auf die Strukturveränderungen der Proteine. Um weitere Strukturmerkmale herauszufinden, welche zu den Konformationsunterschieden führten, wurden die Durchschnittsstrukturen der MD-Simulationen betrachtet und der RMSD-Wert pro Aminosäure (RMSF-Wert) berechnet. In Abbildung 25 werden die Trajektorien von 1NB4 und 1NB7 miteinander verglichen. Die größten Unterschiede befinden sich im Bereich des Motivs F (140-160) zwischen den Aminosäuren 490 und 520, welche den zwei äußeren α -Helices der Daumen-Domäne entsprechen, und dem Linker (530-560). Exemplarisch wurden die Unterschiede des Motivs F und des Linkers in Abbildung B und C dargestellt. Bei Anwesenheit von ssRNA besitzt das Motiv F eine andere Position und ist in Richtung des RNA-Tunnels ausgerichtet. In der MD mit 1NB4 zeigt es dagegen vom Protein weg. Der Linker zeigt in großen Teilen Positionsänderungen auf, verschließt aber in beiden Simulationen immer noch den dsRNA-Ausgang. In der Simulation von 1NB4 konnte sich sogar eine kurze helikale Struktur zwischen den Aminosäuren 542 und 552 ausbilden. Des Weiteren trägt der Bereich zwischen den Aminosäuren 490 und 520 durch seine Positionsveränderung zum verkleinerten Interdomänen-Winkel bei.

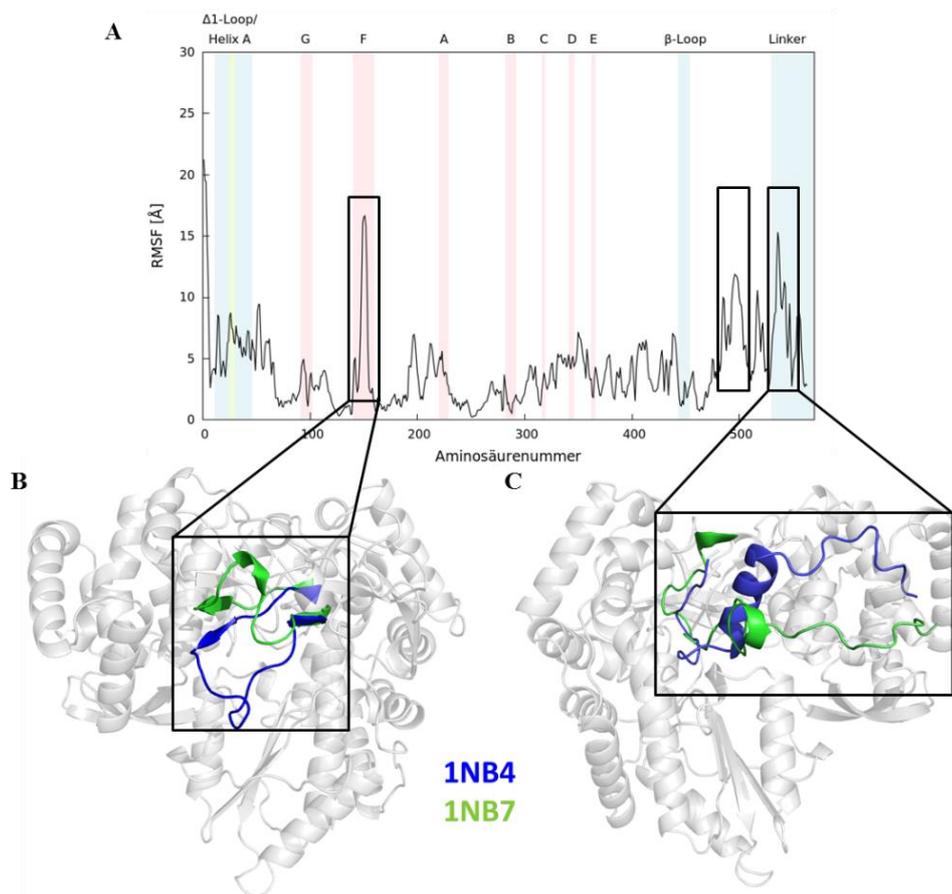


Abbildung 25: RMSF-Vergleich der MD-Durchschnittsstrukturen von 1NB4 und 1NB7.

A) Diagramm mit den RMSF-Werten. Die wichtigsten Strukturmerkmale und Motive sind farblich markiert und über dem Diagramm beschriftet. B) Cartoon-Darstellung der Polymerase mit Blick auf den NTP-Eingang. Das Motiv F ist jeweils für 1NB4 in blau und für 1NB7 in grün dargestellt. C) Cartoon-Darstellung der Polymerase mit Blick auf den verschlossenen dsRNA-Ausgang. Der Linker ist für 1NB4 in blau dargestellt und für 1NB7 in grün.

Die Durchschnittsstrukturen von 2XI2 und 2XI3 wurden ebenfalls zur Berechnung des RMSF-Werts verwendet (vgl. Abbildung 26). Dabei zeigten sich mehrere Unterschiede innerhalb der Finger- und der Daumen-Domäne zwischen den Aminosäuren 17-35 (Δ 1-Loop mit Helix A, gelb markiert), 85-120 (Motiv G, orange markiert), 402-405 (grün markiert), 442-455 (β -Loop, cyan markiert), 503-523 (blau markiert) und 532-562 (Linker, pink markiert). Die Farbcodierung aus Abbildung A wurde in B und C beibehalten. Diese zeigen die Durchschnittsstrukturen von 2XI2 und 2XI3 mit einer Connolly-Oberfläche. Für die Struktur von 2XI3 (Abbildung C) ist eine deutliche Öffnung zwischen Finger- und Daumen-Domäne erkennbar, welche einen Durchmesser von ca. 26 Å besitzt und somit eine Doppelstrang-RNA beherbergen könnte (vgl. Abbildung 69 im Anhang).

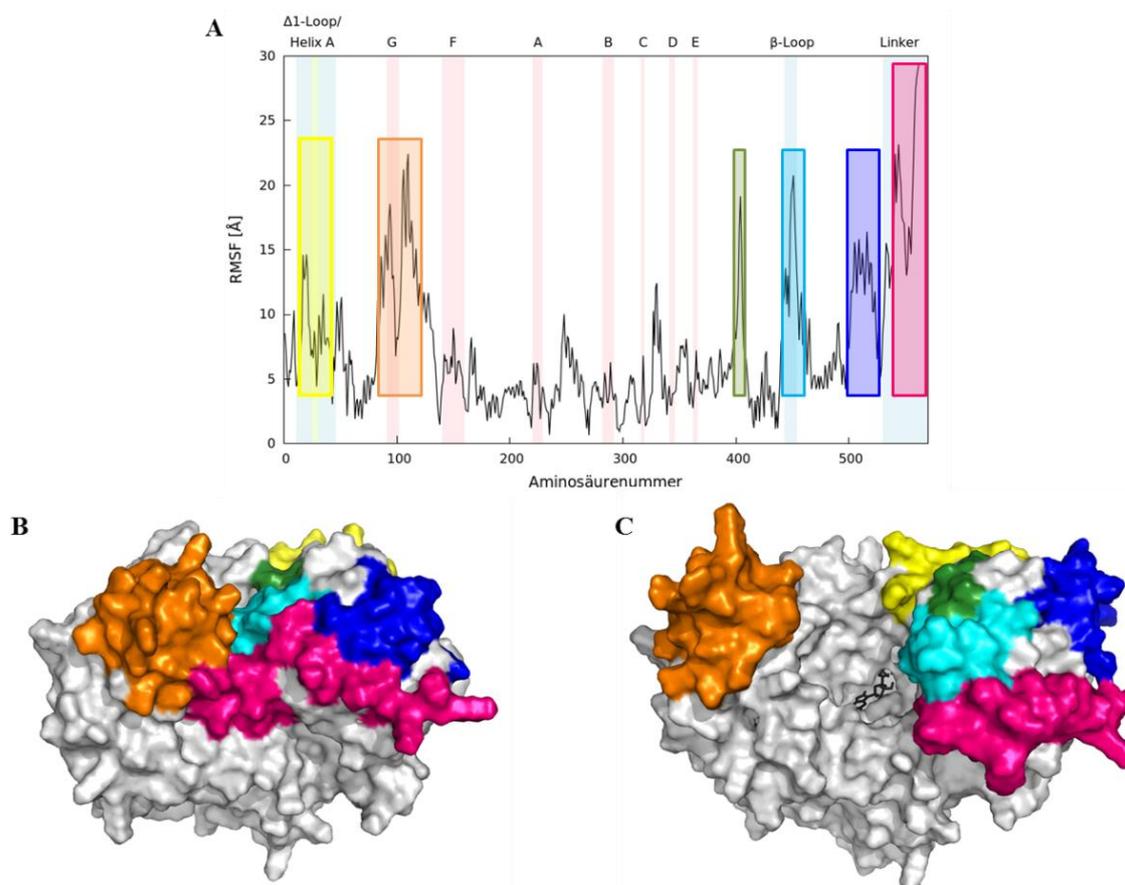


Abbildung 26: Vergleich der MD-Durchschnittsstrukturen aus 2XI2 und 2XI3.

A) RMSF-Werte aus der Simulation von 2XI2 und 2XI3. Die größten Unterschiede wurden farbig markiert. B) und C) Durchschnittsstrukturen aus den MDs mit 2XI2 (links) und 2XI3 (rechts), dargestellt mit einer Connolly-Oberfläche. Die im Diagramm markierten Bereiche wurden entsprechend eingefärbt. Das GTP-Molekül aus der Simulation mit 2XI3 ist schwarz dargestellt und in der Stick-Darstellung abgebildet.

Die wichtigsten Bewegungsabläufe, die anhand der MD von 2XI3 zu der Öffnung von NS5B führten sind:

- β-Loop und Linker bewegen sich aus dem aktiven Zentrum heraus
- Rotation der Daumen-Domäne in Richtung der Helix A
- Bewegung der Helix A bzw. des Δ1-Loops mit der Daumen-Domäne
- der Linker wird flexibel und lagert sich an der Daumen-Domäne an
- nach 50 ns tritt in der MD eine verstärkte Fluktuation des Motivs G und benachbarter Aminosäuren auf

Diese Beobachtungen passen gut zu den bereits postulierten Bewegungsmechanismen von NS5B, wie sie in der Einleitung unter Abschnitt 1.5 erwähnt wurden.^{1,59,62,66,67} Nur die verstärkte Beweglichkeit von Motiv G wurde so noch nicht beschrieben, konnte aber in der Kristallstrukturen-Analyse von Caillet-Saguy *et al.*⁶⁶ beobachtet werden (vgl. Kapitel 4.1.4). Diese Fluktuationen der Fingerspitzen von Motiv G tragen zu einer Vergrößerung des

dsRNA-Ausgangs bei, da sie sich weiter von der Daumen-Domäne wegbewegen. Diese Beobachtung ist gut mit dem allgemeinen Öffnungs- und Schließmechanismus von Viren-DNA-Polymerasen vergleichbar, wobei sich Daumen- und Finger-Domäne aufeinander zu und wieder voneinander weg bewegen^{183–188}.

Interaktionen zwischen NS5B und der ssRNA

Bezüglich der Wechselwirkungen zwischen der Einzelstrang-RNA und der Polymerase existieren unterschiedliche Publikationen mit unterschiedlichen Ergebnissen. O’Farrell *et al.*⁵⁷ und Deval *et al.*¹⁸⁹ bestimmten die Bindestelle der einzelsträngigen RNA durch Kristallisation (PDB-Eintrag 1NB7) und durch RNA-*Footprinting*-Experimente. Wichtige Interaktionspartner der ssRNA in der Kristallstruktur sind: Cys14, Pro93, His95, Ala97, Lys98, Met139, Lys141 und Ile160. Das RNA-*Footprinting*-Experiment definiert Wechselwirkungen mit Lys98, Lys100, Lys106, Arg158, Arg386 und Arg394. In diesen Analysen ist nur die Aminosäure Lys98 identisch. Es ist fraglich, inwieweit die Kristallstruktur 1NB7 und das *Footprinting*-Experiment die native Bindung der ssRNA widerspiegeln können, da die Co-Kristallisation durch *soaking* erfolgte und beim *Footprinting* nur aminreaktive Aminosäuren untersucht wurden. Daher wurde versucht aus den MD-Simulationen detailliertere Einblicke in die ssRNA-NS5B-Wechselwirkungen zu erhalten.

Einhergehend mit den Konformationsveränderungen der Polymerase in den MD-Simulationen kommt es zu Positionsunterschieden der Substrate und damit zu unterschiedlichen Protein-Substrat-Interaktionen. In Abbildung 27 ist die veränderte Bindung der ssRNA aus der Kristallstruktur und der MD-Simulation dargestellt. Deutlich ist eine Positionsänderung in Abbildung B zu beobachten. Die Wasserstoffbrücken zwischen NS5B und der ssRNA sind in der Kristallstruktur und während der MD-Simulation in den Abbildungen C und D zu sehen. Dabei ist eine Zunahme an Wasserstoffbrücken in der Simulation erkennbar. Die Interaktionspartner sind Cys14, Lys98 und Lys100 aus dem Motiv G, Lys141 aus dem Motiv F, Val161, Arg168, Ser282 aus dem Motiv B, Ala450 aus dem β -Loop, sowie Gly557, Gly559 und Ile560 aus dem Linker. So zeigt die MD-Simulation von 1NB7 Wechselwirkungen auf, welche sowohl von O’Farrell *et al.*⁵⁷ als auch von Deval *et al.*¹⁸⁹ bestimmt wurden.

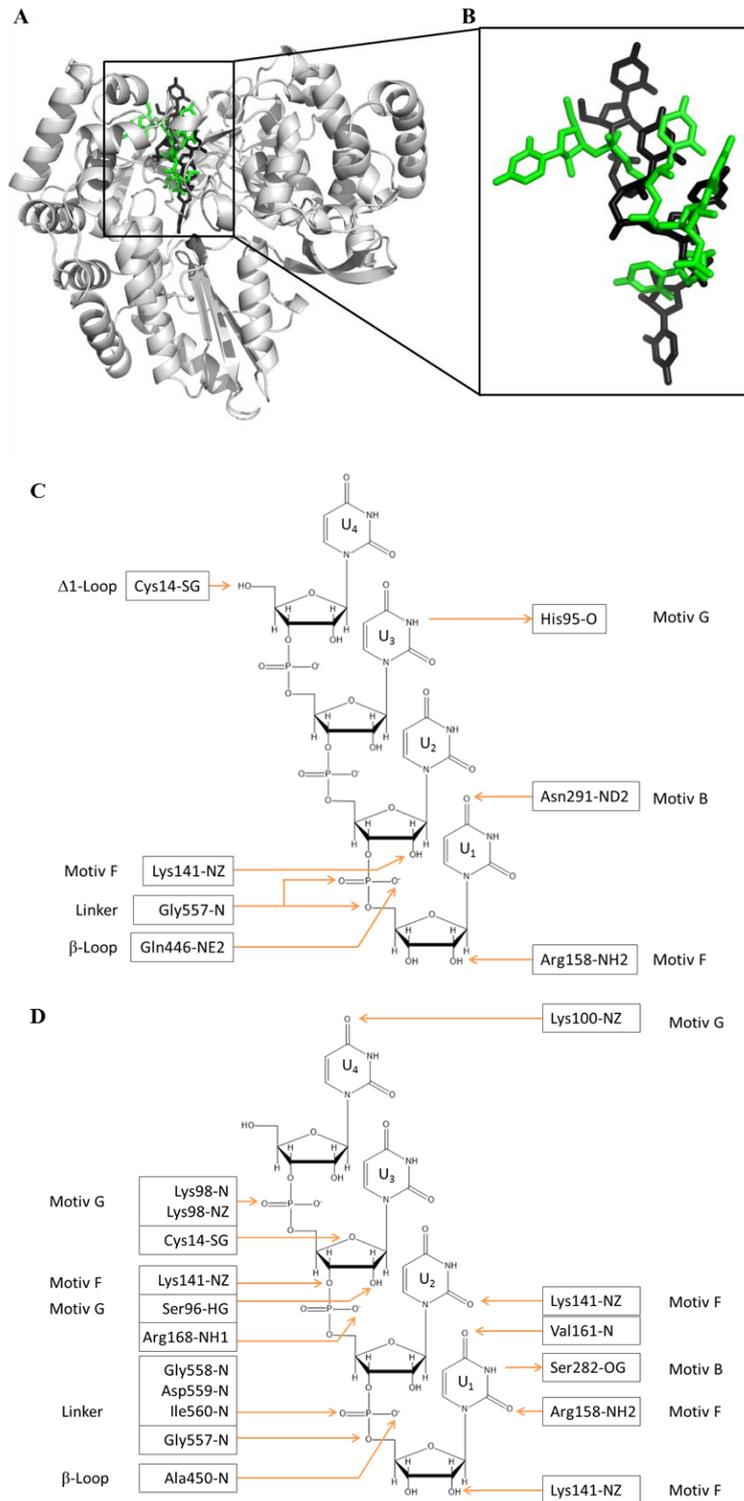


Abbildung 27: Positionsvergleich der ssRNA aus der MD-Simulation mit 1NB7.

A) Die Durchschnittsstruktur der MD-Simulation von 1NB7 ist als Cartoon-Darstellung in grau abgebildet. Die ssRNA dieser Durchschnittsstruktur ist grün eingefärbt. Im Vergleich dazu ist die ssRNA der Kristallstruktur in schwarz dargestellt. **B)** Detailaufnahme der ssRNA-Strukturen aus der MD-Simulation in grün und aus der Kristallstruktur in schwarz. Die Wasserstoffatome sind aus Gründen der Übersichtlichkeit nicht gezeigt. **C)** Darstellung der Wasserstoffbrücken-Wechselwirkungen mittels Pfeilen, welche eine Richtung vom Wasserstoffbrücken-Donor zum -Akzeptor besitzen, zwischen der ssRNA (3'-U₁-U₂-U₃-U₄-5') und der Polymerase aus der Kristallstruktur. **D)** Wasserstoffbrücken-Wechselwirkungen zwischen ssRNA und NS5B in der MD-Simulation.

In der Simulation von 1NB7 reicht das erste 3'-Uracil-Nukleotid am tiefsten in die Polymerase und kann in der Moleküldynamik mit 7 Aminosäureresten Wasserstoffbrücken ausbilden. Dazu gehören Ala450 aus dem β -Loop und Gly557, Gly559 und Ile560 aus dem Linker, welche mit dem Phosphatrest der ssRNA interagieren. Diese Wechselwirkungen sorgen dafür, dass sich der β -Loop und der Linker in der Moleküldynamik nicht aus dem aktiven Zentrum heraus bewegen können. Tatsächlich sollen der β -Loop und der Linker für die *de novo* Synthese der dsRNA wichtig sein, wobei Ala450 als kritische Aminosäure für die RNA-Replikation und für die *de novo* RNA-Synthese identifiziert wurde.^{190–192}

Interaktionen zwischen NS5B und dem GTP-Molekül

Die existierenden Untersuchungen der Interaktion zwischen dem Nukleotid und der Polymerase fallen sehr unterschiedlich aus. Ausgehend von den Kristallstrukturen 1GX6 und 1GX5 beschreiben Bressanelli *et al.*⁵⁴ das katalytische Zentrum der Polymerase als *catalytic site* (C) mit den Aminosäuren Asp220, Asp318 und Asp319, als *priming site* (P) mit Arg158, Ser367, Arg386, Thr390 und Arg394 und als *interrogation site* (I) mit Arg48, Lys51, Lys151, Lys155, Arg158, Leu159 und Asp225. O'Farrel *et al.*⁵⁷ definiert die NTP-Bindestelle mit vergleichbaren Aminosäuren aus der Kristallstruktur 1NB6 mit Arg41, Arg48, Lys51, Lys141, Lys158, Arg222, Ser282, Thr283, Asn291 und Ser556. Wiederum andere Interaktionspartner zeigte die Kristallstruktur 2XI3 von Harrus *et al.*¹ mit den Aminosäuren Arg158, Ser282, Ser367, Arg386, Glu446, Ser556 und Gly557. Auch die hier aufgeführten Kristallstrukturen wurden mittels *soaking* erzeugt. Eine zusammenfassende Auflistung der Wechselwirkungspartner ist im Anhang in Tabelle 8 aufgeführt.

Die Wasserstoffbrücken zwischen GTP und NS5B aus der Kristallstruktur 2XI3 und deren MD-Simulation sind in Abbildung 28 dargestellt. In der Kristallstruktur konnten Wasserstoffbrücken mit den Aminosäuren Arg48, Lys51, Arg158, Asp225 und Asp318 identifiziert werden, welche in den Motiven A, C und F vorkommen (vgl. Abbildung A). In der Simulation bilden vor allem die Reste Lys155, Arg158, Arg222, Cys223 Cys316, Asp318, Asp319, Cys366 und Arg386 mit dem GTP-Molekül Wasserstoffbrücken aus. Diese Aminosäuren sind Bestandteile der Motive A, C, E und F, wobei Motiv E (Cys366) und Motiv C (Asp318 und Asp319) mit der Nukleobase wechselwirken und Motiv A (Arg222 und Cys223) und Motiv F (Lys155 und Arg158) mit der Ribose und dem Triphosphatende (vgl. Abbildung B). Die auffälligsten Unterschiede im Vergleich zu den publizierten Daten sind die Wechselwirkungen mit Motiv E und die Verschiebung der Wasserstoffbrücken von Motiv A mit der Nukleobase im Kristall zu dem Triphosphatende in der Dynamik.

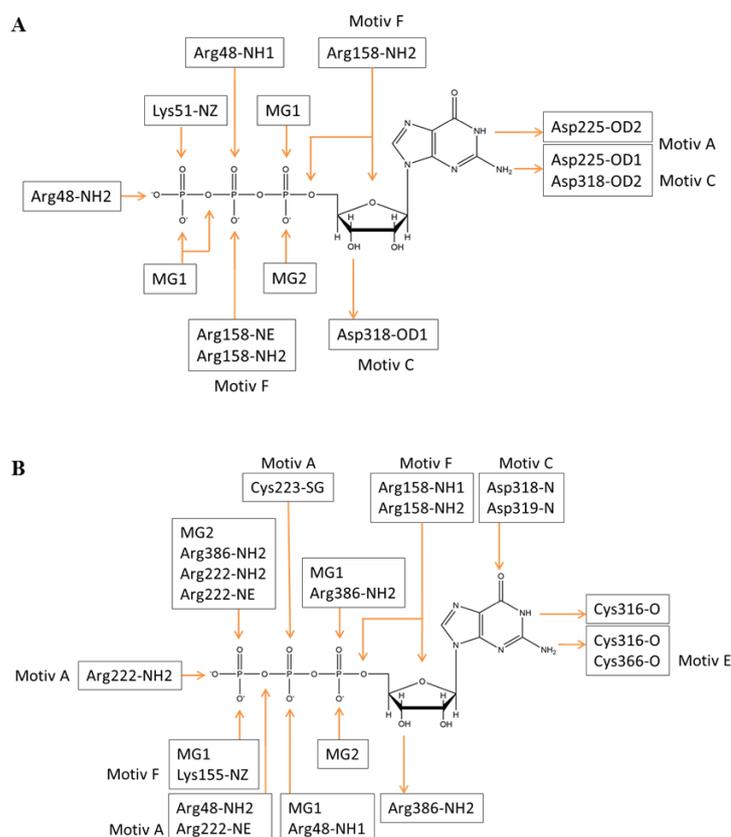


Abbildung 28: Nicht-bindende Wechselwirkungen von NS5B und GTP aus der MD-Simulation von 2XI3. Die Wasserstoffbrückenbindungen sind durch Pfeile vom Wasserstoff-Donor- zum Wasserstoff-Akzeptor-Atom gekennzeichnet.

4.2.2. MD-Simulation einer Doppelstrang-RNA mit NS5B in geöffneter Konformation

Die Dynamik von NS5B mit GTP zeigte vor allem in den Bereichen der Finger- und der Daumen-Domäne Veränderungen der Proteinstruktur hin zur prozessiven Konformation auf. Um diese geöffnete Konformation weiter zu untersuchen, wurde eine Simulation von NS5B in diesem Zustand mit einer Doppelstrang-RNA vorgenommen. Da die kristallisierten NS5B-dsRNA-Komplexe nur sehr kurze dsRNA-Abschnitte enthalten, wurde für die MD-Simulation der PDB-Eintrag 4K4S¹⁹³ verwendet. Darin wurde die RNA-abhängige RNA-Polymerase des Poliovirus mit einer Doppelstrang-RNA kristallisiert, deren Sequenz in Abbildung 29 abgebildet ist. Das erstellte NS5B-dsRNA-Modell entspricht dem Zeitpunkt der Replikation, nachdem die Pyrophosphat-Hydrolyse erfolgte und die Template-RNA einen Platz weitergerückt ist. Die Position +1 des Tochterstrangs ist unbesetzt und bietet Platz für ein NTP-Molekül.

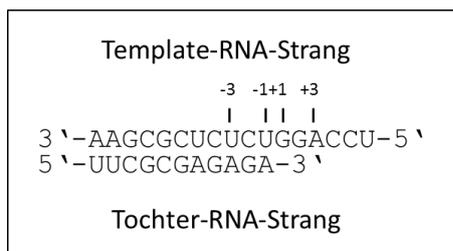


Abbildung 29: Sequenz des RNA-Konstrukts im NS5B-dsRNA-Modell.

Für die Analyse dieser Dynamik sind in Abbildung 30 zum Vergleich die Werte aus den bereits berechneten Simulationen ausgehend von den Kristallstrukturen 2XI2 und 2XI3 dargestellt. Die Simulation von NS5B mit dem dsRNA-Konstrukt zeigt einen sehr stabilen RMSD-Wert, der vergleichbar mit dem von 2XI2 ist (Abbildung A). Der Interdomänen-Winkel und der RNA-Tunnel-Durchmesser sind ebenfalls sehr stabil und zeigen eine durchschnittliche Öffnung des RNA-Tunnels von 28 Å und einen Domänen-Winkel von 76° (vgl. Abbildung B). Im Vergleich zu den MD-Simulationen von 2XI3 oder 2XI2 weisen die Werte hier auf eine geringere Flexibilität der Proteinstruktur und damit auf einen kleineren Konformationsraum hin.

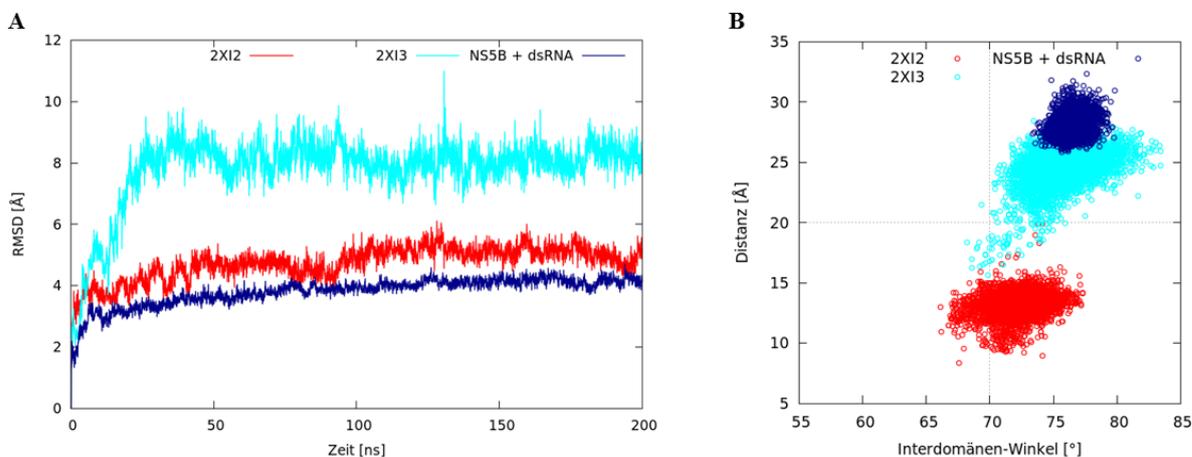


Abbildung 30: Überblick über die Simulation von NS5B in der geöffneten Form mit dsRNA. Zum Vergleich wurden noch einmal die Simulationen von 2XI2 (Apoprotein) und 2XI3 (NS5B mit GTP) in den Diagrammen mit aufgetragen. Die RMSD-Werte der geöffneten Polymerase sind dunkelblau, die von NS5B mit GTP (2XI3) sind cyan und die Werte des Apoproteins (2XI2) sind rot gefärbt. **A)** Die RMSD-Werte wurden anhand der Protein-backbones (N, C α und C) berechnet. **B)** Auftragung des Interdomänen-Winkels gegen den RNA-Tunnel-Durchmesser.

Die strukturellen Unterschiede während der Simulation mit dem dsRNA-Konstrukt sind in Abbildung 31 gezeigt. Anhand des dargestellten RMSF-Wertes der Durchschnittsstrukturen

der Simulationen von 2XI3 und NS5B mit dsRNA können deren Veränderungen identifiziert werden. Auffällige Unterschiede finden sich in dem $\Delta 1$ -Loop mit Helix A zwischen den Aminosäuren 20 und 33, im Motiv G zwischen 86 und 117, im Motiv F zwischen 149 und 157 und innerhalb des Linkers zwischen 546 und 562. Die Positionsänderungen der Helix A und des Motivs G sind exemplarisch in Abbildung B und C dargestellt.

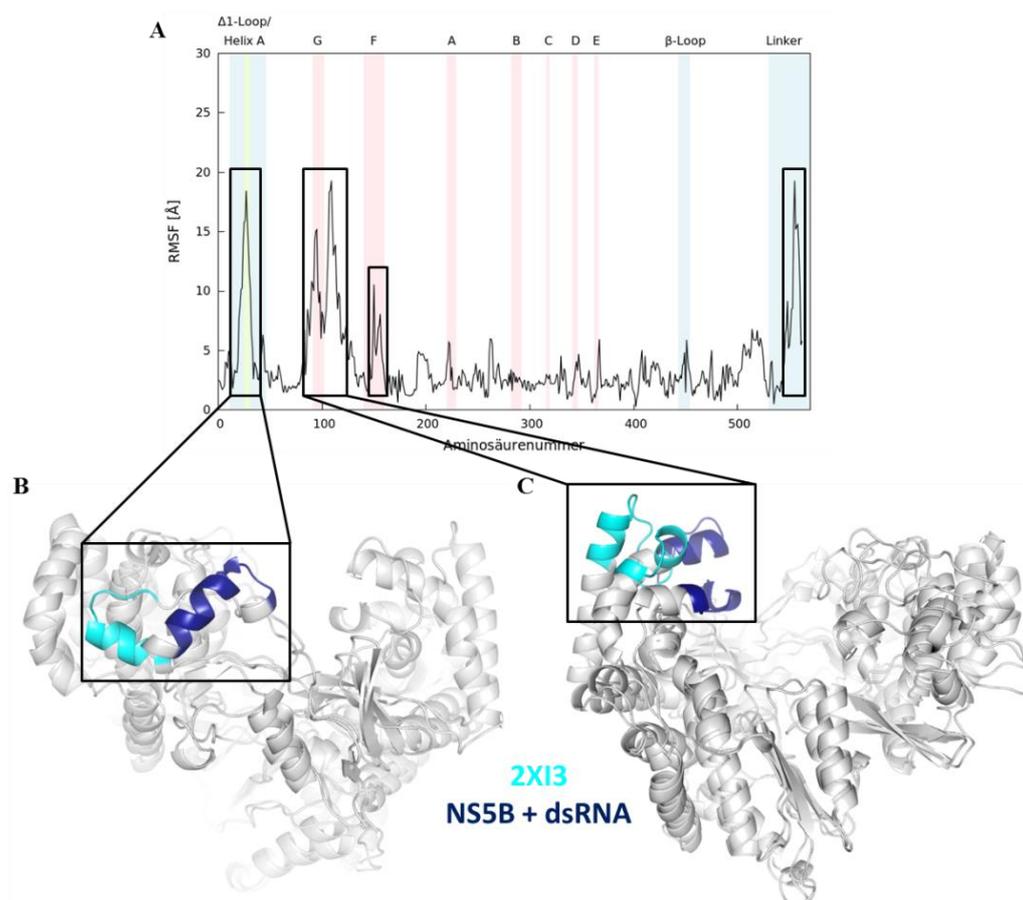


Abbildung 31: Vergleich der MD-Simulationen von 2XI3 und der geöffneten Konformation mit dsRNA. Die dsRNA ist aus Gründen der Übersichtlichkeit nicht gezeigt. **A)** RMSF-Unterschied zwischen der Durchschnittsstruktur von 2XI3 und der MD mit dsRNA. **B)** Cartoon-Darstellung der Polymerase mit Blick auf den NTP-Eingang. Die Helix A ist für 2XI3 cyan und für NS5B mit dsRNA dunkelblau dargestellt. **C)** Cartoon-Darstellung der Polymerase mit Blick auf die dsRNA-Öffnung. Das Motiv G ist entsprechend eingefärbt.

Die C-terminale Positionsänderung des Linkers (Abbildung 31A) resultiert aus dem fehlenden Anschluss an den Membrananker. In der geschlossenen Konformation interagiert der Linker mit der Finger-Domäne. Während der Öffnung der Polymerase geht diese Interaktion verloren, sodass sich der C-Terminus neu orientiert. Das Vorhandensein eines Membranankers würde dem Linker eine Ausrichtung geben und dessen Flexibilität beschränken.

Die Verlagerung des Motivs G resultiert aus der Interaktion mit dem eintretenden Template-RNA-Strang (vgl. Abbildung 32). Diese Wechselwirkungen könnten die Auflösung der dreidimensionalen Struktur des *Stem-Loops* in der nicht-translatierten Region am 3'-Ende des HCV-Genoms erleichtern. Durch diese Interaktionen mit dem Motiv G würde der zufällig aufgefaltete Bereich der Template-RNA, welcher als nächstes in die Polymerase eintreten würde, seine zuvor ausgebildeten Wasserstoffbrücken-Bindungen innerhalb der RNA nur erschwert wiederherstellen können. Dies würde eine Verschiebung des Gleichgewichtes in Richtung der gestreckten RNA bedeuten. Untersuchungen bezüglich der Wasserstoffbrücken im Stem-Loop der 3'-NTR-Region wurden bereits von Reich *et al.* durchgeführt.¹⁹⁴

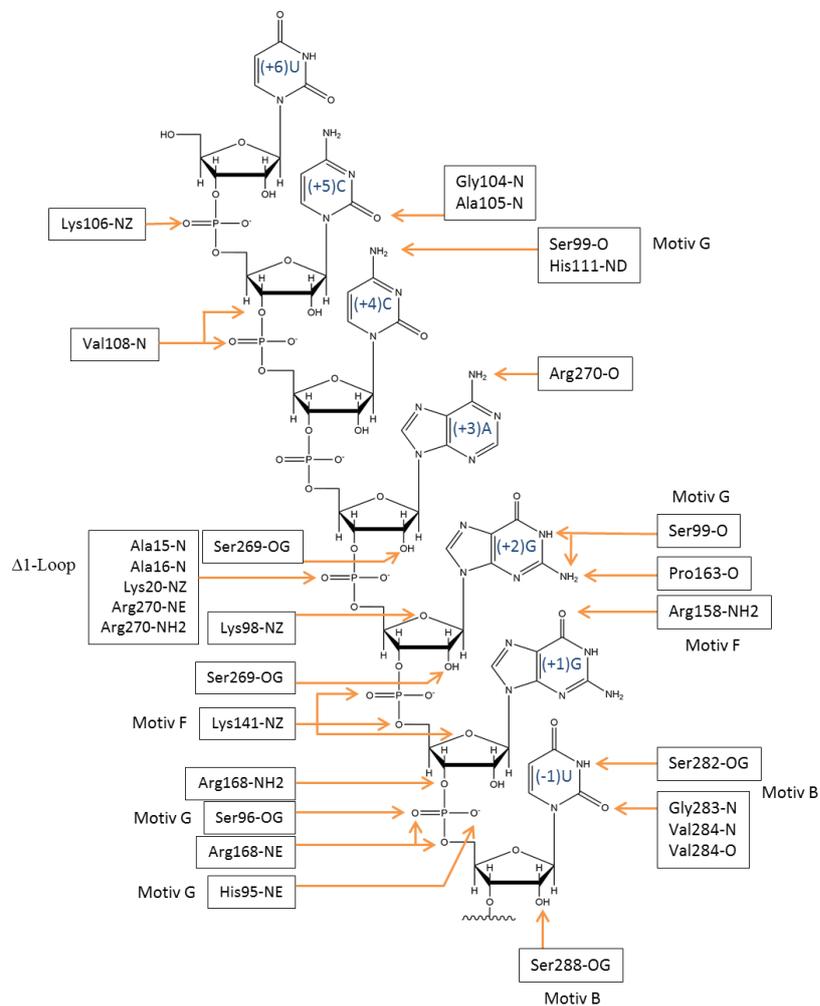


Abbildung 32: Schematische Darstellung der Wasserstoffbrücken zwischen NS5B und der Template-RNA. Die Wasserstoffbrückenbindungen sind durch Pfeile vom Wasserstoff-Donor- zum Wasserstoff-Akzeptor-Atom gekennzeichnet.

Das fehlende Nucleotid an Position +1 des Tochter-RNA-Strangs führt zu der Positionsänderung von Motiv F in Richtung des RNA-Tunnels, deren Ausrichtung vergleichbar mit dem Motiv F aus der MD-Simulation mit 1NB7 ist. Auch die Verlagerung

der Helix A ist der Ausrichtung in der Simulation von 1NB7 (NS5B + ssRNA) ähnlich (vgl. Abbildung 33A). Es existieren *backbone*-Wasserstoffbrücken zwischen den β -Faltblättern des Δ 1-Loops und des Motivs F, wie in Abbildung B gezeigt. Diese β -Faltblätter sind wie folgt definiert: β 2 (37-40), β 3 (136-140), β 4 (144-146) und β 5 (159-162), wobei β 2 mit β 4 interagiert und β 3 mit β 5. Positionsänderungen des Motivs F durch die Interaktion mit dem Triphosphatende eines Nukleotids könnten sich so auch auf die Helix A auswirken.

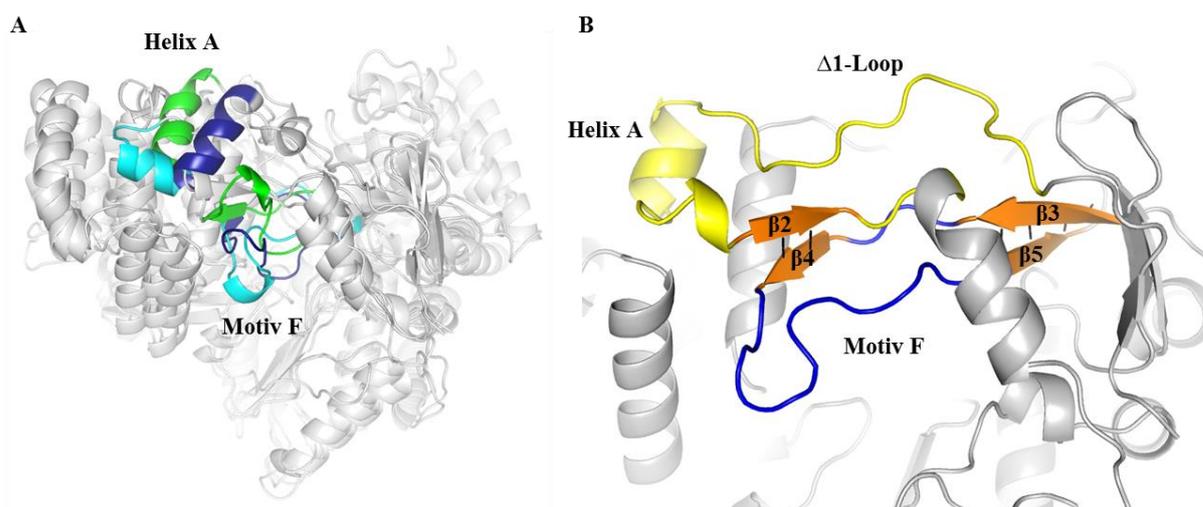


Abbildung 33: Detaillierte Darstellung des Δ 1-Loops mit der Helix A und dem Motiv F.

A) Darstellung der Durchschnittsstrukturen aus den Simulationen mit Blick auf den NTP-Eingang. Die Helix A und das Motiv F sind je nach der verwendeten Kristallstruktur eingefärbt. 1NB7 ist grün, 2XI3 ist cyan und die MD-Simulation von NS5B mit der dsRNA ist dunkelblau. **B)** Fokussierung auf den Δ 1-Loop in gelb und das Motiv F in blau. Die dazugehörigen β -Faltblätter (β 2 bis β 5) sind orange dargestellt.

4.3. Molekülmechanische Untersuchungen der Interaktionen von NS5B mit Nicht-Nukleotid-Inhibitoren

Die Experimente von Reich *et al.*^{67,194} zeigten nicht nur eine Tertiärstrukturveränderung bei der Bindung von Nukleotiden, sondern auch bei der Bindung des Inhibitors HCV-796. Der Mechanismus begründet sich in der Störung der dsRNA-Bildung.⁶⁷ Diese experimentellen Beobachtungen sollten mit Hilfe von Moleküldynamik-Simulationen auf mikroskopischer Ebene untersucht werden. Der Inhibitor HCV-796 bindet an der allosterischen Stelle NNI-4 in der Hand-Domäne des Proteins. Neben dieser Bindestelle wurden auch die von NNI-2, welche sich in der Daumen-Domäne befindet, von NNI-3 und von NNI-5, welche in der Hand-Domäne liegen, untersucht. Eine Simulation mit einem NNI-1 konnte nicht durchgeführt werden, da die vorhandenen Kristallstrukturen zu viele fehlende Aminosäuren aufwiesen. Eine Übersicht der verwendeten PDB-Strukturen ist im Anhang in Tabelle 9

aufgelistet. Die Wasserstoffbrücken, die über die Simulationszeit detektiert werden konnten, sowie die RMSD-Werte der Schweratome der Inhibitoren sind zum Vergleich in Tabelle 10 und Tabelle 11 im Anhang gezeigt.

4.3.1. MD-Simulationen mit dem Inhibitor HCV-796

Der Inhibitor HCV-796 ist in der Kristallstruktur 3FQK⁷⁷ enthalten und befindet sich in der NNI-4-Bindestelle innerhalb der Hand-Domäne, welche in der Nähe des NTP-Eingangs liegt (siehe Abbildung 34). HCV-796 hat ein Molekulargewicht von 446,49 g/mol, einen $\log P$ von 2,24, gehört zur Klasse der Benzofurane und besitzt einen IC_{50} -Wert von 0,081 μM .⁶⁷ Die Bindestellen für NNI-4 und NNI-3 liegen beide in der Hand-Domäne. Auffällig ist aber, dass sich die Bindetasche für NNI-4 sehr viel tiefer im Protein befindet. Die Aminosäure Arg200 versperrt in der Apoprotein-Struktur den Zugang zu dieser Tasche, weshalb die Rotation von Arg200 als geschwindigkeitsbestimmender Schritt für die langsame Bindung der Benzofurane angenommen wird.⁷⁷

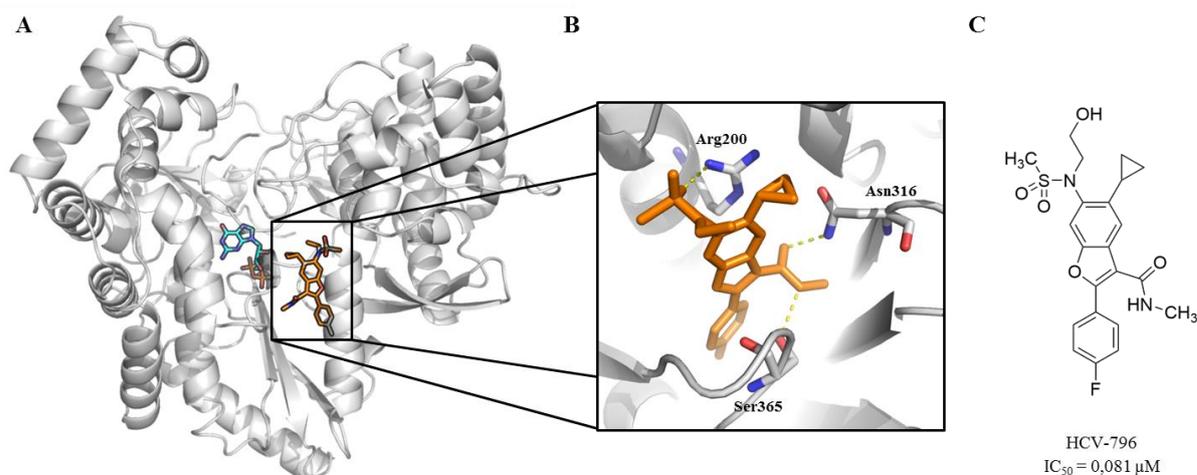


Abbildung 34: Die Position des Inhibitors HCV-796 im Protein und dessen Molekularstruktur.

A) Lage des Inhibitors (orange) in der Polymerase (grau) aus der Kristallstruktur 3FQK im Vergleich zum GTP-Molekül (cyan) aus 2XI3. **B)** Detailaufnahme der Bindestelle von HCV-796. Die Wasserstoffbrücken zwischen NS5B und HCV-796 sind durch zwei gelb gestrichelte Linien gekennzeichnet. **C)** Molekularstruktur von HCV-796.

Die Simulation von NS5B mit HCV-796 erfolgte ausgehend von der Kristallstruktur 3FQK. Da Hang *et al.*⁷⁷ neben diesem Proteinkomplex kein Apoprotein unter den gleichen Bedingungen kristallisiert haben, war kein direkter Vergleich mit einer verwandten Apoproteinstruktur möglich. In Abbildung 35A ist der RMSD-Wert über die Simulationsdauer dargestellt, welcher sich bei durchschnittlich 5 Å stabilisiert. In Abbildung

35B sind die Werte des RNA-Tunnel-Durchmessers und des Interdomänen-Winkels angegeben. Die durchschnittlichen RNA-Tunnel-Abstände von 21 Å und Interdomänenwinkel von 74° weisen auf eine Verschiebung der Konformation in Richtung der geöffneten Form der Polymerase hin. Welchen Einfluss die einzelnen Bereiche des Proteins auf diese Verschiebung haben, kann durch die Analyse der Durchschnittsstrukturen untersucht werden.

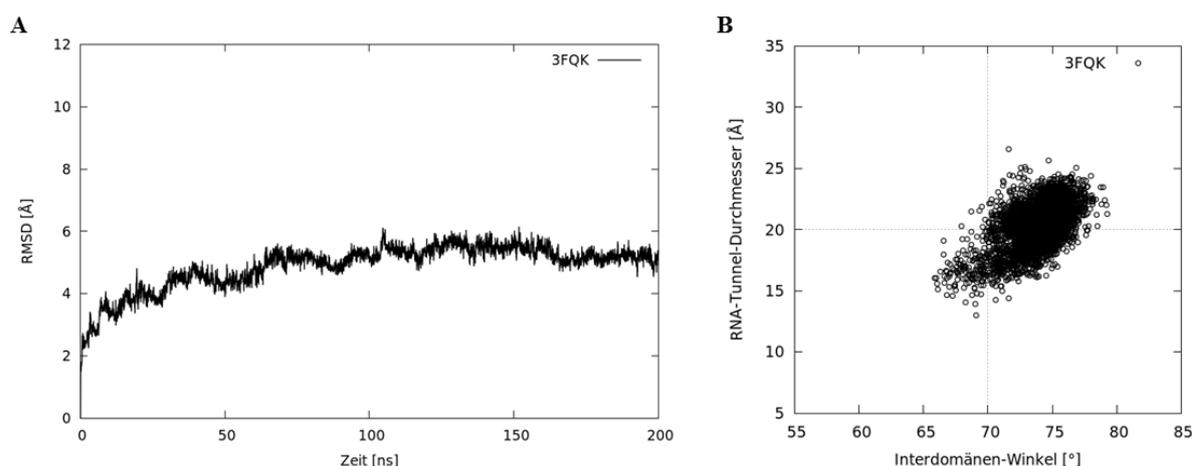


Abbildung 35: Überblick über die Simulation von NS5B mit HCV-796.

A) RMSD-Wert des Proteinrückgrats der Atome N, C α und C über die Simulationszeit der Dynamik von 3FQK. B) Auftragung des RNA-Tunnel-Durchmessers aufgetragen gegen den Interdomänen-Winkel von NS5B.

Für die Abbildung 36 wurden die Durchschnittsstruktur der MD-Simulation und die Kristallstruktur von 3FQK übereinandergelegt, verglichen und die RMSF-Werte bestimmt. Größere Unterschiede finden sich vor allem in der Daumen-Domäne innerhalb der Aminosäuren 401 bis 406 und 491 bis 522, aber auch innerhalb des Linkers zwischen den Aminosäuren 551 und 556 (im Diagramm markiert). Die veränderten Koordinaten des Loops mit den Aminosäuren 401 bis 406 bewirken eine Vergrößerung des RNA-Tunnel-Durchmessers, was bereits in Abbildung 35B dargestellt wurde. Dabei erfolgte die Berechnung anhand der Aminosäuren Met139 und Val405 (in Abbildung 36 als Kalottenmodelle dargestellt), wobei Valin seine Position zusammen mit dem Loop ändert, während Methionin unverändert bleibt.

Die Bewegungen der Daumen-Domäne, insbesondere der Helix 28 (Aminosäuren 497-514), resultieren in einer Rotation des Daumens, was eine Vergrößerung des Interdomänen-Winkels zur Folge hat. Diese Rotation in Anwesenheit von HCV-796 ist vergleichbar mit der Dynamik von NS5B mit GTP, reicht aber nicht für eine vollständige Öffnung der Polymerase. Der

β -Loop und der Linker blockieren in der Durchschnittsstruktur immer noch das aktive Zentrum und den dsRNA-Ausgang.

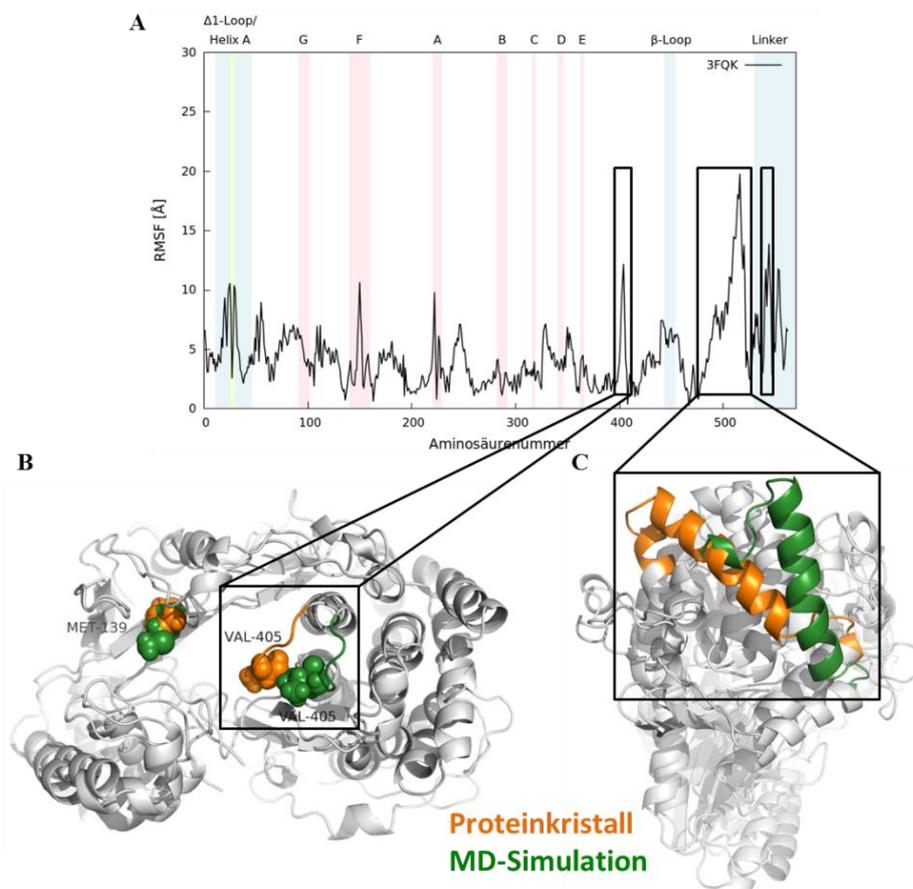


Abbildung 36: Vergleich der Durchschnittsstruktur der Simulation von 3FQK mit der Kristallstruktur. **A)** RMSF-Werte aus dem Vergleich der Durchschnittsstruktur mit der Kristallstruktur. Große Unterschiede sind eingrahmt. **B)** *Cartoon*-Darstellung der Polymerase mit Blick auf den RNA-Tunnel. Der Loop zwischen den Aminosäuren 401 und 406 ist farblich markiert. Die Aminosäuren Met139 und Val405 sind als Kugeln dargestellt. **C)** *Cartoon*-Darstellung der Polymerase mit Blick auf die Daumen-Domäne. Die Aminosäuren 491 bis 522 sind entsprechend eingefärbt. Die Durchschnittsstruktur ist dunkelgrün, die Kristallstruktur ist orange gefärbt.

Interaktionen von HCV-796 mit NS5B

Der Inhibitor HCV-796 selbst zeigt keine auffälligen Konformations- oder Positionswechsel während der Moleküldynamik. Die Wasserstoffbrücken-Bindungen zwischen HCV-796 und NS5B sind in Abbildung 37 dargestellt. Die Interaktion mit Arg200 ist dabei am stabilsten. Im Verlauf der Simulation bildet HCV-796 weitere Wasserstoffbrücken mit Asn316, welches benachbart zum Motiv C liegt, sowie mit Ser365 (Motiv E), Ser367, Ser368, Arg386, Asn411, Tyr448 (β -Loop) und mit Trp550 (Linker). Damit stimmen diese Ergebnisse weitestgehend mit den Untersuchungen von Hang *et al.*⁷⁷ und Howe *et al.*¹⁹⁵ überein. Diese

beschreiben die Bindestelle von HCV-796 mit den Aminosäuren Arg200, Leu314, Asn316, Ile363, Ser365, Ser368 und Met414.

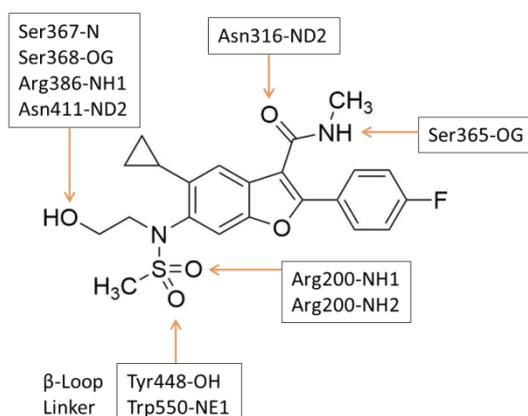


Abbildung 37: Wasserstoffbrücken-Bindungen zwischen NS5B und dem Inhibitor HCV-796 in der Simulation mit 3FQK.

Die Wasserstoffbrückenbindungen sind durch Pfeile vom Wasserstoff-Donor- zum Wasserstoff-Akzeptor-Atom gekennzeichnet.

Die inhibitorische Wirkung von HCV-796 begründet sich offensichtlich nicht darin, NS5B in einer starren Konformation zu belassen. Zumindest zeigt die Simulation, dass die Rotation der Daumen-Domäne und die Vergrößerung des RNA-Tunnels eintreten können. Andererseits bewirken die Wechselwirkungen mit den Aminosäuren Tyr448 und Trp550 eine Bewegungseinschränkung des β -Loops und des Linkers. Ein weiterer Grund für die hohe biologische Aktivität von HCV-796 liegt in der Art der Bindung an NS5B. Der Inhibitor liegt in einer Tasche, welche in den Kristallstrukturen ohne NNI-4 Substrate nicht existiert. Dabei befindet sich HCV-796 in einer Scharnierregion des Proteins, welche die Hand- und die Daumen-Domäne miteinander verbindet (Aminosäuren 356-375). Bei der Öffnung der Polymerase bewegt sich diese Loop-Struktur zwischen den Aminosäuren 356 und 369 auf die gegenüberliegende Helix (Aminosäuren 195-210) der Hand-Domäne zu. In diesem Zwischenraum bindet HCV-796 und fungiert dadurch als Stopper, sodass die Daumen-Domäne keine vollständige Rotation ausführen kann, (dargestellt in Abbildung 38). Des Weiteren reicht HCV-796 bis in die NTP-Binderegion des Proteins hinein (Interaktion mit Motiv C und E durch Asn316, Ser365) und würde im geöffneten Zustand des Proteins mit den Nukleotiden in Konkurrenz treten und somit den Verdopplungsmechanismus der RNA behindern, wie es auch von Davis *et al.* und Reich *et al.* angenommen wurde.^{67,163}

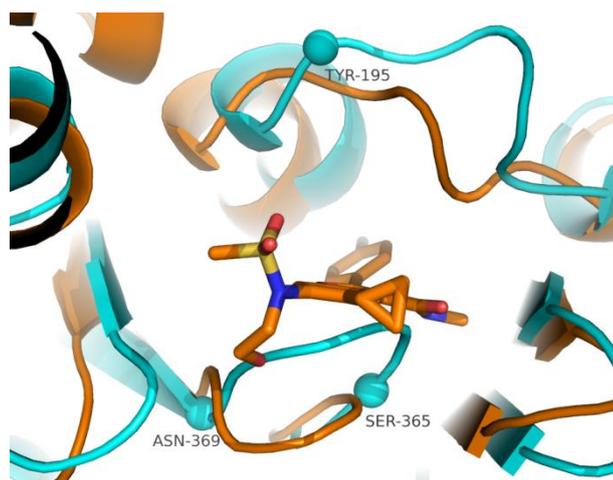


Abbildung 38: Vergleich der Bindestelle von HCV-796 aus den Simulationen mit 2XI3 und 3FQK. Zu sehen ist die Bindestelle des Inhibitors HCV-796 aus der Durchschnittsstruktur der MD mit 3FQK (orange). Darüber gelegt ist die Durchschnittsstruktur der Simulation von 2XI3 mit GTP (cyan). Die Proteine sind in der *Cartoon*-Darstellung abgebildet. Die Bewegung des Loops zwischen den Aminosäuren Ser365 und Asn369 zu der Helix zwischen den Aminosäuren Pro195 und Ser210 wird durch den Inhibitor behindert (*Stick-Präsentation*). Die Wasserstoffatome des Inhibitors sind aus Gründen der Übersichtlichkeit nicht gezeigt.

4.3.2. MD-Simulationen mit einem NNI-2-Inhibitor

Neben der Simulation von NS5B mit dem NNI-4 HCV-796 wurden auch andere allosterische Bindestellen der Nicht-Nukleosid-Inhibitoren untersucht. Die Inhibitoren der Gruppe NNI-2 assoziieren an eine Stelle des Proteins, welche sich ca. 30 Å entfernt vom aktiven Zentrum und innerhalb der Daumen-Domäne befindet (vgl. Abbildung 39A). Als repräsentative Strukturen für NNI-2 wurden die PDB-Einträge 3UDL¹⁹⁶ und 3CJ2¹⁹⁷ ausgewählt. Sie besitzen eine Auflösung von 2,17 Å und 1,75 Å und Qmean6-Werte von 0,784 und 0,798 (siehe Tabelle 7 im Anhang). Die Inhibitoren sind in Abbildung 39C und D abgebildet. Bei KLI (aus 3UDL) handelt es sich um ein Chinolon- und bei SX3 (aus 3CJ2) um ein Anthranilsäureamid-Derivat. Dabei wurden zwei Inhibitoren mit einer hohen (KLI, $IC_{50}=0,014 \mu\text{M}$) bzw. einer niedrigen biologischen Aktivität (SX3, $IC_{50}=17 \mu\text{M}$) gewählt.^{196,197} In Abbildung 39B ist die Bindestelle von KLI und SX3 dargestellt. Der Inhibitor SX3 bildet Wasserstoffbrücken mit der Seitenkette von Arg501 aus und KLI mit den Amidgruppen des Proteinrückgrats von Ser476 und Tyr477. In Abbildung 39B nicht hervorgehoben sind π - π -Interaktionen zwischen Trp528 und dem Phenylring von KLI bzw. dem aromatischen Ring von SX3, sowie weitere hydrophobe Interaktionen der Inhibitoren mit den Aminosäureresten Leu419, Ile482 oder Met423.⁷⁴

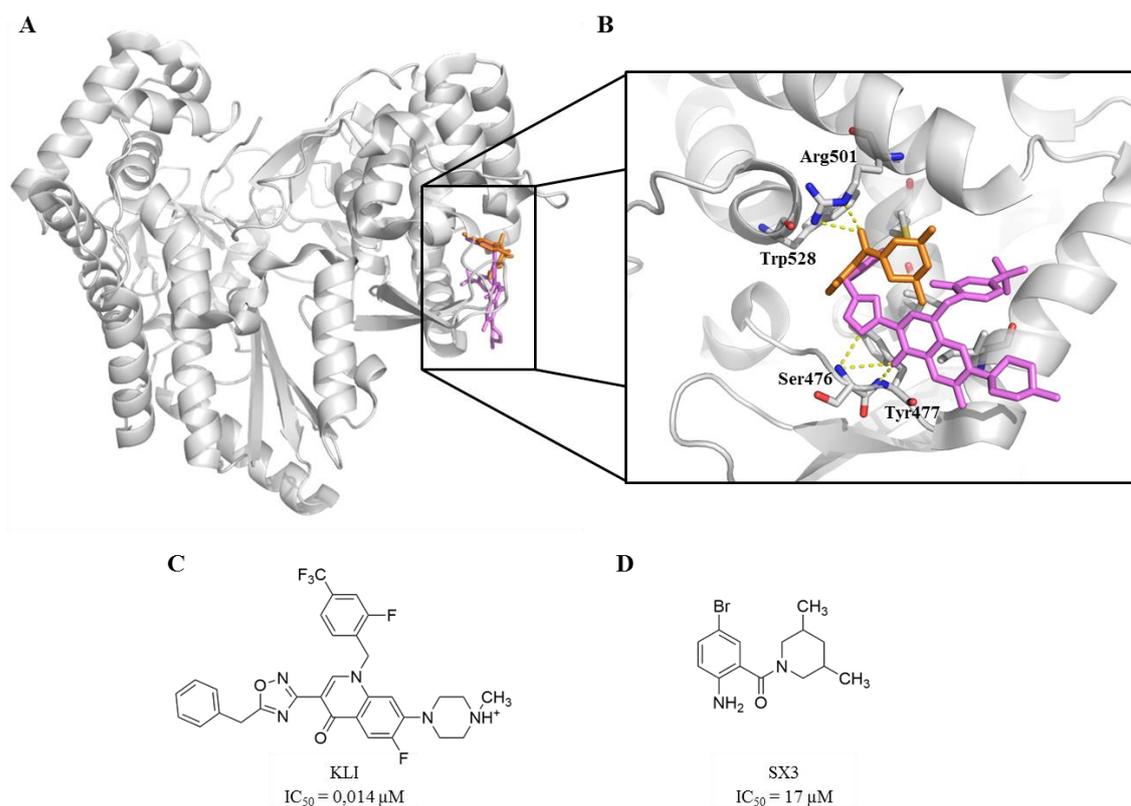


Abbildung 39: Bindestelle der verwendeten NNI-2-Inhibitoren im Protein und deren Molekularstruktur. **A)** NS5B ist in grau und die NNI-2-Inhibitoren sind als *Sticks* dargestellt. KLI ist violett gefärbt und SX3 orange. **B)** Detailaufnahme der Bindetasche der NNI-2. Die Wasserstoffbrücken sind mittels gelb gestrichelter Linien dargestellt. **C)** Molekularstruktur von KLI aus 3UDL und **D)** von SX3 aus 3CJ2.

In Abbildung 40A sind die RMSD-Werte der Simulationen von 3UDL (mit KLI) und 3CJ2 (mit SX3) abgebildet. Die RMSD-Werte der Simulation von 3UDL liegen bei durchschnittlich 5,3 Å und für 3CJ2 bei 4,0 Å. Die Moleküldynamik von 3CJ2 wurde nach 100 ns abgebrochen, da sich der Ligand während der Simulation vom Protein löste und sich immer weiter von diesem entfernte. Ist der Inhibitor nicht mehr an das Protein assoziiert, wird er auf Grund der Simulationsbedingungen (implizit, keine Randbedingungen) von dem Protein weggetrieben.

Einen weiteren Aufschluss über die Konformationsveränderungen der Polymerase während der MD bietet die Berechnung des RNA-Tunnel-Durchmessers und des Interdomänen-Winkels (siehe Abbildung 40B). Für 3CJ2 bleiben der Durchmesser und der Winkel über die Zeit der Simulation relativ stabil bei 21,9 Å und 70,5° und liegen damit knapp über der Grenze zur geöffneten Konformation. Die zwei Anfangswerte von 3CJ2 wurden in Abbildung B durch ein grünes Kreuz markiert, das knapp unterhalb dieser Grenze liegt (18,8 Å und 67,8°). Ähnlich wie bei der Simulation von 3FQK scheint es eine Verschiebung in Richtung der geöffneten Konformation zu geben. Allerdings oszilliert der Interdomänen-Winkel von

3CJ2 über die gesamte Zeit der Simulation sehr stark um den Grenzwert von 71° , sodass eine stabile geöffnete Konformation nicht ausgebildet wird. In der MD-Simulation von 3UDL zeigen sich deutlich größere Veränderungen der Messwerte. Die Startwerte sind als blaues Kreuz gekennzeichnet ($18,6 \text{ \AA}$ und $70,0^\circ$). Über die Simulationszeit betrachtet ist eine Verschiebung in Richtung der geschlossenen Konformation zu beobachten (die Mittelwerte liegen bei $18,8 \text{ \AA}$ und $66,0^\circ$), wobei sich gegen Ende der Simulation der RNA-Tunnel-Durchmesser verkleinert (die Endwerte sind $12,0 \text{ \AA}$ und $64,9^\circ$). Damit zeigen 3CJ2 und 3UDL eine gegensätzliche Beeinflussung des Interdomänenwinkels und des RNA-Tunnel-Durchmessers.

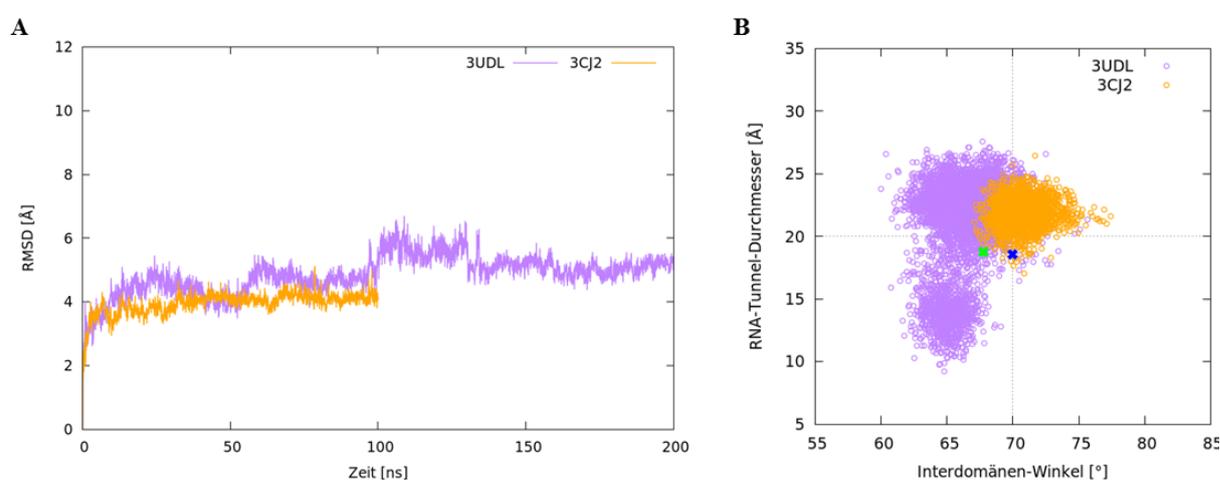


Abbildung 40: Überblick über die Simulationen von NS5B mit NNI-2.

A) RMSD-Wert für die MD-Simulationen von 3UDL (mit KLI) in violett und 3CJ2 (mit SX3) in orange. **B)** Auftragung des RNA-Tunnel-Durchmessers gegen den Interdomänen-Winkel mit der gleichen Farbkodierung. Die Startwerte von 3CJ2 sind durch ein grünes Kreuz gekennzeichnet, die Startwerte von 3UDL durch ein blaues.

Die Berechnung der RMSF-Werte erfolgte anhand der Durchschnittsstrukturen der MD-Simulationen und der jeweiligen Kristallstrukturen von NS5B. Diese sind in Abbildung 41 dargestellt. Diese sind in vielen Bereichen der Proteine vergleichbar. Für 3UDL finden sich signifikante Unterschiede zur Kristallstruktur zwischen den Aminosäuren 403 bis 408 (Loopstruktur am RNA-Tunnel), 440 bis 453 (β -Loop), 531 bis 533 und 547 bis 557 (Linker). Die Bewegungen im Bereich der Aminosäuren 403 bis 408 gehen mit einer Verkleinerung des RNA-Tunnel-Durchmessers einher (vgl. Abbildung 40B). Auch wenn der β -Loop und der Linker größere Unterschiede zur Kristallstruktur aufweisen und nicht mehr so tief in das aktive Zentrum der Polymerase hineinreichen (vgl. Abbildung 41B), bleibt NS5B in einem geschlossenen Zustand. Die fehlende Rotation der Daumen-Domäne und die fehlende

Bewegung der Helix A innerhalb des $\Delta 1$ -Loops verhindern den Übergang zur geöffneten Konformation.

Im Vergleich der Durchschnittsstruktur von 3CJ2 zur entsprechenden Kristallstruktur finden sich Differenzen in den folgenden Abschnitten: zwischen 148 bis 151 (Motiv F), 402 bis 405 (Loopstruktur des RNA-Tunnels) und 513 bis 544 (Daumen-Domäne und Linker). Die Flexibilität der Aminosäuren 402 bis 405 führt zur Fluktuation des RNA-Tunnel-Durchmessers. Besonders starke Strukturunterschiede finden sich in der Daumen-Domäne und im Linker (513-544), was in Abbildung 41C dargestellt ist. Diese Strukturmerkmale werden von der veränderten Position des Inhibitors SX3 hervorgerufen, worauf in den nächsten Abschnitten näher eingegangen wird. Die Positionsunterschiede der Aminosäuren im Motiv F (148-151) in beiden Simulationen sind auf die intrinsische Beweglichkeit dieser Aminosäuren zurückzuführen, da diese bereits in den Kristallstrukturen nicht aufgelöst werden konnten und für die MD-Simulation modelliert werden mussten.

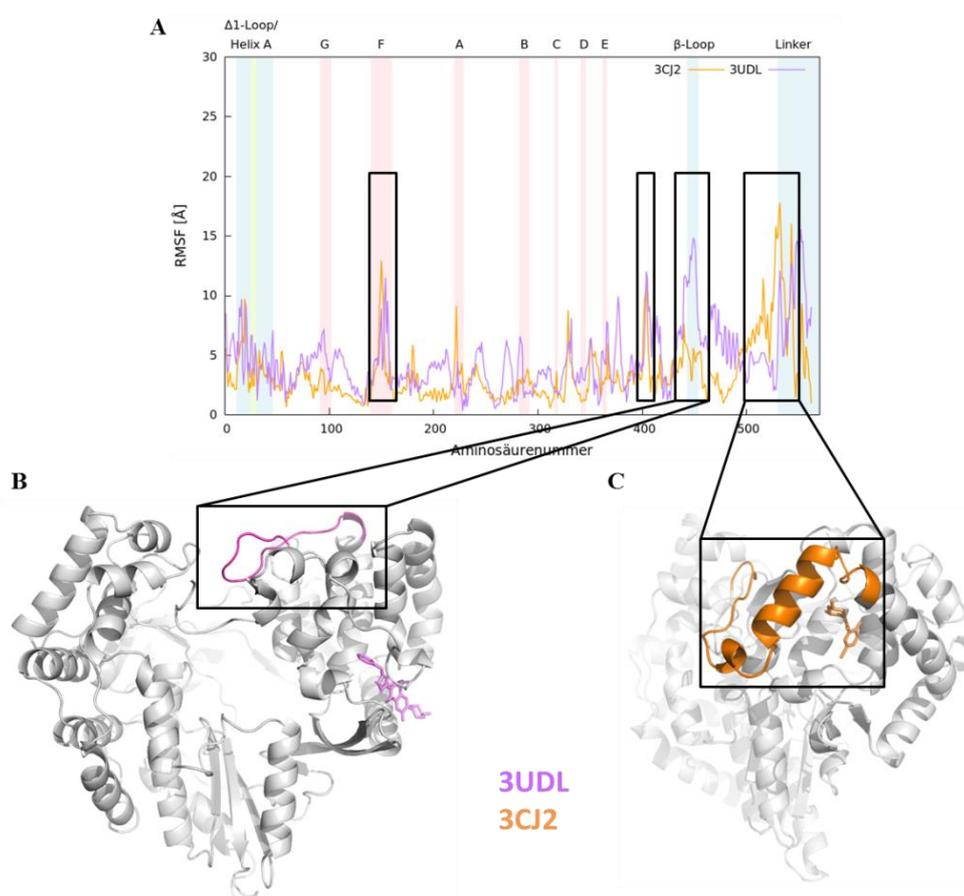


Abbildung 41: Vergleich der NNI-2-Durchschnittsstrukturen mit den Kristallstrukturen.

A) RMSF-Werte der Simulationen mit 3CJ2 in orange und 3UDL in violett. **B)** Durchschnittsstruktur aus der MD mit 3UDL als *Cartoon*-Darstellung mit Sicht auf den dsRNA-Ausgang. Teile des β -Loops (440-453) und der Inhibitor sind violett eingefärbt. **C)** Durchschnittsstruktur aus der MD mit 3CJ2 als *Cartoon*-Darstellung mit Blick auf die NNI-2-Bindestelle. Teile der Daumen-Domäne und des Linkers (513-544), sowie der Inhibitor SX3 sind orange gefärbt.

Interaktionen von KLI mit NS5B

Mit einem relativen Molekulargewicht (MW) von 596,58 g/mol und einem $\log P$ von 6,42 handelt es sich bei KLI um ein sehr großes und lipophiles Molekül, das 7 mögliche Wasserstoffbrücken-Donor oder Akzeptoratome besitzt.¹⁹⁸ In der Kristallstruktur werden zwei Wasserstoffbrücken zu den Resten Ser476 und Tyr477 sowie viele hydrophobe Wechselwirkungen zu Leu419, Met423, Ile482, Val485, Ala486, Leu489 und Leu497 beschrieben.^{74,199} Während der MD-Simulation kommt es allerdings zu einer Repositionierung des Inhibitors (siehe Abbildung 42). Durch die Rotation der die Tasche begrenzenden Seitenketten Arg422, Leu497 und Trp524 verändert sich die Form der Bindestelle und KLI kann noch tiefer in das Protein hineinreichen (vgl. Abbildung 42B, C und D). Die Aminosäuren Met426, Ile463, Phe472, Tyr524 und Leu534 stellen nun die neue Barriere der Bindetasche dar. Dieser Positionswechsel wird durch die Ausbildung von zwei neuen Wasserstoffbrücken mit Asn483 und Lys531 begünstigt, sowie durch die hydrophoben Wechselwirkungen mit Ile424 (siehe Abbildung 42A). Im Vergleich zu der Kristallstruktur findet sich keine Interaktionen mit der Aminosäure Leu419.

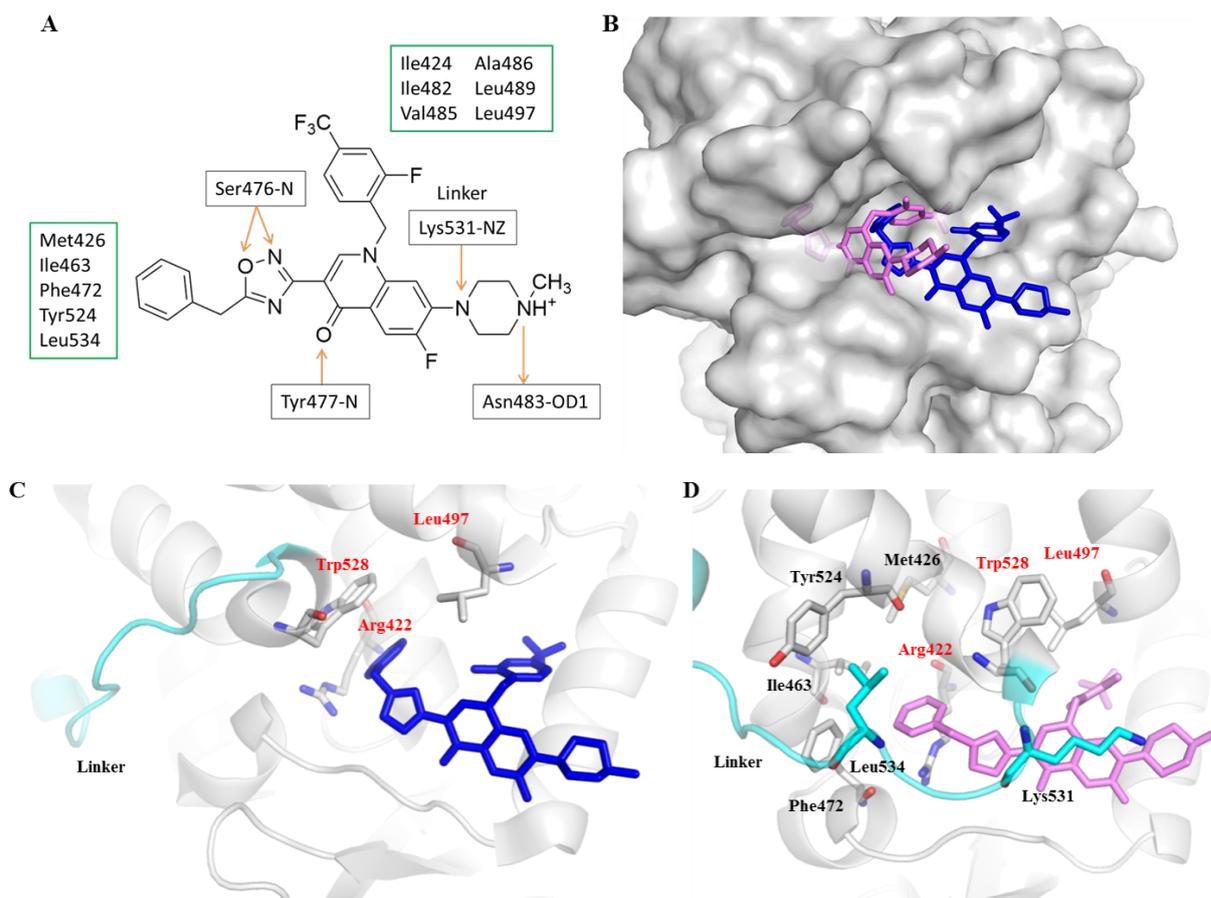


Abbildung 42: Repositionierung des Inhibitors KLI während der MD-Simulation.

A) Die Wasserstoffbrückenbindungen sind durch Pfeile vom Wasserstoff-Donor- zum Wasserstoff-Akzeptor-Atom gekennzeichnet. Die hydrophoben Wechselwirkungen sind grün umrandet. **B)** Vergleichende Darstellung der Positionsänderung von KLI während der Simulation. NS5B (PDB: 3UDL) ist als Connolly-Oberfläche dargestellt mit Blick auf die NNI-2-Bindestelle. Der Inhibitor KLI aus der Kristallstruktur ist in blau dargestellt, KLI aus der MD-Durchschnittsstruktur in violett. **C)** Darstellung der NNI-2-Bindestasche aus der Kristallstruktur 3UDL in der *Cartoon*-Darstellung. Das Protein ist grau gefärbt, der Linker cyan und der Inhibitor KLI ist in blau in der *Sticks*-Darstellung gezeigt. Die Aminosäuren Arg422, Leu497 und Trp528 sind als *Sticks* dargestellt. **D)** Darstellung der NNI-2-Bindestasche aus der MD-Durchschnittsstruktur in der *Cartoon*-Darstellung. Das Protein ist grau gefärbt, der Linker cyan. Der Inhibitor KLI ist violett und wie die Aminosäuren Met426, Arg422, Ile463, Phe472, Leu497, Tyr524, Trp528, Lys531 und Leu534 als *Sticks* dargestellt. Wasserstoffatome sind aus Gründen der Übersichtlichkeit nicht gezeigt.

Der Inhibitor KLI bindet in seiner allosterischen Bindestelle an einen zentralen Punkt der Daumen-Domäne. Hier treffen fünf α -Helices (418-436, 455-467, 478-493, 496-513, 515-530) aufeinander (vgl. Abbildung 43A), mit denen KLI nichtbindende Wechselwirkungen eingeht. Durch die Repositionierung des Inhibitors wird die intrinsische Flexibilität der Daumen-Domäne gestört. Dies trägt zum Inhibierungsmechanismus bei und verhindert eine Umlagerung des Linkers in diesem Bereich und die Rotation der Daumen-Domäne, welche für den Öffnungsmechanismus benötigt wird (vgl. Abbildung 43B). Die Untersuchungen von Boyce *et al.* zeigten ebenfalls, dass eine Deletion des Linkers zu einer Reduktion der NNI-2

Inhibierung führt.²⁰⁰ Die Repositionierung des Inhibitors in der MD-Simulation führt zu einer Zunahme der Interaktionen mit dem Linker und unterstützt somit diese Beobachtung.

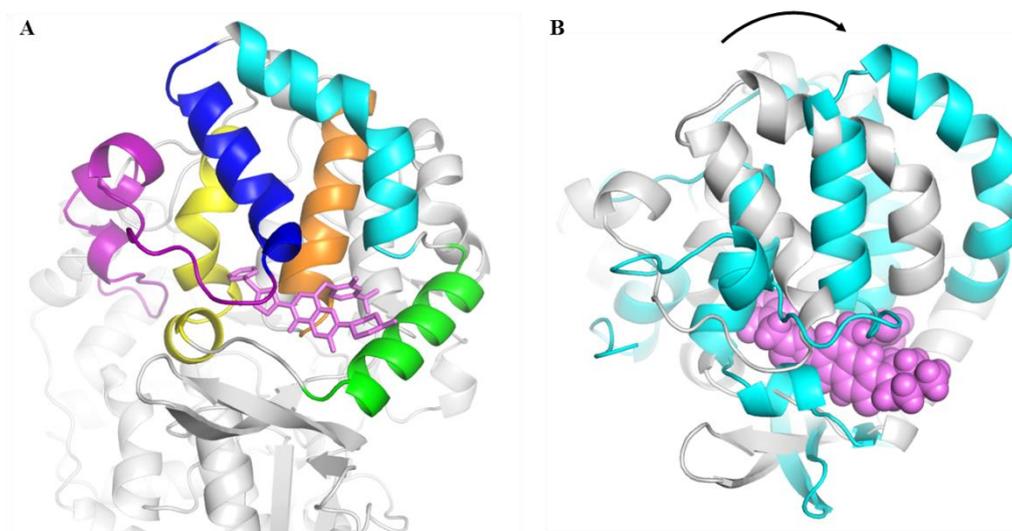


Abbildung 43: Darstellung der NNI-2-Bindestelle im Vergleich der Simulation von 2XI3 und 3UDL.

A) Die linke Abbildung zeigt die MD-Durchschnittsstruktur von 3UDL. Die fünf α -Helices, mit denen KLI interagiert, sind folgendermaßen gefärbt: 418-436 ist orange, 455-467 ist gelb, 478-493 ist grün, 496-513 ist cyan und 515-530 ist blau. Der Linker und der Inhibitor sind violett gefärbt. **B)** Vergleichende Darstellung der MD-Durchschnittsstrukturen in der *Cartoon*-Darstellung von 3UDL in grau und 2XI3 in cyan. KLI ist violett gefärbt und als Kalottenmodell dargestellt. Die Pfeilrichtung deutet die Rotation beim Öffnungsmechanismus an.

Interaktion von SX3 mit NS5B

Wie bereits erwähnt, wurde die MD-Simulation von NS5B mit SX3 nach 100 ns abgebrochen, da der Inhibitor nach ca. 80 ns die NNI-2-Bindestelle verlassen hatte. Bei SX3 handelt es sich um ein sehr kleines (MW = 311,22 g/mol) und hydrophobes ($\log P = 3,25$) Molekül, welches drei Wasserstoffbrücken-Donor- und Akzeptor-Atome besitzt.¹⁹⁸ Während der MD-Simulation interagiert der Inhibitor SX3 sporadisch mit Met426 und Arg508 (vgl. Abbildung 44A und Tabelle 11 im Anhang). Die Wasserstoffbrücken mit Arg501 und Trp528 sowie die hydrophoben Wechselwirkungen mit der Polymerase gehen im Zuge des Positionswechsels von SX3 verloren. In Abbildung 44B ist sehr gut zu erkennen, dass sich der Inhibitor aus seiner eigentlichen Bindestelle herausdreht und sich zwischen zwei α -Helices der Daumen-Domäne lagert. Dabei wird die α -Helix (Aminosäuren 518-530) aus ihrer eigentlichen Position verdrängt und bildet anschließend einen Winkel von 95° zur vorhergehenden α -Helix (Aminosäuren 497-514), welcher im Kristall 45° beträgt. Durch diese Verschiebung vergrößert sich die Bindestelle des Inhibitors, was zu einer verschlechterten Interaktionsmöglichkeit für SX3 mit dem Protein führt.

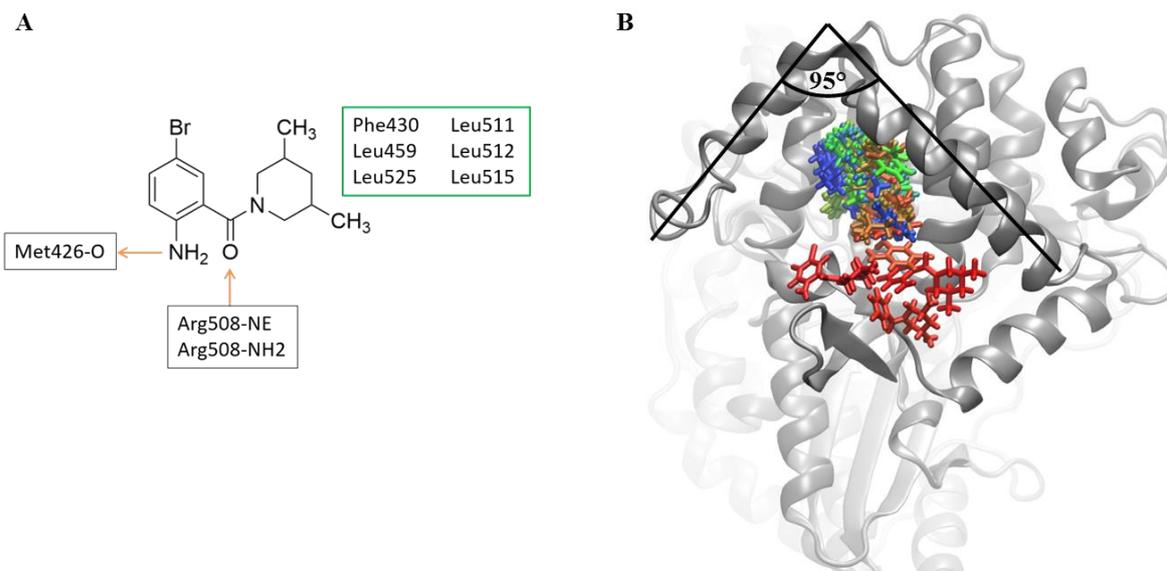


Abbildung 44: Positionswechsel und nichtbindende Wechselwirkungen von SX3 während der MD-Simulation.

A) Die Wasserstoffbrückenbindungen sind durch Pfeile vom Wasserstoff-Donor- zum Wasserstoff-Akzeptor-Atom gekennzeichnet. Die hydrophoben Wechselwirkungen sind grün umrandet. **B)** Die MD-Durchschnittsstruktur von 3CJ2 ist als graue *Cartoon*-Darstellung gezeigt. Der Inhibitor SX3 ist als *Sticks* gezeigt. Die Farbcodierung von SX3 entspricht dem zeitlichen Verlauf der MD-Simulation von rot über gelb und grün nach blau.

Vergleich der Inhibierungsmechanismen der NNI-2

Die beiden Inhibitoren KLI und SX3 besitzen zwar die gleiche allosterische Bindestelle, scheinen aber unterschiedliche Wechselwirkungen mit dem Protein einzugehen. Trotzdem zeigt sich in der MD-Simulation, dass beide Moleküle die vollständige Öffnung der Polymerase verhindern. In einer Arbeit von Davis *et al.*²⁰¹ wurde eine Moleküldynamik-Simulation von NS5B mit dem NNI-2 Inhibitor VGI (Isochinolin-Derivat) aus der Kristallstruktur 2WHO²⁰² durchgeführt, was eine hypergeschlossene Konformation von NS5B zur Folge hatte. Die Simulation wurde mit dem NNI-2 Liganden PFI (Dihydropyron-Derivat) aus 2HAI²⁰³ wiederholt, welche aber zu keiner hypergeschlossenen Konformation führte.¹⁶³ Aus dem unterschiedlichen Bindungsverhalten der Inhibitoren schlossen die Autoren auf eine unterschiedliche inhibitorische Aktivität der Substanzen. Diese Annahme konnte in Bezug auf die unterschiedlichen Auswirkungen auf die Proteinstruktur in den Simulationen mit 3UDL (KLI) und 3CJ2 (SX3) bestätigt werden. Weiterhin beschreiben Davis *et al.*, dass der aktivere Inhibitor die Polymerase in einem sehr geschlossenen Zustand hält.¹⁶³ Diese Beobachtung konnte in der hier beschriebenen Simulation mit 3UDL in Bezug auf den RNA-Tunnel-Durchmesser und den Interdomänen-Winkel gleichfalls festgestellt werden (vgl. Abbildung 40B).

4.3.3. MD-Simulationen mit einem NNI-3-Inhibitor

Die Bindestelle NNI-3 befindet sich in der Hand-Domäne in der Nähe des aktiven Zentrums der Polymerase. In der Literatur wurden viele unterschiedliche Substanzklassen für diese Bindestasche beschrieben. Ein Grund hierfür ist die große Ausdehnung dieser Bindestelle, die gleichfalls der Ort ist, an dem die Bildung des RNA-Doppelstrangs erfolgt und dieser aus der Polymerase heraustritt. Für die Simulation von NS5B mit einem Inhibitor der Bindestelle NNI-3 wurden die Kristallstrukturen 3U4O und 3U4R als repräsentative Strukturen ausgewählt. Diese wurden von Anilkumar *et al.*²⁰⁴ unter den gleichen Bedingungen kristallisiert und besitzen eine Auflösung von 1,8 Å und 2,0 Å. Die Inhibitoren von 3U4O und 3U4R wurden in den PDBs mit den Kürzeln O8E und O8F bezeichnet und sind in Abbildung 45 aufgezeigt. Sie besitzen einen vergleichbaren IC₅₀-Wert von 0,053 µM für O8E aus 3U4O und 0,039 µM für O8F aus 3U4R.

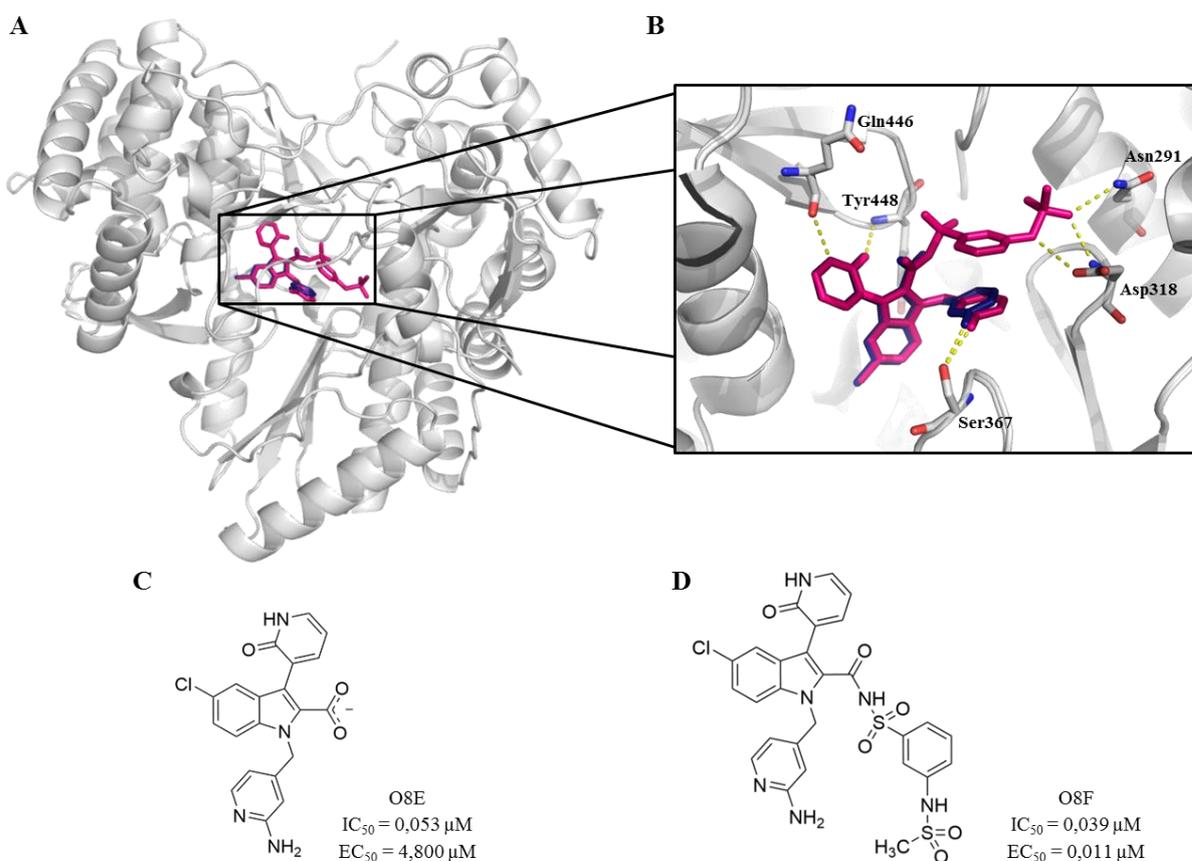


Abbildung 45: Bindestelle und Molekularstruktur der in den Simulationen verwendeten NNI-3. **A)** Darstellung von NS5B in grau als *Cartoon* mit Blick auf den NTP-Eingang. **B)** Detailaufnahme der NNI-3-Bindestelle mit O8E und O8F. **C)** Molekularstruktur von O8E aus 3U4O und **D)** von O8F aus 3U4R.

Die Simulationen der Komplexe mit 3U4O und 3U4R wurden für 500 ns berechnet. Die RMSD-Werte dieser Simulationen sind in Abbildung 46A zu sehen und liegen innerhalb der letzten 100 ns für 3U4O bei 4,3 Å und für 3U4R bei 4,0 Å. Sie sind damit vergleichbar mit den Simulationen der Apoproteinstrukturen 1NB4 und 2XI2 (siehe Kapitel 4.2.1). Die RNA-Tunnel-Durchmesser und die Interdomänen-Winkel der MD-Simulationen sind in dem Diagramm in Abbildung 46B aufgetragen. Bei beiden Simulationen bleiben die Proteinstrukturen im Bereich einer geschlossenen Konformation von NS5B. Dabei wurden für die Simulation von 3U4O etwas kleinere Werte mit einem durchschnittlichen Winkel von 69° und einem RNA-Tunnel-Durchmesser von 12 Å im Vergleich zu 70° und 15 Å für 3U4R erhalten.

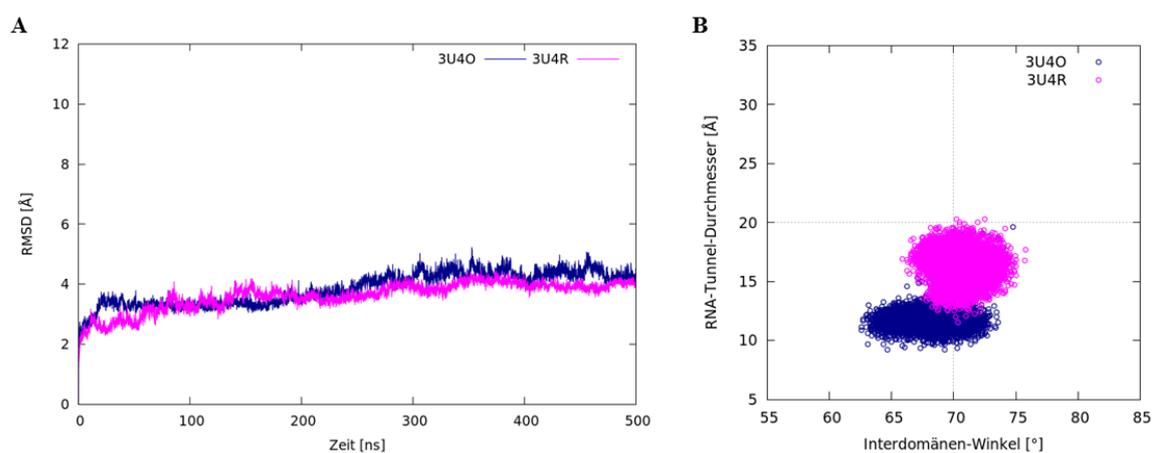


Abbildung 46: Überblick über die Simulationen von NS5B mit NNI-3.

A) RMSD-Werte der Simulationen von 3U4O (O8E) in dunkelblau und 3U4R (O8F) in magenta. **B)** Auftragung des RNA-Tunnel-Durchmessers gegen den Interdomänen-Winkel.

Die letzten 100 ns der MD-Simulationen dienten zur Berechnung der Durchschnittsstrukturen und der RMSF-Werte im Vergleich mit den Kristallstrukturen. Diese Werte sind in Abbildung 47A dargestellt. Im Bereich des Δ 1-Loops und des Linkers zeigen beide Simulationen auffällige Fluktuationen. Für den C-terminalen Abschnitt der Daumen-Domäne finden sich ebenfalls größere RMSF-Werte, welche die größeren Fluktuationen des Interdomänen-Winkels der Simulationen erklären. Im Allgemeinen zeigt die Simulation mit 3U4O in den Bereichen zwischen den Aminosäuren 16 bis 25 (Δ 1-Loop), 386 bis 401 und 505 bis 518 (Daumen-Domäne), sowie zwischen 542 bis 556 und 565 bis 568 (Linker) größere Unterschiede zur Kristallstruktur. Bei 3U4R sind es die Regionen zwischen den Aminosäuren 4 bis 18 (Δ 1-Loop), 148 bis 153 (Motiv F), 328 bis 335 (Hand-Domäne), 515 bis 517

(Daumen-Domäne) und zwischen 528 bis 548 (Linker). Bei 3U4R bildet sich zusätzlich im Bereich des Linkers zwischen den Aminosäuren 540 und 554 eine Helixstruktur aus (vgl. Abbildung 47C). Die überwiegende Anzahl der NS5B-Kristallstrukturen (ca. 70 %) zeigt im Bereich von 540 bis 545 eine helikale Struktur, deren Existenz in der Literatur bisher keine Bedeutung zugesprochen worden ist. Wäre der Linker nicht an das aktive Zentrum assoziiert, könnte sich eventuell eine längere Helix bis zum Membrananker bilden.

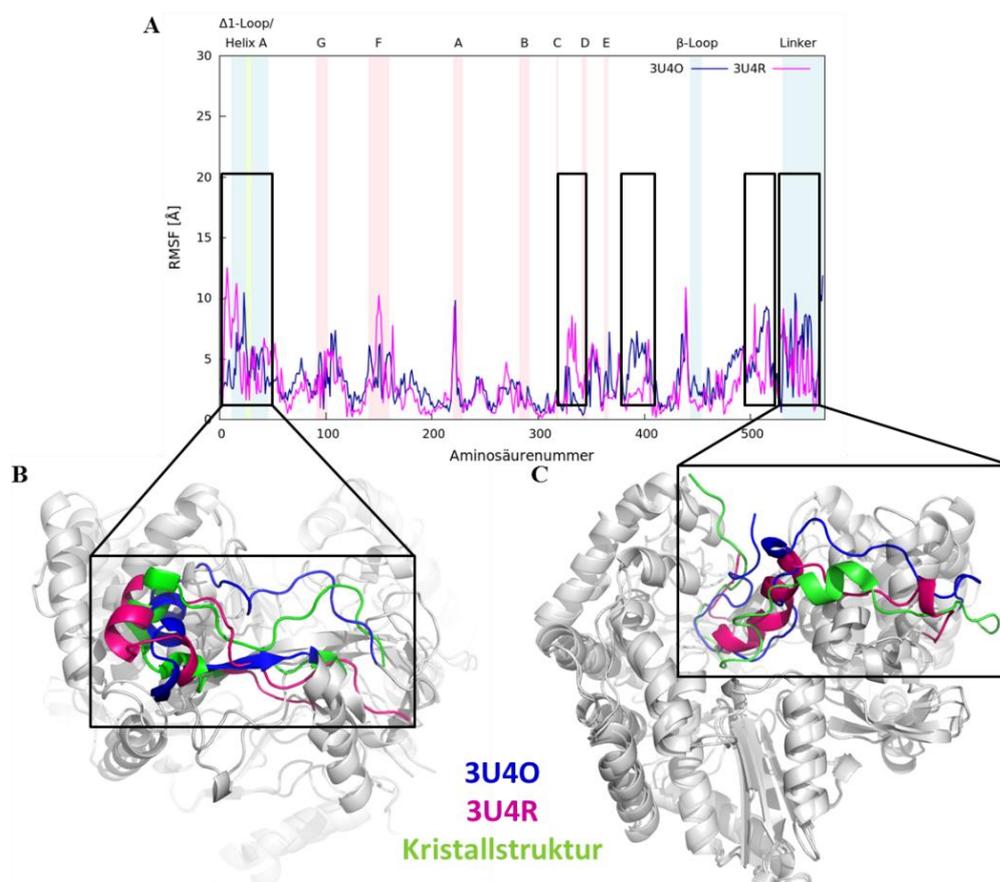


Abbildung 47: Vergleich der NNI-3-Durchschnittsstrukturen mit der Kristallstruktur.

A) RMSF-Werte der Simulation mit 3U4O in dunkelblau und 3U4R in magenta. **B)** Vergleich des $\Delta 1$ -Loops und **C)** des Linkers der Durchschnittsstrukturen von 3U4O und 3U4R mit der Kristallstruktur 3U4O (grün).

Interaktionen von O8E und O8F mit NS5B

Bei den mittels MD-Simulation untersuchten Inhibitoren handelt es sich um Indol-Derivate. Der Inhibitor O8E (3U4O) besitzt einen $\log P$ von 1,52 und ein Molekulargewicht von 394,81 g/mol. Der Inhibitor O8F (3U4R) besitzt eine höhere Hydrophobizität mit einem $\log P$

von 2,95 und ein Molekulargewicht von 627,10 g/mol. Beide Moleküle besitzen einen ähnlichen IC_{50} -Wert von 0,059 μ M für O8E und 0,039 μ M für O8F.

Da sich die Positionen von O8E und O8F in ihren Bindestellen nach 180 ns bzw. nach 100 ns verändern (vgl. Abbildung 48), wurden die Dynamiken jeweils für 500 ns berechnet. Die beschriebenen Translokationen sind in der Abbildung 48A-D anhand der RMSD-Werte des Proteinrückgrats und des Inhibitors (Abbildungen A und B) verdeutlicht sowie anhand der berechneten Wasserstoffbrücken zwischen den Inhibitoren und NS5B (Abbildung 48C und D). In den Abbildungen E und F sind vergleichsweise die Kristallstrukturen und die MD-Durchschnittsstrukturen der letzten 100 ns dargestellt. Dabei verändert sich die Position von O8F nur geringfügig (Abbildung 48F), wohingegen die Position von O8E eine fundamentale Veränderung aufweist (Abbildung 48E).

In der Simulation mit 3U4O bildet die 2-Aminopyridin-Gruppe von O8E mit Arg158 aus Motiv F eine Wasserstoffbrücke aus, die Carboxylgruppe interagiert mit Arg394 und Arg386 und die 2-Pyridon-Gruppe interagiert hauptsächlich mit Arg386 und Tyr448 aus dem β -Loop (Abbildung 48G). Die Translokation von O8E ruft eine Verkleinerung der Bindetasche hervor. Dies wird wiederum durch Interaktionen mit Arg386 und Arg394 bewirkt (Abbildung 48I). Die Bewegung dieser Region zwischen den Aminosäuren 386 und 401 ist bereits aus den RMSF-Werten ersichtlich (Abbildung 47) und trägt zur Verkleinerung des Interdomänen-Winkels bei.

Der Inhibitor O8F ist größer und kann den Innenraum der Bindetasche in der Polymerase besser ausfüllen. Die 2-Aminopyridin-Gruppe interagiert mit Asp220 aus Motiv A und die 2-Pyridon-Gruppe bildet beständige Wasserstoffbrücken mit Gln446 und Gly449 aus dem β -Loop (siehe Abbildung 48H). Die Carboxylgruppe in O8E wurde in O8F durch eine N-({3-[(Methylsulfonyl)amino]phenyl}sulfonyl)-carboxamid-Gruppe ersetzt. Diese interagiert mit Lys141 aus Motiv F, Asp318 aus Motiv C, Gln446 aus dem β -Loop und mit Ser556 und Gly557 aus dem Linker. Die Bindetasche NNI-3 bleibt in diesem Fall weitgehend erhalten, nur Asp220 bildet mit dem Inhibitor eine Wasserstoffbrücke und verändert dadurch geringfügig seine Position (vgl. Abbildung 48J).

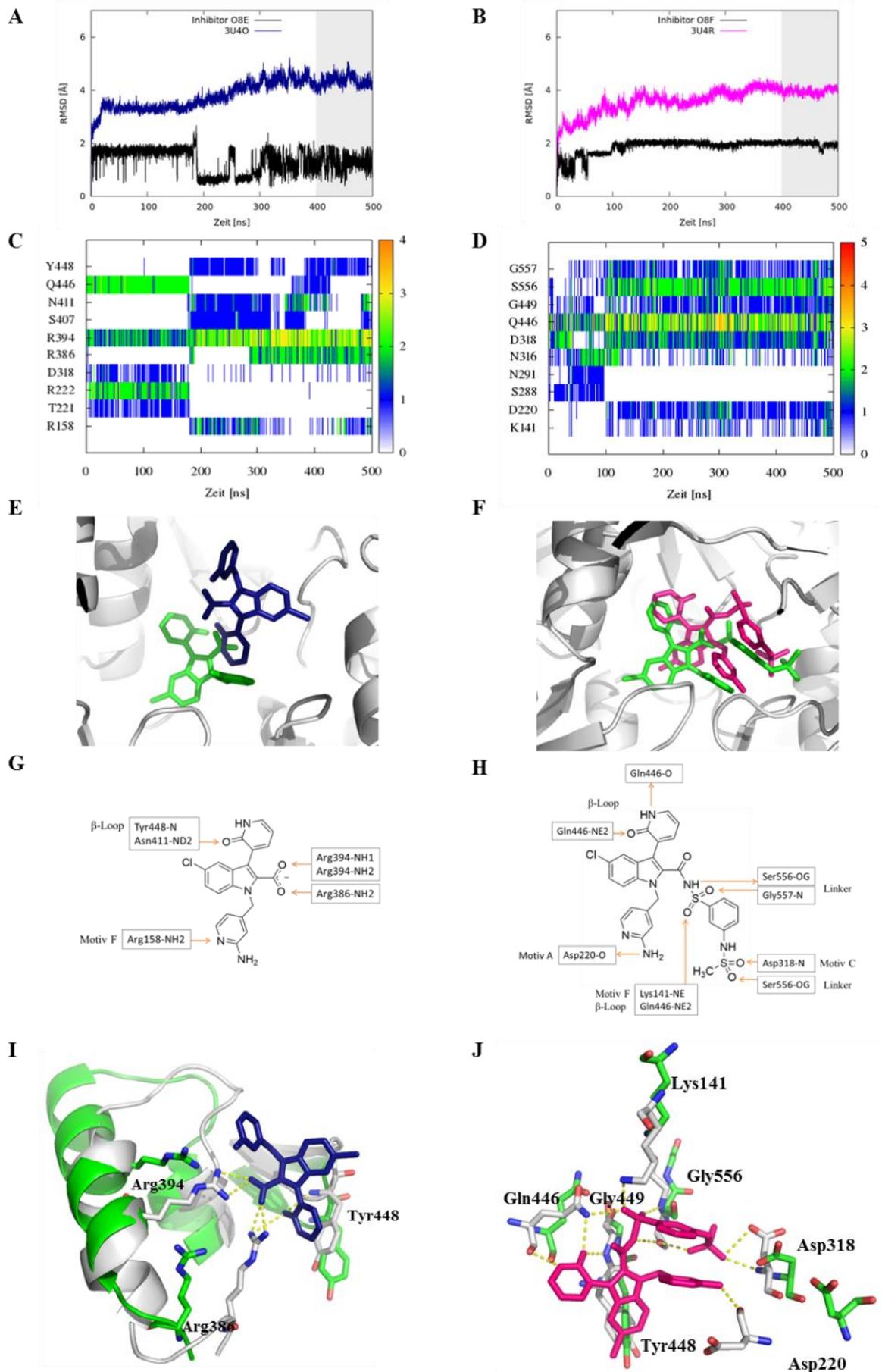


Abbildung 48 (vorherige Seite): RMSD-Werte und Wasserstoffbrücken-Wechselwirkungen zwischen NNI 3 und NS5B aus den Simulationen mit den Kristallstrukturen 3U4O und 3U4R.

A) Vergleich der RMSD-Werte der Proteinstruktur 3U4O und des Liganden O8E. **B)** Vergleich der RMSD-Werte der Proteinstruktur 3U4R und des Liganden O8F. **C)** Wasserstoffbrücken-Wechselwirkungen der Aminosäuren über die Simulationszeit mit O8E. **D)** Wasserstoffbrücken-Wechselwirkungen der Aminosäuren über die Simulationszeit mit O8F. **E)** und **F)** zeigen vergleichende Darstellungen der NNI-3-Positionen zwischen der Kristallstruktur in grau und dem Inhibitor in grün mit O8E (dunkelblau) im Protein aus der MD-Simulation mit 3U4O und O8F (magenta) aus der MD-Simulation mit 3U4R. **G)** Wasserstoffbrücken zwischen NS5B und O8E. **H)** Wasserstoffbrücken zwischen NS5B und O8F. Die Wasserstoffbrückenbindungen sind durch Pfeile vom Wasserstoff-Donor- zum Wasserstoff-Akzeptor-Atom gekennzeichnet. **I)** Darstellung der Kristallstruktur 3U4O in grün. Position von O8E (dunkelblau) im Protein aus der MD-Simulation in grau. **J)** Darstellung der Kristallstruktur 3U4R in grün. Position von O8F (magenta) im Protein aus der MD-Simulation mit 3U4R in grau.

In den Kristallstrukturen besitzen die Inhibitoren O8E und O8F nahezu identische Bindemodi. In der MD-Simulation zeigen sich jedoch unterschiedliche Interaktionen mit denen die Inhibitoren mit der Polymerase wechselwirken. Beide Inhibitoren interagieren über fünf bzw. sechs Aminosäuren mit NS5B, darunter auch mit den Aminosäuren aus den Motiven, welche eigentlich für die NTP-Bindung verantwortlich sind (Motiv A, Motiv C und Motiv F), sowie mit dem β -Loop und dem Linker (Gln446, Tyr 448, Ser556 und Gly557). Zusätzlich zeigt die Polymerase bei Anwesenheit des Inhibitors O8E kleinere Werte des RNA-Tunnel-Durchmessers und des Interdomänen-Winkels. Beide Inhibitoren binden an eine Stelle des Proteins, an welcher der RNA-Doppelstrang die Polymerase verlässt. Eine Störung in der dsRNA-Bildung führt zum Abbruch der Synthese. Vergleicht man die MD-Durchschnittsstrukturen mit der von 2XI3 und 1NB7, so lagern sich die Inhibitoren O8E und O8F zwischen GTP und ssRNA (siehe Abbildung 49) und könnten die Ausrichtung zueinander blockieren.

Davis *et al.*¹⁶³ haben nicht nur MD-Simulationen mit NNI-2 untersucht, sondern auch mit NNI-3. Dabei zeigten sich Veränderungen des Wasserstoffbrückenbindungsmusters zugunsten der Polymerase und den Inhibitoren, da neue Interaktionen mit N291 und Y448 beobachtet wurden.¹⁶³ Die unterschiedlichen Wasserstoffbrücken führten sie auf die unterschiedliche Form der Liganden zurück, obwohl es sich in beiden Fällen um Benzothiadiazin-Derivate handelte. Die inhibitorische Wirkungsweise der untersuchten Inhibitoren schlossen sie aus dem Vergleich mit der Kristallstruktur 1NB7. Dabei zeigte sich, dass die ssRNA daran gehindert wird das katalytische Zentrum zu erreichen.

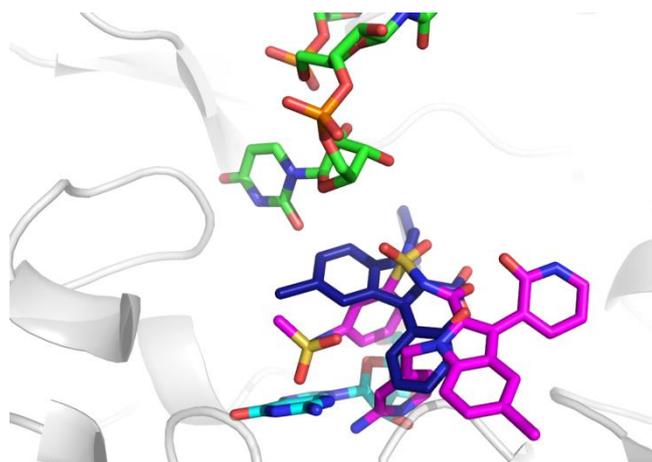


Abbildung 49: Vergleichende Darstellung der Liganden der Durchschnittsstrukturen aus den Simulationen von 1NB7, 2XI3, 3U4O und 3U4R.

Abbildung der Liganden (*Sticks*) der Durchschnittsstrukturen von 1NB7 (ssRNA) in grün, 2XI3 (GTP) in cyan, 3U4O (O8E) in dunkelblau und 3U4R (O8F) in magenta. Die *Cartoon*-Darstellung des Proteins in grau ist aus der MD-Struktur von 2XI3. Wasserstoffatome sind aus Gründen der Übersichtlichkeit nicht gezeigt.

4.3.4. MD-Simulationen mit einem NNI-5-Inhibitor

Die Bindestelle für NNI-5 wurde erstmals 2010 von Cheng *et al.*⁷⁹ in der Kristallstruktur 3LKH detektiert. Insgesamt existieren nur 4 Kristallstrukturen (3LKH, 3QGH, 3QGI und 3VQS) mit Inhibitoren, die an dieser Stelle binden und welche von Gentles *et al.*⁸⁰ veröffentlicht wurden (Stand August 2017). Für die MD-Simulation wurde der PDB-Eintrag 3VQS⁷⁸ ausgewählt. Dieser weist nur wenig fehlende Aminosäuren und eine Auflösung von 1,90 Å auf. Des Weiteren wurde für 3VQS ein Qmean6-Wert von 0,851 errechnet (vgl. Tabelle 7 im Anhang). Diese Kristallstruktur wurde mit einem Piperazin-Derivat co-kristallisiert. Interessanterweise befindet sich der Linker-Abschnitt der Polymerase nicht mehr so tief im aktiven Zentrum, wie bei der Mehrzahl der Kristallstrukturen (in Abbildung 50A magenta gefärbt). Dieser Bereich wird nun durch den Inhibitor eingenommen.

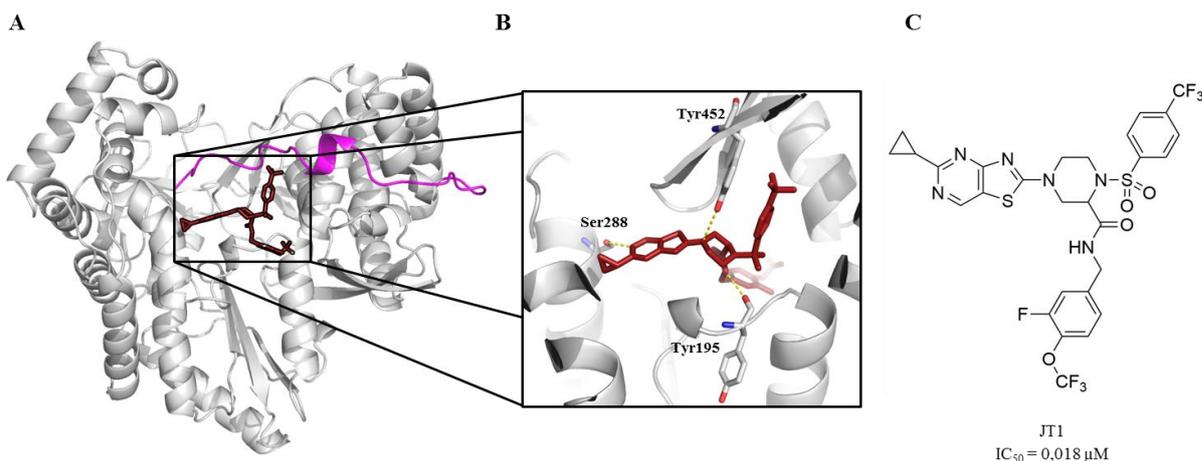


Abbildung 50: Darstellung der Kristallstruktur 3VQS, der Bindestelle und der Molekülstruktur von JT1. **A)** Darstellung von 3VQS in grau. Der Linker ist magenta eingefärbt und JT1 ist braun dargestellt. **B)** Detailaufnahme der Bindestelle von JT1 im Protein. Wasserstoffatome sind aus Gründen der Übersichtlichkeit nicht gezeigt. **C)** Molekülstruktur von JT1.

Die RMSD-Werte der Simulation mit der Kristallstruktur 3VQS sind im Diagramm von Abbildung 51A aufgetragen und wurden für die Atome N, C_α und C berechnet. Die Durchschnittsstruktur der MD-Simulation besitzt eine Abweichung von 5 Å im Vergleich zur Startstruktur. Zusätzlich ist in Abbildung 51B der RNA-Tunnel-Durchmesser gegen den Interdomänen-Winkel aufgetragen. Während der MD-Simulation schwankten die Werte für den Tunnel-Durchmesser um den Grenzwert von 20 Å und um einen Winkel von 70°. Der Vergleich der MD-Simulation mit der Kristallstruktur sollte eine Erklärung für die gemessenen Werte geben.

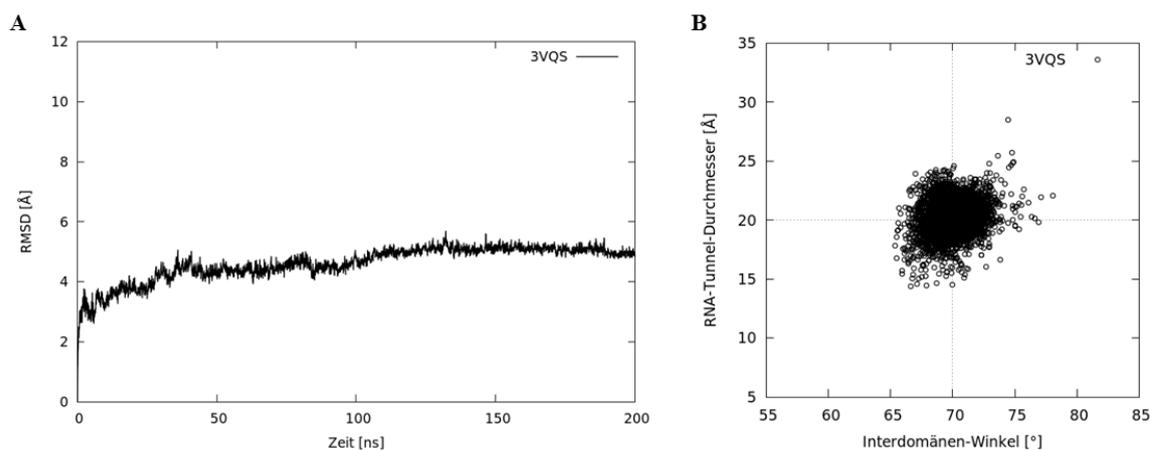


Abbildung 51: Überblick über die Simulation mit der Kristallstruktur 3VQS. **A)** Darstellung der RMSD-Werte über die Simulationszeit der Dynamik mit 3VQS. **B)** Auftragung des RNA-Tunnel-Durchmessers gegen den Interdomänen-Winkel von NS5B aus der Dynamik.

In Abbildung 52 wurden die RMSF-Werte durch den Vergleich der MD-Durchschnittsstruktur mit der Kristallstruktur berechnet. Signifikante Unterschiede befinden sich in den Bereichen der Aminosäuren 7 bis 16 (Δ 1-Loop), 148 bis 156 (Motiv F), 371 bis 380 (Daumen-Domäne) und 531 bis 548 (Linker) (vgl. Abbildung 52A). Die Beweglichkeit des Motivs F ist aus den vorherigen MD-Simulationen bekannt. Jedoch positioniert sich in diesem Fall dieser Bereich in Richtung des RNA-Tunnel-Eingangs, vergleichbar mit der Simulation von 1NB7 (vgl. Abbildung 52B). Die veränderte Position des Abschnitts zwischen den Aminosäuren 371 und 380 in der Daumen-Domäne war in den bisher beschriebenen Simulationen nicht aufgetreten. Dieser Bereich besteht aus zwei antiparallelen β -Faltblättern, welcher den Übergang von der Hand- zur Daumen-Domäne kennzeichnen (vgl. Abbildung 52C). Diese veränderte Position wirkt sich unter anderem auf den schwankenden Wert des Interdomänenwinkels aus.

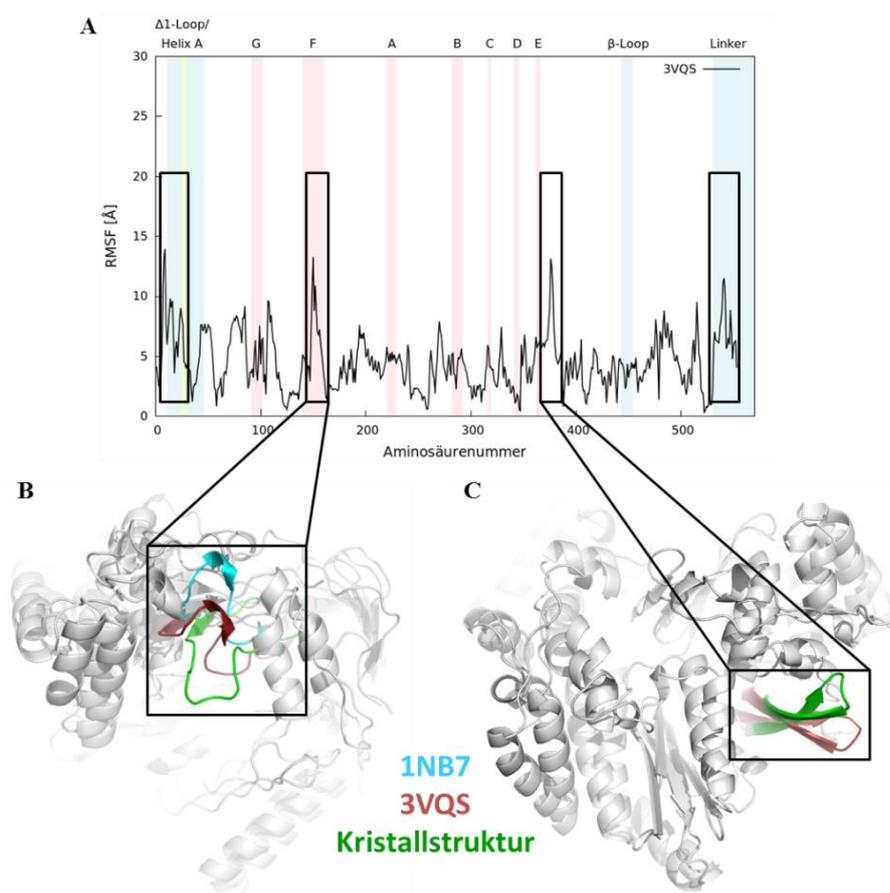


Abbildung 52: RMSF-Werte aus dem Vergleich zwischen der Durchschnittsstruktur der MD-Simulation mit 3VQS und der Kristallstruktur.

A) RMSF-Werte der Durchschnittsstruktur der Simulation von 3VQS im Bezug zur Kristallstruktur. **B)** Darstellung von NS5B mit Blick auf den NTP-Eingang bzw. auf das Motiv F. Die Färbung entspricht 1NB7 in cyan, 3VQS in braun und der Kristallstruktur in grün. **C)** Darstellung von NS5B mit Blick auf den dsRNA-Ausgang.

Interaktionen von JT1 mit NS5B

Der in der Kristallstruktur enthaltene Inhibitor wird als JT1 bzw. als JTK-853 bezeichnet. Dabei handelt es sich um ein sehr großes und hydrophiles Molekül mit einem $\log P$ von 4,53 und einem Molekulargewicht von 704,64 g/mol. Der Inhibitor selbst befindet sich in der Kristallstruktur am dsRNA-Ausgang. Zu den Interaktionsstellen von NNI-5 bzw. JT1 gehören die Aminosäuren Val179, Tyr191, Phe193, Tyr195, Ser288, Met414 und Tyr452.⁷⁸ Zu Beginn der Dynamik zeigt sich ein fundamentaler Positionswechsel von JT1, wie es in Abbildung 53A dargestellt ist. Dies kann ebenfalls anhand der RMSD-Werte des Inhibitors gezeigt werden (siehe Abbildung 53B). Entscheidend trägt dazu die Konformationsveränderung des Piperazin-Rings des Inhibitors aus der Sessel- in die Twist-Konformation bei (siehe Abbildung 53C). Dies geht einher mit der Rotation der Seitenketten von Tyr448, Tyr452 und Phe551, sowie der Verlagerung der Loop-Struktur zwischen den Aminosäuren Met187 und Pro197 in der Hand-Domäne (Abbildung 53D). Der Energieaufwand für den Positionswechsel wird durch zusätzliche Wasserstoffbrückenbindungen ausgeglichen. Der Inhibitor JT1 dreht sich dadurch um 45° und interagiert anschließend über seine Sulfonylgruppe mit Ser196 und Gln199 aus der Hand-Domäne (siehe Abbildung 53E). Weitere Wasserstoffbrücken konnten zu den Aminosäuren Tyr452 (β -Loop), Arg465 und Phe550 (Linker) identifiziert werden. Die Wasserstoffbrücke zu Ser288, welche in der Kristallstruktur vorliegt, ging während der MD-Simulation verloren.

JT1-resistente Zellen zeigten Mutationen von NS5B an den Aminosäuren C316Y, M414T, Y452H und L466V, sowie die Doppelmutationen C316Y Y452H und Y452H L466V.⁷⁸ Dies ist in Bezug auf die Mutationen Y452H und L466V aus der MD-Simulation erklärbar, da JT1 nichtbindende Wechselwirkungen mit Tyr452, Arg465 und Leu466 zeigt.

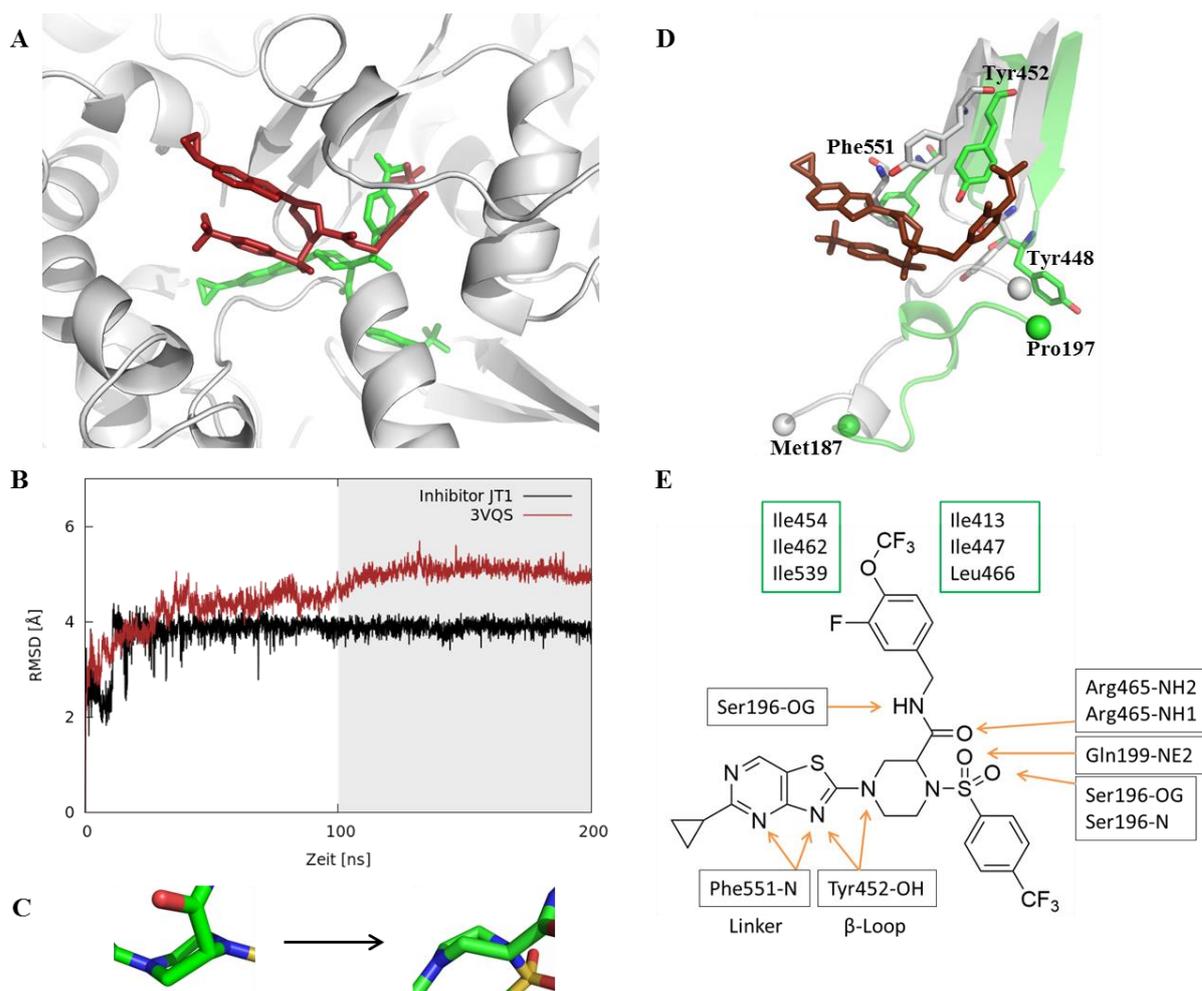


Abbildung 53: Positionswechsel von JT1 in der MD-Simulation mit 3VQS.

A) Positionswechsel des Inhibitors JT1 zu Beginn der Dynamik in grün im Vergleich zur Durchschnittsstruktur der letzten 100 ns in braun. Das Protein ist grau in der *Cartoon*-Darstellung gezeigt. **B)** RMSD-Werte des Proteinrückgrats von NS5B und der Schweratome des Inhibitors JT1. **C)** Abbildung des Inhibitors JT1 mit Fokus auf den Piperazin-Ring. Dieser wechselt die Konformation im Laufe der MD-Simulation von der Sessel- zur Twist-Konformation. **D)** Vergleichende Darstellung der Rotation der Aminosäuren Tyr448, Tyr452 und Phe551, sowie der Loop-Struktur zwischen den Resten Met187 und Pro197. Die grüne Darstellung entspricht der Kristallstruktur 3VQS, die graue bzw. braune der MD-Durchschnittsstruktur. **E)** Wasserstoffbrücken von JT1 mit NS5B während der Simulation. Die Wasserstoffbrückenbindungen sind durch Pfeile vom Wasserstoff-Donor- zum Wasserstoff-Akzeptor-Atom gekennzeichnet. Hydrophobe Wechselwirkungen bestehen mit den grün eingerahmten Aminosäuren.

Durch die nichtbindenden Wechselwirkungen des Inhibitors werden die Hand- und die Daumen-Domäne, sowie der β -Loop und der Linker miteinander vernetzt, was zu einer verminderten Bewegungsfreiheit des Proteins führt. JT1 verschließt so den dsRNA-Ausgang (vgl. Abbildung 54A). Der inhibitorische Mechanismus von JT1 ist zumindest im geschlossenen Zustand der Polymerase erkennbar. Um die inhibitorische Wirkung im geöffneten Zustand zu untersuchen, wurde das NS5B-dsRNA-Modell (siehe Kapitel 4.2.2) mit der MD-Durchschnittsstruktur von 3VQS überlagert. (siehe Abbildung 54B). In diesem Zustand könnten die nichtbindenden Wechselwirkungen weiterhin mit der Hand- und

teilweise mit der Daumen-Domäne bestehen. Würden diese für eine Bindung ausreichen, so würde JT1 im aktiven Zustand der Polymerase den Template-Strang der dsRNA beim Austritt aus dem Protein behindern.

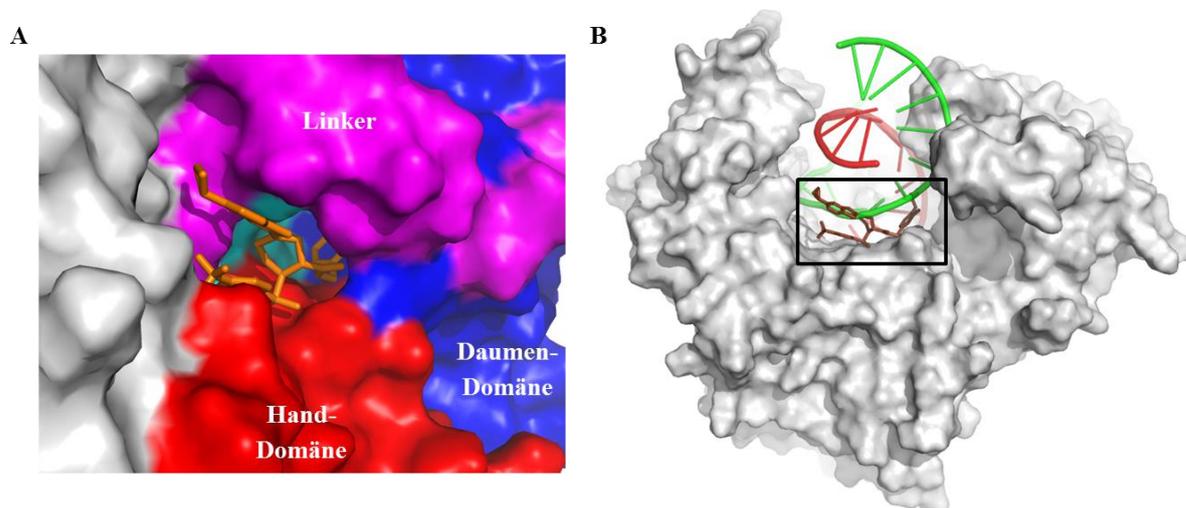


Abbildung 54: Position von JT1 im geschlossenen und geöffneten Protein.

A) Darstellung der MD-Durchschnittsstruktur mit einer Connolly-Oberfläche. Die Finger-Domäne ist grau gefärbt, die Hand-Domäne rot, die Daumen-Domäne blau, der β -Loop cyan und der Linker magenta. Der Inhibitor JT1 ist in der *Stick*-Darstellung und orangefarben dargestellt. B) Abbildung des NS5B-dsRNA-Modells mit dem Inhibitor JT1. Das Protein ist in grau in der Connolly-Oberfläche dargestellt, die RNA als *Cartoon* und JT1 als *Sticks*. Der Template-Strang der RNA ist grün und der Tochterstrang rot gefärbt. JT1 ist braun dargestellt. Die Umrandung hebt die Lage des Inhibitors im geöffneten Protein hervor.

4.4. Vergleich der MD-Simulationen

4.4.1. RMSD-Matrix

Anhand der Moleküldynamik-Simulationen sollte der Einfluss der Liganden und Inhibitoren auf die Struktur von NS5B untersucht werden. Hierfür bietet sich ein Vergleich der RMSD-Werte und der Durchschnittsstrukturen der MD-Simulationen an. Einen guten Überblick gibt die RMSD-Matrix der MD-Durchschnittsstrukturen in Tabelle 3. Im Vergleich zu den Simulationen der Apoproteinstrukturen 1NB4 und 2XI2 zeigt 3U4O gefolgt von 3U4R und 3CJ2 die geringsten Unterschiede. Die Inhibitoren der Gruppe NNI-3 scheinen demnach das Protein in einen apo-ähnlichen Zustand zu halten. Die größten Unterschiede finden sich, wie zu erwarten, bei 2XI3, deren Simulation in einer geöffneten Konformation resultiert. Die Durchschnittsstruktur, welche 2XI3 am ähnlichsten ist, ist 3FQK. Die Werte der RMSD-Matrix sind vergleichbar mit den zuvor beschriebenen Ergebnissen der RNA-Tunnel-Durchmesser- und Interdomänen-Winkel-Berechnungen, bei denen 3FQK die gleiche

Tendenz wie 2XI3 aufweist und bei denen 3U4O und 3U4R ähnlich geringe Werte wie 1NB4 und 1NB7 besitzen.

Tabelle 3: RMSD-Matrix der Durchschnittsstrukturen der Moleküldynamik-Simulationen. Die Angaben sind in Ångström (Å). Die farbliche Codierung erfolgt mit steigendem RMSD-Wert von grün (4-5 Å) über hellgrün (5-6 Å), gelb (6-7 Å) und orange (7-8 Å) nach rot (8-9 Å).

	1NB4 (Apo)	1NB7 (ssRNA)	2XI2 (Apo)	2XI3 (GTP)	3FQK (HCV-796)	3UDL (KLI)	3CJ2 (SX3)	3U4O (O8E)	3U4R (3U4R)	3VQS (JT1)
1NB4 (Apo)	Apo	5,13	5,01	8,21	5,3	5,72	4,38	4,09	4,21	5,48
1NB7 (ssRNA)	5,13	ssRNA	6,27	8,07	5,24	6,4	4,92	5,17	5,12	5,56
2XI2 (Apo)	5,01	6,27	Apo	8,27	5,58	5,59	5,11	4,83	4,89	6,37
2XI3 (GTP)	8,21	8,07	8,27	GTP	6,39	7,53	7,71	7,5	8,08	8,89
3FQK (HCV-796)	5,3	5,24	5,58	6,39	NNI-4	6,28	4,58	4,9	5,37	6,24
3UDL (KLI)	5,72	6,4	5,59	7,53	6,28	NNI-2	5,76	5,87	5,87	6,91
3CJ2 (SX3)	4,38	4,92	5,11	7,71	4,58	5,76	NNI-2	4,74	4,22	5,35
3U4O (O8E)	4,09	5,17	4,83	7,5	4,9	5,87	4,74	NNI-3	4,23	5,87
3U4R (O8F)	4,21	5,12	4,89	8,08	5,37	5,87	4,22	4,23	NNI-3	4,93
3VQS (JT1)	5,48	5,56	6,37	8,89	6,24	6,91	5,35	5,87	4,93	NNI-5

4.4.2. Hauptkomponentenanalyse

Eine weitere Möglichkeit, die Moleküldynamik-Simulationen zu vergleichen, besteht in der Untersuchung der Hauptbewegungen der Proteine. Dies erfolgte mit Hilfe der Hauptkomponentenanalyse. Dabei wird davon ausgegangen, dass die größte kollektive Bewegung, die durch die Analyse gefiltert wird, dem Konformationsraum des Proteins entspricht.

Die Eigenwerte der Hauptkomponentenanalyse sind in Abbildung 55 graphisch dargestellt und in Tabelle 12 im Anhang hinterlegt. Diese wurden für die ersten 100 ns (bzw. 50 ns für 3CJ2) in Abbildung 55A und für die letzten 100 ns der Simulation in Abbildung 55B

dargestellt. Die Eigenwerte der MD-Simulation von 2XI3 besitzen in den ersten 100 ns den höchsten Wert von 2285 \AA^2 und zeigen damit eine große Veränderung im Vergleich zu 3U4R, welche den geringsten Eigenwert von 288 \AA^2 aufweist. Innerhalb der letzten 100 ns zeigen die Simulationen von 2XI3 und 3UDL die größten Eigenwerte (1050 \AA^2 und 744 \AA^2) und 3U4O und 3U4R die geringsten (149 \AA^2 und 197 \AA^2).

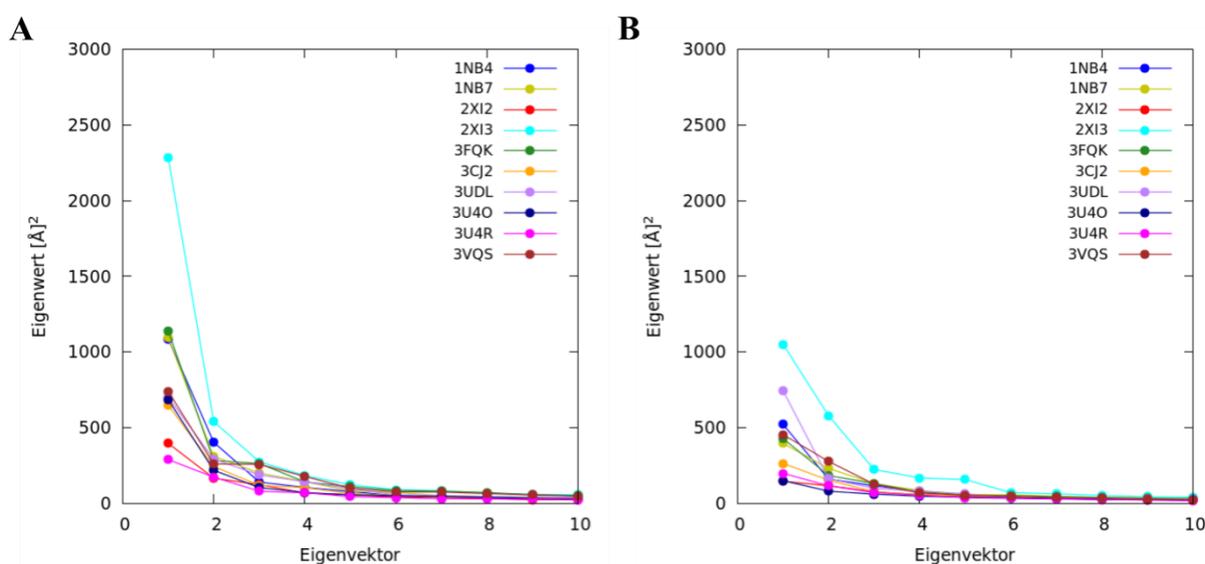


Abbildung 55: Auftragung der Eigenwerte der 10 ersten Eigenvektoren (*principle components*) der Hauptkomponentenanalyse.

A) Darstellung der Eigenwerte für die ersten 100 ns (50 ns für 3CJ2) der jeweiligen MD-Simulation. **B)** Darstellung der Eigenwerte berechnet aus den letzten 100 ns der MD-Simulation.

In Abbildung 56 sind die Eigenwerte der ersten beiden Hauptkomponenten (PC1 und PC2) für die jeweilige MD-Simulation gegeneinander aufgetragen. Diese Werte wurden für die letzten 100 ns der jeweiligen Simulation berechnet. Die in Abbildung 56 gezeigten Auftragungen der Hauptkomponenten veranschaulichen den Bewegungsraum des jeweiligen Proteins während der MD-Simulation. Je weiter der Eigenwert vom Koordinatenursprung abweicht, desto größer ist die gerichtete Bewegung im Protein. Gut zu erkennen ist, dass 2XI3 die größte gerichtete Bewegung aufweist, was vergleichbar mit den vorherigen Untersuchungen ist, in denen 2XI3 die größte Konformationsänderung aufweist. Die Simulationen von 3U4O und 3U4R nehmen nur einen kleinen Konformationsraum ein. 1NB4 und 1NB7 zeigen zudem einen vergleichbaren Bewegungsraum auf, wohingegen sich die Auftragungen von 3CJ2 und 3UDL am meisten unterscheiden. Diese Ergebnisse sind vergleichbar mit den zuvor beschriebenen Untersuchungen bezüglich der Berechnungen des RMSD-Werts, des RNA-Tunnel-Durchmessers und des Interdomänen-Winkels.

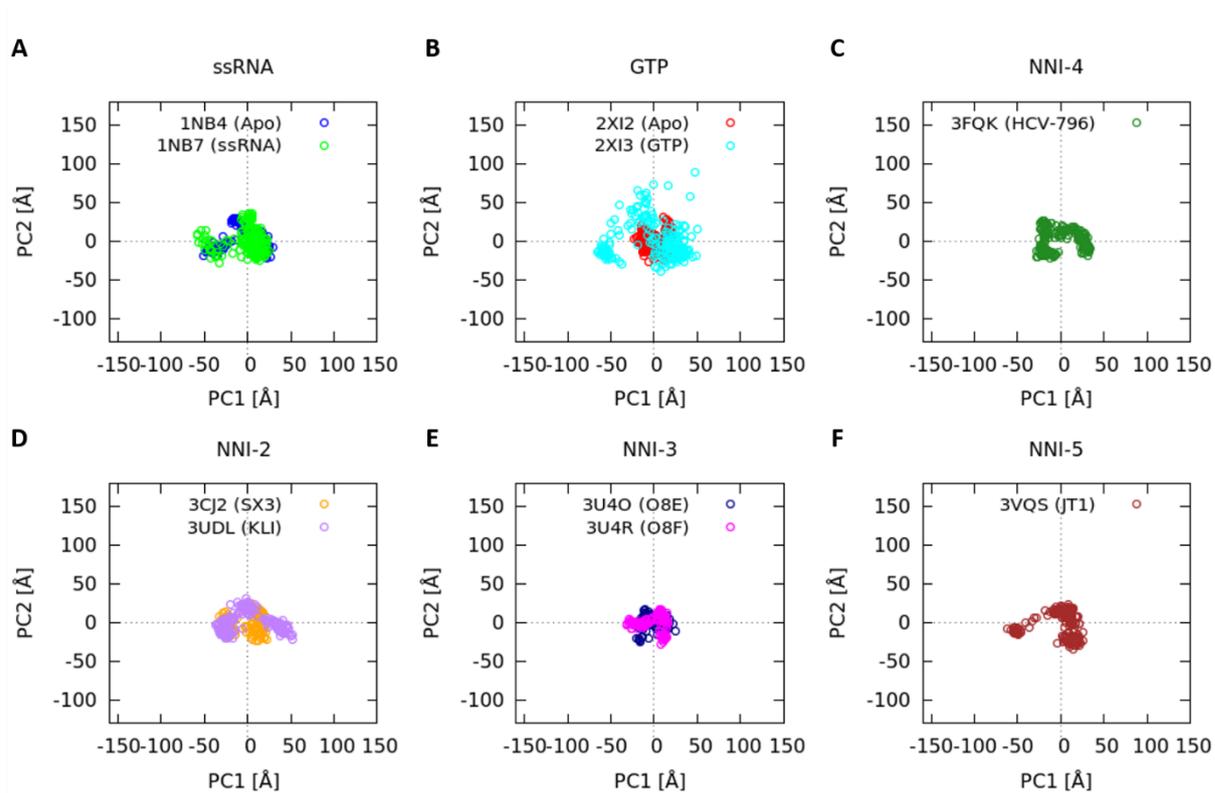


Abbildung 56: Darstellung der ersten zwei Hauptkomponenten (PC1 und PC2) für die Moleküldynamik-Simulationen von NS5B.

Die Auftragung von PC1 gegen PC2 erfolgte für **A)** 1NB4 in blau und 1NB7 in grün, **B)** 2XI2 in rot und 2XI3 in cyan, **C)** 3FQK in dunkelgrün, **D)** 3CJ2 in orange und 3UDL in violett, **E)** 3U4O in dunkelblau und 3U4R in magenta und **F)** 3VQS in braun.

4.5. Modellierung des Membranankers

Die Kristallstruktur-Untersuchungen zeigten bereits bei der Hauptkomponenten- und der RMSD-Analyse eine Veränderung der Proteinstruktur, welche auf die Länge des Linkers zurückgeführt werden konnte. Die Moleküldynamik-Simulationen zeigten ebenfalls eine große Beweglichkeit des Linkers auf, welche in der Simulation von NS5B mit GTP dazu führte, dass sich der Linker aus dem aktiven Zentrum bewegte und sich an der Daumen-Domäne neu anlagerte. Diese Untersuchungen und Beobachtungen lassen vermuten, dass die Position des Linkers von NS5B *in vivo* mit dem assoziierten Membrananker eine andere Konformation aufweisen kann, als bisher aus den Kristallstrukturen ersichtlich ist. Die vollständige Modellierung von NS5B mit dem Membrananker scheint daher essentiell zu sein, um den nativen Zustand der Polymerase analysieren zu können.

Die Modellierung des gesamten Proteins einschließlich Membrananker erwies sich als schwierig. Die hydrophobe Umgebung des Membranankers kann bei der Modellierung mit dem cytosolischen Teil von NS5B nicht berücksichtigt werden und kann aufgrund dessen zu unzuverlässigen Ergebnissen führen. Daher wurde zunächst die Sequenz des Membranankers untersucht und modelliert, um später beide Proteinteile zusammenzufügen.

4.5.1. Sequenzalignment des C-Terminus

Die Kristallstruktur von NS5B mit dem längsten C-Terminus reicht bis Aminosäure 573 (PDB-Index: 3Q0Z). Die vollständige Polymerase besitzt allerdings 591 Aminosäuren. Das Sequenzalignment der C-terminalen Aminosäuren 500 bis 591 der 8 Subtypen, die bisher kristallisiert wurden, ist in Abbildung 57 dargestellt. Das obere Alignment ist nach der BLOSUM62-Matrix, das untere Alignment nach der Hydrophobizität eingefärbt, wobei hydrophobe Aminosäuren rot gekennzeichnet sind und hydrophile blau.

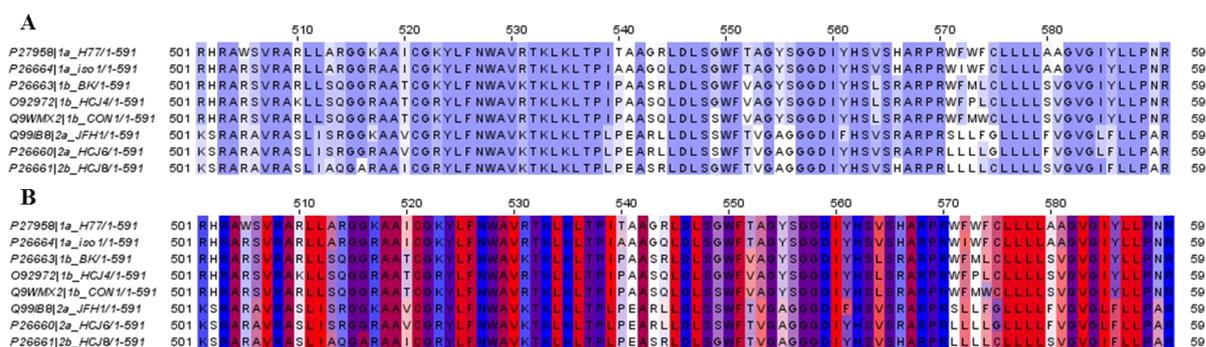


Abbildung 57: Sequenzalignment der bisher kristallisierten NS5B-Genotypen.

A) Das obere Alignment ist anhand der BLOSUM62-Matrix eingefärbt. **B)** Das untere Alignment ist zusätzlich nach der Hydrophobizität der Aminosäurereste codiert. Rot kennzeichnet eine hydrophobe Aminosäure und blau eine hydrophile. Übergänge zwischen diesen Eigenschaften sind violett gekennzeichnet.¹⁵⁵

Die letzten 20 Aminosäuren (571-591) von NS5B gehören zum Membrananker, sind sehr konserviert und sehr hydrophob (siehe Abbildung 57B) und wurden von Robetta¹⁶⁵, HMMTOP²⁰⁵, Porter¹⁶⁴, PSI-Pred und TMPred als Transmembranhelix vorhergesagt. Der Anfang des Linkers (Aminosäuren 530-539) ist ebenfalls stark konserviert und enthält ein auffällig amphiphiles Muster mit der Sequenz: 532-TKCLKLTP-538. Zudem befinden sich an den Positionen 538 und 569 zwei konservierte Prolinreste, welche als Helixbrecher fungieren können. Sie könnten beispielsweise durch eine cis/trans-Isomerisierung auch zur Aktivierung der Polymerase dienen.

4.5.2. Modellierung des Membranankers

In diesem Abschnitt wird die Homologie-Modellierung der Aminosäuren 562 bis 591 von HCV NS5B mit dem Server MUSTER¹⁷² beschrieben. Für die Modellierung wurden die Genotyp-Varianten 1a, H77, 1b, BK und 2a, JFH1 mit den entsprechenden Templates 2K9Y, 1L7V und 2KAI verwendet, welche vom Server ausgesucht wurden.

In Abbildung 58 ist exemplarisch der vom Server modellierte Membrananker von NS5B mit dem Subtyp 1b, BK in verschiedenen Darstellungen gezeigt. Die erste Darstellung rechts im Bild zeigt das Modell in der *Cartoon*-Darstellung mit einer Einfärbung vom N- zum C-Terminus von blau nach rot. Dabei beschreibt die Hauptform des Modells eine α -Helix mit einem leichten Bogen. Die letzten 3 Aminosäuren des C-Terminus sind unstrukturiert. Das mittlere und das linke Bild von Abbildung 58 zeigen den Membrananker mit einer *Connolly*-Oberfläche, die nach der Hydrophobizität der Aminosäurereste eingefärbt ist. Die orange gefärbten Bereiche sind hydrophob und befinden sich im Kernbereich des α -helikalen Membranankers. Der N-Terminus des Membranankers und die letzten 3 Aminosäuren sind dagegen hydrophiler und reichen vermutlich nur peripher in die Membran. Das MUSTER-Modell gibt einen guten ersten Einblick in die Konformation des Membranankers und seine Lage in der Membran.

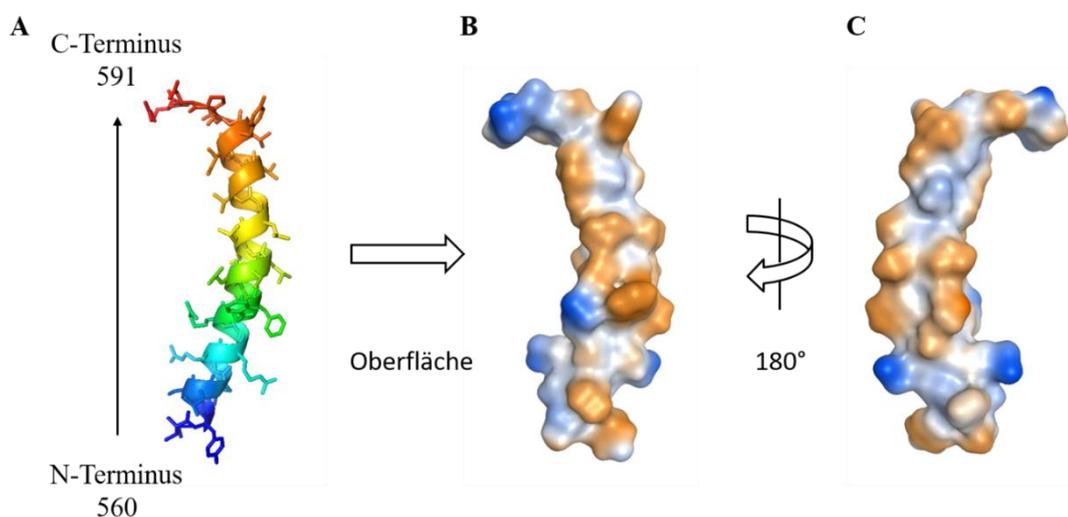


Abbildung 58: Verschiedene Darstellungen des modellierten Membranankers.

A) Das Peptid ist in der *Cartoon*-Darstellung abgebildet und vom N- zum C-Terminus von blau nach rot gefärbt. Aus Gründen der Übersichtlichkeit sind die Wasserstoffatome nicht gezeigt. **B)** Das mittlere Bild zeigt die *Connolly*-Oberflächendarstellung des Membranankers. Die hydrophoben Bereiche sind orange eingefärbt, die hydrophilen blau. **C)** Gleiche Darstellung des Membranankers wie in der Mitte nur um 180° gedreht. In allen Abbildungen zeigt der C-Terminus nach oben und der N-Terminus nach unten.

4.5.3. Vergleich mit der NMR-Struktur des Membranankers von NS5B

Seit April 2016 ist die NMR-Struktur des Membranankers von NS5B des Genotyps 1a, H77 in der PDB unter dem Eintrag 2N1P zugänglich. Sie enthält 41 berechnete Konformere des Membranankers. Dieser besteht im Bereich des C-Terminus bis zur Hälfte aus einer Helix mit einer durchschnittlichen Länge von 23,1 Å. Der Rest besitzt eine *coil*-Struktur, welche stark innerhalb der verschiedenen Modelle der NMR-Struktur variiert. Da eine dazugehörige Veröffentlichung zurzeit noch nicht publiziert ist, konnten keine genaueren Informationen zu dieser Struktur bezogen werden (Stand Februar 2017).

Zum Vergleich wurden die mit MUSTER erstellten Modelle mit der NMR-Struktur in Abbildung 59 übereinander gelagert dargestellt. Die C-terminalen Bereiche der helikalen Strukturen liegen gut übereinander. So besitzt beispielsweise das MUSTER-Modell der Variante 1b, BK einen durchschnittlichen RMSD-Wert der C_{α} -Atome von 3,5 Å im Vergleich zu den NMR-Strukturen. Die vom MUSTER-Server vorhergesagten Helices sind jedoch im Vergleich zu den NMR-Modellen länger (37-40 Å versus 23 Å). Diese Länge würde besser in eine ER-Membran passen, welche z.B. bei der Mäuseleber eine Dicke von durchschnittlich 37,5 Å besitzt.²⁰⁶ Für die vollständige Modellierung der Polymerase mit dem Membrananker wurden daher die MUSTER-Modelle und nicht die NMR-Strukturen verwendet.

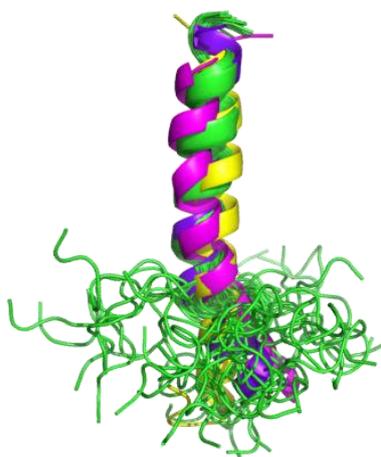


Abbildung 59: Vergleichende Darstellung der Membrananker-Modelle und der NMR-Struktur. Die Zustände der NMR-Struktur (2N1P) sind in grün und die modellierten Membrananker von GT-1a-H in gelb, GT-1b, BK in pink und GT-2a, JFH1 in violett zu sehen.

4.5.4. *Highly Mobile Membrane-Mimetic* (HMMM) -Builder

Der HMMM-Builder ist ein Webinterface von CHARMM-GUI zur Erstellung von Membran-Protein-Modellen.¹⁷³ Für die Generierung eines solchen Modells wurde das MUSTER-Modell für 1b, BK verwendet. Das dabei entstandene Membrananker-Membran-Modell hat einen

Durchmesser von ca. 40 Å. Allerdings reicht der hydrophile N-Terminus des Membranankers etwas weit in den hydrophoben Bereich der Membran hinein. Des Weiteren kann beobachtet werden, dass die letzten 3 hydrophilen Aminosäurereste des C-Terminus auf den polaren Kopfgruppen der Lipide aufliegen.

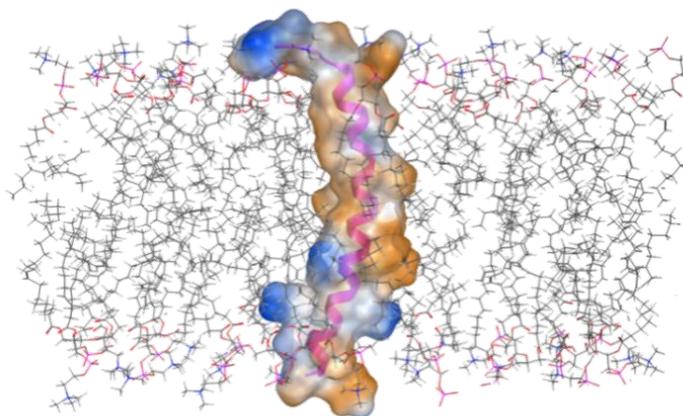


Abbildung 60: Das Ergebnis des HMMM-*Builders* ist die Modellierung des Membranankers von NS5B in einer Lipiddoppelschicht.

Die Lipide sind als graue Linien dargestellt, der Membrananker ist in der *Cartoon*-Darstellung in pink abgebildet, sowie dessen Oberfläche mit hydrophoben und hydrophilen Bereichen in orange und blau.

4.5.5. Die Position des Linkers in Anwesenheit des Membranankers

In den Kristallstrukturen reicht der Linker in das aktive Zentrum der Polymerase und verschließt so den dsRNA-Ausgang (vgl. Abbildung 6, Kapitel 1.6). Während der Moleküldynamik-Simulation des Proteins mit GTP wurde eine Positionsänderung des Linkers beobachtet, welcher sich anschließend an die Daumen-Domäne assoziierte (vgl. Kapitel 4.2.1). In diesem Kapitel soll nun geklärt werden, welche Auswirkung die Position des Linkers für die Modellierung des gesamten Proteins hat.

Würde auf den Linker gleich der Membrananker folgen, so läge das Protein entweder mit seinem RNA-Eingang oder mit seinem dsRNA-Ausgang direkt an der Membran, wie es in Abbildung 61A und B beispielhaft dargestellt ist. Im Gegensatz dazu würde der Linker einen sehr viel längeren Abstand zur Membran realisieren, wenn er sich nicht erst in das Proteinzentrum legen, sondern nur an der Daumen-Domäne assoziiert wäre. Ohne die Rotation der Daumen-Domäne kann das Protein aber weiterhin als geschlossen betrachtet werden, da der β -Loop noch immer in das katalytisch aktive Zentrum des Proteins hineinreicht und dieses verschließt. Um eine Überschneidung von Polymerase und Membran im Vollängenmodell zu vermeiden, wurde die Modellierung mit einer Kristallstruktur durchgeführt, welche einen verkürzten Linker besitzt. Hierfür wurde der PDB-Eintrag 4OBC

ausgewählt, in dem 553 Aminosäuren aufgelöst werden konnten, und eine MD-Simulation durchgeführt. Die Durchschnittsstruktur der MD-Simulation wurde anschließend für die Modellierung verwendet, wobei der Linker geradlinig verlängert und mit dem Membranankermodell verbunden wurde. Das Ergebnis ist in Abbildung 61C und D gezeigt. In diesem Aufbau wäre zwischen der Membran und dem Protein noch ausreichend Platz um eine Einzelstrang-RNA zu binden, da der Abstand zwischen Polymerase und Membran im Modell zwischen 26 Å und 37 Å beträgt. In diesem zweiten Modellvorschlag entspricht der Linker größtenteils einer gestreckten Konformation. Grundsätzlich sollte nicht ausgeschlossen werden, dass dieser Bereich auch andere Sekundärstruktur-Elemente annehmen kann, was den Abstand zwischen Protein und Membran wieder verkürzen würde.

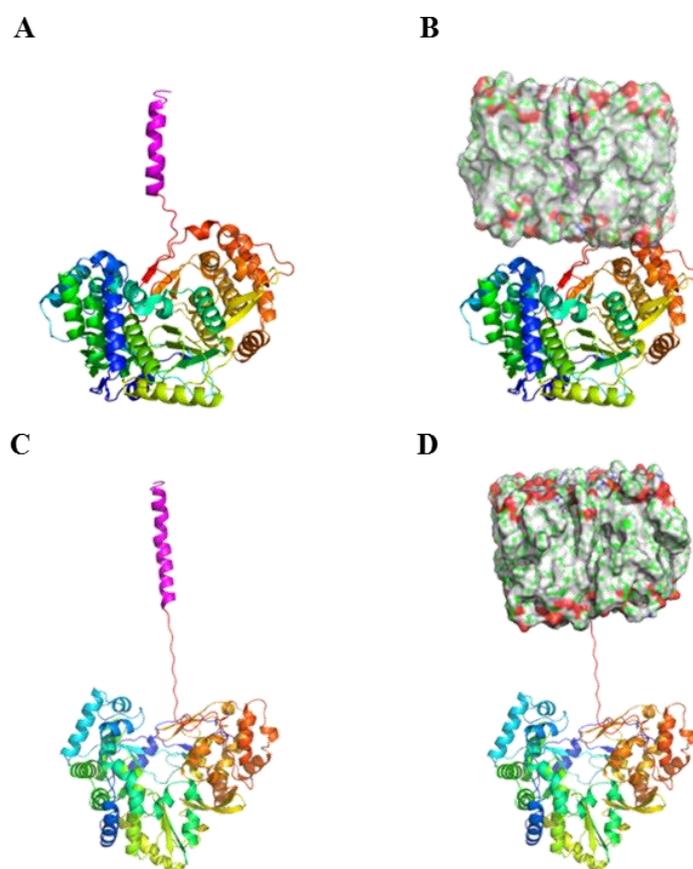


Abbildung 61: Modellierung des gesamten NS5B-Proteins.

Das Protein ist in der *Cartoon*-Darstellung abgebildet und vom N- zum C-Terminus von blau nach rot gefärbt. Der Membrananker ist magenta eingefärbt. **A)** Das dargestellte Modell entspricht einer Kombination aus der Kristallstruktur 3Q0Z und dem Muster-Modell für 1b, BK. **B)** Zusätzlich ist die Membran abgebildet, in der sich der Membrananker befindet. **C)** Die Modellierung des NS5B-Proteins erfolgte anhand der minimierten MD-Durchschnittsstruktur von 4OBC und dem zuvor erstellten Membrananker für 2a, JFH1. **D)** Zusätzlich ist die Membran abgebildet, in der sich der Membrananker befindet.

Das oben abgebildete Protein stellt dabei nur ein Beispiel für die gesamte Polymerase und die Anordnung des Linker-Bereichs dar. Die genaue Position des Linkers kann nur in einer Kristallisation der gesamten Polymerase mit Membrananker geklärt werden. Zudem spielen wahrscheinlich die Interaktionen anderer HCV Nicht-Struktur-Proteine eine entscheidende Rolle für die replikationskompetente Form von NS5B. Aus Publikationen sind die Wechselwirkungen mit NS4A und NS3 bekannt²⁰⁷⁻²⁰⁹. Diese könnten die Konformation der Polymerase in einer Weise verändern, die aus den Kristallstrukturen von NS5B nicht ersichtlich ist. Eine Möglichkeit, die Wechselwirkungen zwischen NS5B, NS3 und NS4A zu untersuchen, wären *Crosslink*-Experimente, die zumindest die Interaktionsbereiche zwischen diesen Proteinen klären könnten. Auf Grundlage dessen könnten wiederum neue Konformations-Modelle von NS5B erstellt werden.

Die *Crosslink*-Experimente im folgenden Abschnitt dienen zwar nicht zur Untersuchung der Interaktion von NS5B mit anderen HCV Proteinen, sondern der Validierung der bisher analysierten MD-Simulationen und der konformationellen Änderung von NS5B in Anwesenheit verschiedener Substrate bzw. des Inhibitors HCV-796.

4.6. *Crosslink*-Experimente

Um die konformationellen Änderungen von NS5B in Anwesenheit von NTPs, ssRNA und des Inhibitors HCV-796 experimentell nachzuweisen, eignet sich die Methode des *Crosslinkings*. Hierbei werden spezifische Aminosäuren mit funktionellen Seitenketten im Protein miteinander kovalent verknüpft, das Protein proteolytisch verdaut und die über den *Crosslinker* verknüpften Fragmente massenspektrometrisch identifiziert. Um eine Quervernetzung eindeutig nachzuweisen, ist es hilfreich den *Crosslinker* zu markieren, zum Beispiel durch eine Deuterierung, oder einen durch CID zerfallenden *Crosslinker* zu verwenden.²¹⁰

Die *Crosslink*-Experimente wurden mit der aminreaktiven *Crosslinker*-Reagenz BS²G (Bis(sulfosuccinimidyl)glutarat) durchgeführt. Dieser *Crosslinker* reagiert mit den funktionellen Gruppen von Lys (sowie Ser, Thr und Tyr) und hat eine Überbrückungsdistanz von 7,7 Å (vgl. Kapitel 2.7.2). Je nach Ausrichtung der Aminosäureseitenketten und der Flexibilität des Proteins, können die entsprechenden C_α-Atome im Protein einen Abstand von bis zu 26 Å Entfernung aufweisen.²¹¹ Die *Crosslink*-Experimente erfolgten mit HCV NS5B der Variante 1b, BK (Uniprot-ID P26663) und 2a, JFH1 (Uniprot-ID Q99IB8). Das Protein wurde ohne und mit Zugabe von NTP, ssRNA und dem Inhibitor HCV-796 quervernetzt.

Dabei wurde BS²G im nicht-deuterierten (d₀) und deuterierten Zustand (d₄) als *Crosslink*-Reagenz verwendet, wie im Kapitel 3.7.5 beschrieben. Durch die Massenunterschiede der Deuterierung können die *Crosslink*-Produkte nicht nur qualitativ, sondern auch quantitativ untersucht werden, indem die Peakflächen der entsprechenden Massenspektren verglichen werden.

Die Verteilung der Aminosäure Lysin fällt in den einzelnen Domänen von NS5B sehr unterschiedlich aus (vgl. Abbildung 62). Viele Lysinseitenketten befinden sich in der Finger-Domäne (39), dagegen nur wenige in der Hand-Domäne (10) und der Daumen-Domäne (7). Betrachtet man zusätzlich die Verteilung der Aminosäuren Serin, Threonin und Tyrosin, die ebenfalls mit BS²G reagieren können, sieht die Verteilung von potentiellen *Crosslink*-Partnern vorteilhafter aus. Ausschlaggebend für den Erfolg des *Crosslink*-Versuchs ist aber die Zugänglichkeit dieser Seitenketten, damit die Verknüpfung mit BS²G erfolgen kann.

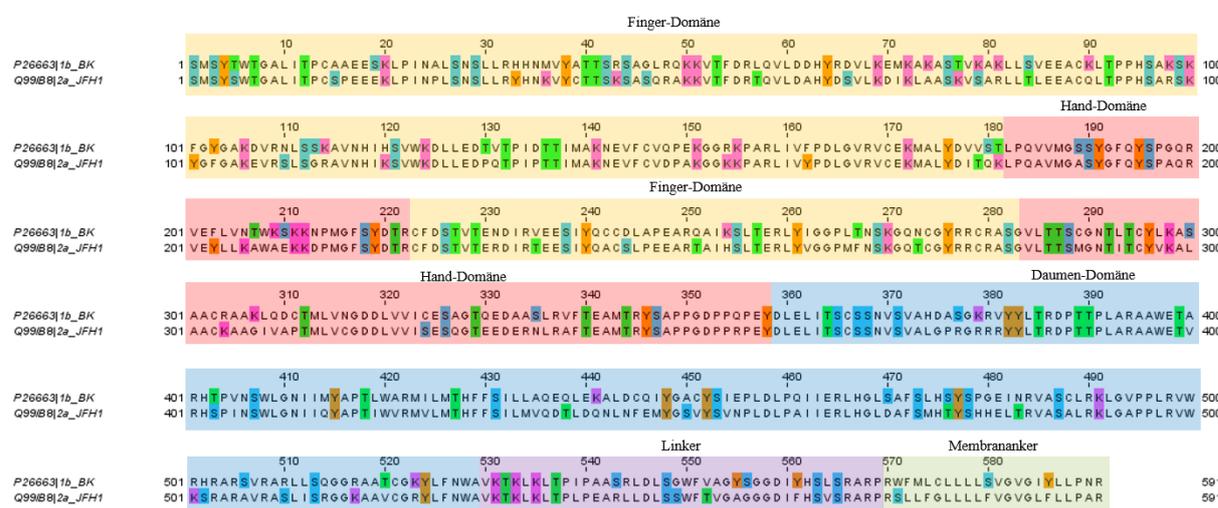


Abbildung 62: Aminosäuresequenzen der NS5B-Varianten 1b, BK und 2a, JFH1.

Die Proteinabschnitte sind farblich gekennzeichnet, gelb für die Finger-Domäne, rot für die Hand-Domäne, blau für die Daumen-Domäne, violett für den Linker und grün für den Membrananker. Zusätzlich sind die Lysine violett, Serine blau, Threonine grün und Tyrosine orange markiert.

Die *Crosslink*-Ansätze wurden mittels einer SDS-PAGE aufgetrennt (siehe Abbildung 63). Die Referenz-Bande konnte auf Höhe von 65 kDa detektiert werden und entspricht einer Monomerbande. Bei der Zugabe des *Crosslinkers* zeigte sich ein Schmier oberhalb des Molekulargewichts der Polymerase, welcher auf die Quervernetzung des Proteins zurückzuführen ist. Ebenfalls konnte bei Anwesenheit von BS²G eine Dimerbande ab einem Molekulargewicht von 130 kDa und darüber detektiert werden. Für die Untersuchung der konformationellen Änderung von NS5B sind jedoch ausschließlich die intramolekularen *Crosslinks* und damit nur die Monomerbande von Interesse, welche für die anschließenden

Tandem-MS-Untersuchungen verwendet wurde. Die Banden der SDS-Page wurden entsprechend der schwarzen Markierungen in Abbildung 63 auf Höhe der Referenz und darüber im Bereich des Schmiers ausgeschnitten, welche im weiteren Verlauf der Arbeit als obere und untere Bande bezeichnet werden. Einen Einfluss der Liganden auf die Vernetzung der Polymerase ist aus der SDS-PAGE nicht ersichtlich.

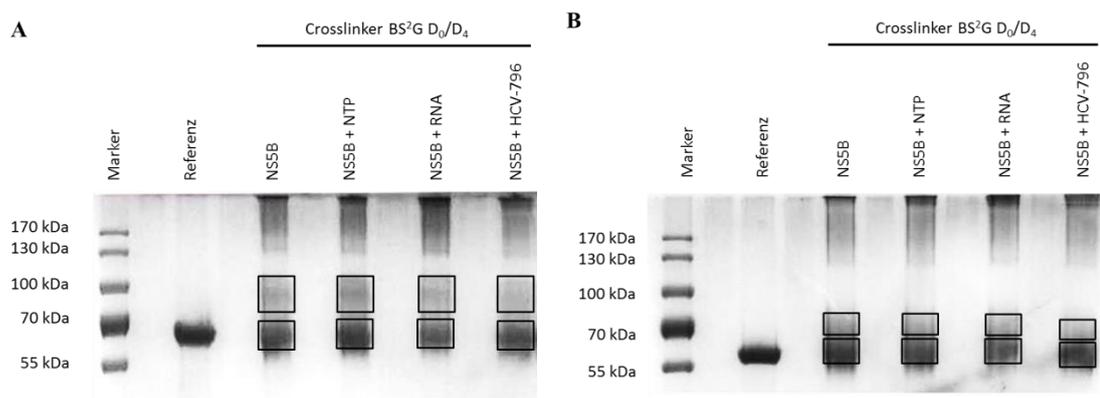
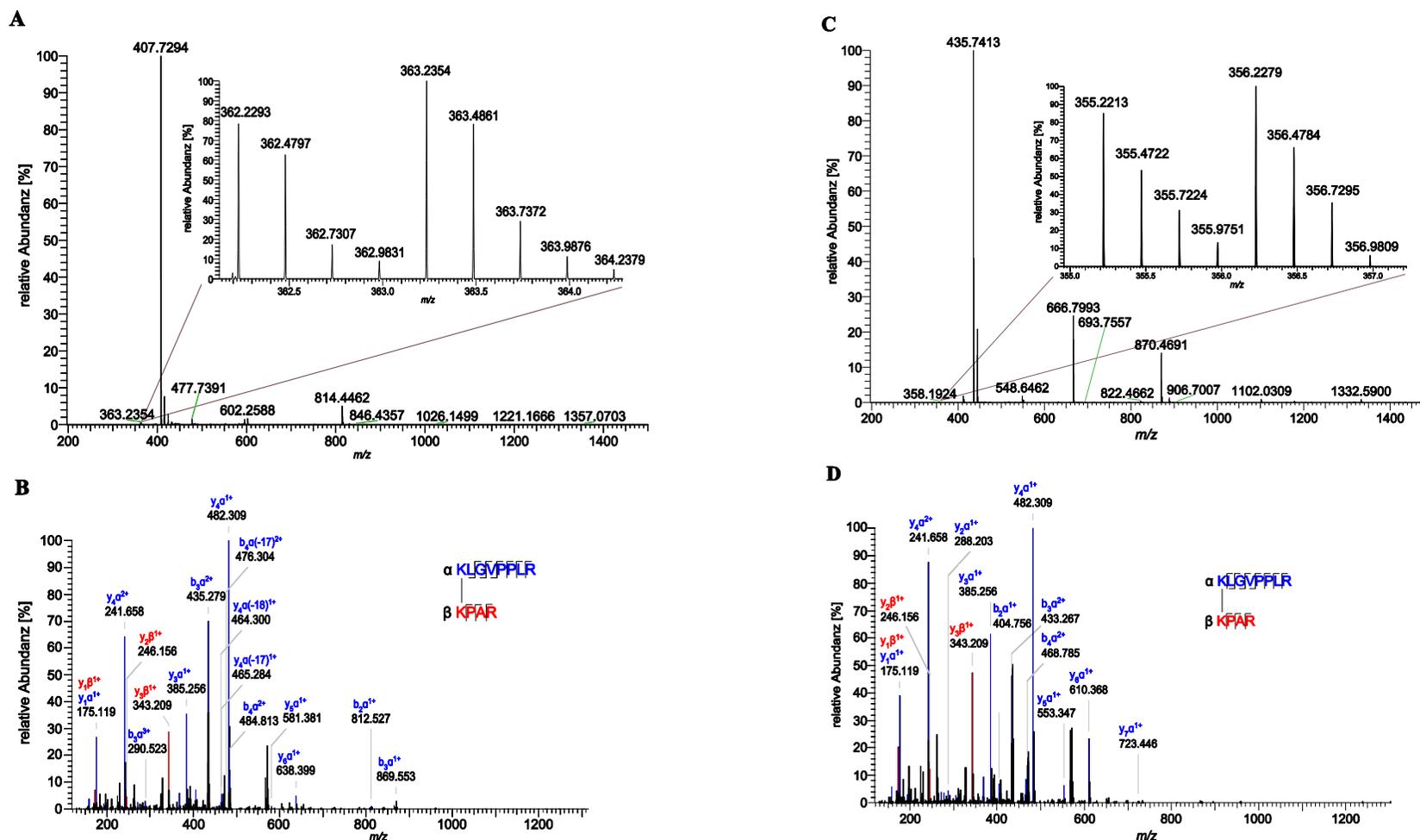


Abbildung 63: 10 %-iges SDS-Gel der quantitativen *Crosslink*-Versuche mit NS5B und BS²G d₀/d₄. Die Polymerase wurde im Apo-Zustand und unter Zugabe von NTP, RNA oder HCV-796 vernetzt. Als Referenz diente NS5B ohne die Zugabe von BS²G. **A**) In der linken Darstellung ist der Versuch mit der Variante 1b, BK gezeigt, **B**) die rechte Abbildung zeigt das Experiment mit der Variante 2a, JFH1.

Als Beispiel eines interpeptidalen *Crosslink*-Produkts sind die Massenspektren und die Fragmentionen-Massenspektren zwischen Lys155 und Lys491 der Genotypen 1b, BK und 2a, JFH1 in Abbildung 64A-D zu sehen. Auch wenn das Signal des *Crosslinks* im Massenspektrum nur eine geringe Intensität aufweist, kann es doch anhand des Deuterierungsmusters und des Fragmentionensmusters identifiziert werden.

Abbildung 64: MS- und MS²-Spektren der *Crosslink*-Versuche mit BS²G und NS5B.

A) Massenspektrum von GT-1b, BK bei 43,68 min. Das d₀/d₄-Signal des *Crosslink*-Produkts bei *m/z* 362,2293 und einer Ladung von 4⁺ ist vergrößert dargestellt. **B**) Fragmentionen-Massenspektrum des *Crosslink*-Produkts zwischen Lys155 (α -Sequenz, 155-162, blau) und Lys491 (β -Sequenz, 491-494, rot). **C**) Massenspektrum von GT-2a, JFH1 bei 38,88 min. Das d₀/d₄-Signal des *Crosslink*-Produkts bei *m/z* 355,2213 und einer Ladung von 4⁺ ist vergrößert dargestellt. **D**) Fragmentionen-Massenspektrum des *Crosslink*-Produkts zwischen Lys155 (α -Sequenz, 155-162, blau) und Lys491 (β -Sequenz, 491-494, rot).

Mit Hilfe der MS- und MS²-Spektren konnten die *Crosslinks* von NS5B identifiziert und analysiert werden. Diese befinden sich alle in den Abschnitten der Finger- und der Daumen-Domäne. Dabei zeigten sich Unterschiede zwischen den Genotypen 1b, BK und 2a, JFH1. Diese sind in Abbildung 65 für die Variante 1b, BK mit gelb eingefärbten Seitenketten, für die Variante 2a, JFH1 mit blauen Seitenketten und für gemeinsame *Crosslinks* mit grün eingefärbten Seitenketten dargestellt. Die übereinstimmenden *Crosslinks* befinden sich zwischen den Seitenketten Lys51 und Lys155, sowie zwischen Lys155 und Lys491 (grün eingefärbte Seitenketten in Abbildung 65). Sequenzspezifische *Crosslinker* wurden für die Variante 1b, BK innerhalb der Finger-Domäne zwischen den Seitenketten von Lys98 und Lys114, sowie zwischen Lys106 und Lys114 gefunden. Für die Variante 2a, JFH1 existieren spezifische *Crosslinks* in der Daumen-Domäne zwischen den Seitenketten von Lys491 und Lys501, Lys501 und Lys531, sowie zwischen Lys501 und Lys533. Diese Genotyp-abhängigen *Crosslinks* wurden in Abbildung 65 schwarz eingekreist.

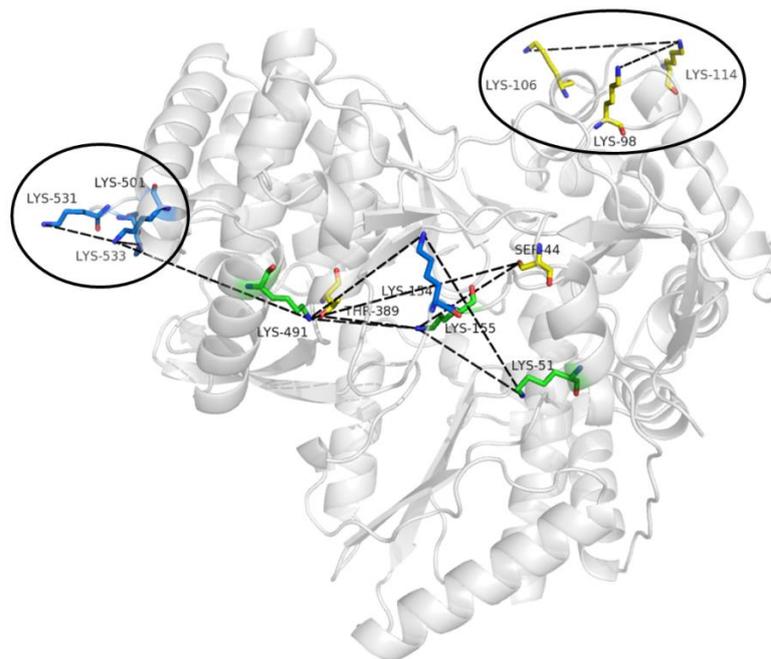


Abbildung 65: Darstellung der *Crosslinks* aus den Versuchen mit den NS5B-Varianten 1b, BK und 2a, JFH1.

NS5B ist grau mit Blick auf den NTP-Eingang abgebildet. Die an den *Crosslinks* beteiligten Aminosäureseitenketten von 1b, BK sind gelb eingefärbt, die aus 2a, JFH1 sind blau und übereinstimmende Vernetzungen sind grün. Die jeweiligen *Crosslinks* sind durch schwarz gestrichelte Linien gekennzeichnet.

Die *Crosslinks* wurden im Massenspektrum entsprechend ihrer Retentionszeit in d₀- und d₄-Fragmente eingeteilt, deren Peakflächen bestimmt und daraus das d₀/d₄-Verhältnis errechnet.

Diese d_0/d_4 -Verhältnisse sind in Tabelle 13 für GT-1b, BK und in Tabelle 14 für GT-2a, JFH1 im Anhang angegeben. Für eine zusammenfassende Auflistung der *Crosslinks* wurden die d_0/d_4 -Verhältnisse gleicher verlinkter Peptidabschnitte arithmetisch gemittelt und auf den entsprechenden Versuch mit NS5B ohne Substratzugabe normiert. Diese normierten Werte sind in Tabelle 4 zusammengefasst, zusammen mit den entsprechenden C_α -Abständen aus der Kristallstruktur 2XI2¹. Die gemessenen C_α -Abstände der jeweiligen *Crosslinks* befinden sich alle in einem für BS²G möglichen Überbrückungsbereich, sodass allein anhand dieser Distanzen keine qualitative Aussage zur Konformationsänderung des Proteins bei Zugabe der Liganden gemacht werden kann. Anhand des d_0/d_4 -Verhältnisses lassen sich jedoch Positionsveränderungen im Protein ableiten. Eine Abnahme des Verhältnisses von BS²G- d_0 zu BS²G- d_4 bedeutet eine Zunahme an deuterierten *Crosslinks* (d_4) bei Substratzugabe, eine Zunahme des Verhältnisses bedeutet eine Abnahme an deuterierten *Crosslinks* (d_4) bei Substratzugabe im Vergleich zum nicht-deuterierten *Crosslinker* (d_0).

Um diese Verhältnisse graphisch darzustellen, sind in Abbildung 66 die *Crosslinks* als Kreisdiagramme gezeigt. Die Quervernetzungen sind dabei entsprechend der Zu- oder Abnahme farblich gekennzeichnet. Bleibt das Verhältnis von BS²G- d_0 und BS²G- d_4 in etwa gleich, sind die *Crosslinks* als schwarze Linien dargestellt. Eine Abweichung zu Gunsten von BS²G- d_4 ist grün, eine Abweichung zu Gunsten von BS²G- d_0 ist rot gefärbt. Gut zu erkennen ist, dass bei der Zugabe von NTP zum Versuchsansatz eine Zunahme an deuterierten *Crosslinks* bei beiden Genotypen zu beobachten ist. Bei dem Versuchsansatz mit ssRNA und GT-1b, BK ist eine Tendenz zur Abnahme an *Crosslink*-Produkten zu erkennen. Das Experiment mit dem Inhibitor HCV-796 zeigt in beiden Genotypen sehr unterschiedliche Tendenzen auf.

Tabelle 4: Auflistung der identifizierten *Crosslinks* von NS5B mit BS²G und verschiedenen Liganden.

Für die intrapeptidalen *Crosslinks* sind die gefundenen Sequenzen und die dazugehörigen Aminosäuren aufgelistet. Daneben wurde vermerkt, in welchem Verhältnis die vernetzten Fragmente identifiziert wurden und welcher Distanz (C_{α} - C_{α}) diese in der Kristallstruktur 2XI2 entsprechen. Abschnitte aus der Finger-Domäne sind gelb hinterlegt, aus dem Motiv F grün und aus der Daumen-Domäne und dem Linker blau.

Aminosäure 1	Sequenz 1	Aminosäure 2	Sequenz 2	Verhältnis d_0/d_4 für 1b, BK				Verhältnis d_0/d_4 für 2a, JFH1				Abstand im Kristall [Å]
				Apo	NTP	ssRNA	HCV-796	Apo	NTP	ssRNA	HCV-796	
K98	[LTPPHSA <u>K</u> SK]	K114	[NLSS <u>K</u> AVNHIHSVWK]	1,00	0,67	0,92	0,64					8,7
K106	[FGYGA <u>K</u> DVR]	K114	[NLSS <u>K</u> AVNHIHSVWK]	1,00	0,51	0,92	0,59					11,8
K155	[<u>K</u> PAR]	S44	[<u>S</u> AGLR]	1,00	0,83	2,56	2,29					8,4
K154	[GG <u>K</u> KPAR]	K51	[<u>K</u> VTFDR]					1,00	0,61	0,54	0,86	14,2
K155	[<u>K</u> PAR]	K51	[<u>K</u> VTFDR]	1,00	0,82	0,88	0,86	1,00	0,91	0,86	0,86	11,4
K155	[<u>K</u> PAR]	K51	[QK <u>K</u> VTFDR]	1,00	0,96	1,50	1,20	1,00	0,75	0,97	0,56	14,6
K155	[<u>K</u> PAR]	T389	[DPTTPLAR]	1,00	0,73	1,57	0,46					24,5
K491	[<u>K</u> LGAPPLR]	S44	[<u>S</u> AGLR]	1,00	0,50	1,29	3,03					15,7
K491	[<u>K</u> LGAPPLR]	K154	[GG <u>K</u> KPAR]					1,00			0,33	17,9
K491	[<u>K</u> LGAPPLR]	K155	[<u>K</u> PAR]	1,00	0,83	2,28	3,43	1,00	0,66	1,33	1,05	19,8
K491	[<u>K</u> LGAPPLR]	K501	[VW <u>K</u> SR]						0,21			10,9
K531	[YLFNWAV <u>K</u> TK]	K501	[VW <u>K</u> SR]					1,00	0,70	0,38	0,77	14,0
K533	[TK <u>L</u> K]	K501	[VW <u>K</u> SR]					1,00	0,78	0,33		8,7

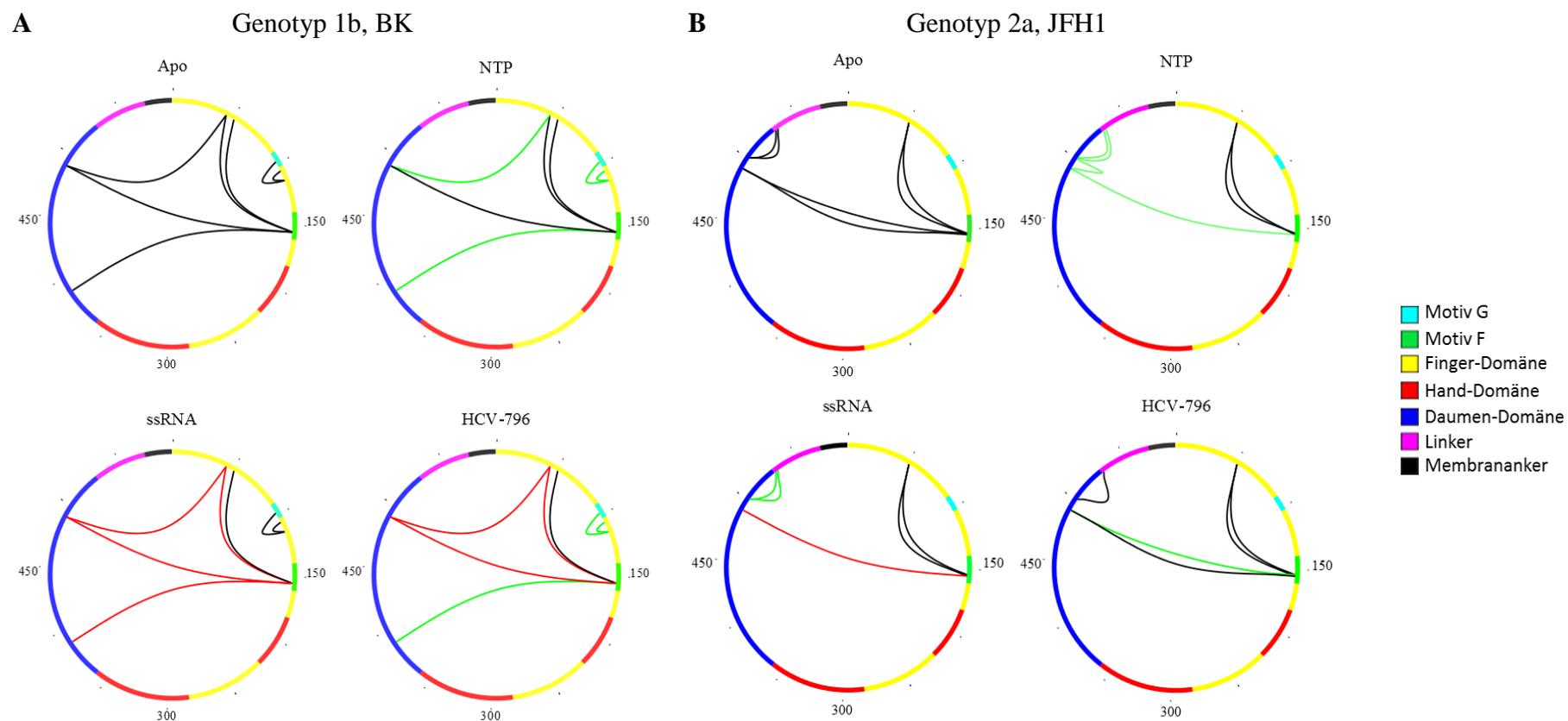


Abbildung 66: Kreisdiagramme der *Crosslink*-Versuche mit NS5B Genotyp 1b, BK und Genotyp 2a, JFH1. Die *Crosslinks* sind durch Linien im Kreisdiagramm gekennzeichnet. Große Abweichungen des d_0/d_4 -Verhältnisses zu Gunsten von d_4 sind grün und zu Gunsten von d_0 rot gefärbt. Die Farbzugeweisungen bestimmter Motive und Proteinabschnitte sind in der Legende angegeben.

In Tabelle 5 ist ein Vergleich der gemittelten C_{α} -Atom-Abstände aus den MD-Simulationen entsprechend der *Crosslinks* aufgeführt. Veränderungen in den Distanzen während der Simulationen können somit mit den unterschiedlichen *Crosslinks* verglichen werden.

Tabelle 5: Abstandsmessung der quervernetzten Aminosäuren aus den MD-Simulationen.

Berechnet wurden jeweils die C_{α} -Atom-Abstände der letzten 100 ns der Dynamiken. Aminosäuren aus der Finger-Domäne sind gelb hinterlegt, aus dem Motiv F grün und aus der Daumen-Domäne und dem Linker blau.

<i>Crosslinks</i> zwischen den Aminosäuren		Abstandsmessung der C_{α} -Atome					
		Kristall [Å]			MD-Simulation [Å]		
Aminosäure 1	Aminosäure 2	2XI2	1NB4 (apo)	2XI2 (apo)	1NB7 (ssRNA)	2XI3 (GTP)	3FQK (HCV-796)
K98	K114	8,7	12,6	15,6	12,4	11,8	11,9
K106	K114	11,8	11,3	12,9	11,3	11,7	12,1
K155	S44	8,4	7,9	10,2	7,2	7,9	8,3
K154	K51	14,2	13,5	19,8	19,4	15,6	18,5
K155	K51	11,4	11,8	16,9	17,2	12,3	17,9
K155	T389	14,6	9,9	19,7	14,2	12,9	12,5
K491	S44	24,5	25,5	32,1	22,6	18,6	23,1
K491	K154	15,7	17,9	20,9	16,4	10,3	15,3
K491	K155	17,9	19,6	22,5	18,7	11,4	15,9
K491	K501	19,8	20,2	20,4	17,4	20,7	19,1
K531	K501	10,9	14,7	13,3	19,7	15,3	22,7
K533	K501	14,0	20,4	18,9	25,5	17,6	26,2

Für den Genotyp 1b, BK führte die Zugabe von NTP zum Reaktionsansatz in allen identifizierten *Crosslinks* zu einer Abnahme des *Crosslink*-Verhältnisses zwischen BS²G-d₀ und BS²G-d₄ (siehe Tabelle 4 und Abbildung 66). Dies bedeutet eine Zunahme der Quervernetzungen im Protein bei Anwesenheit von NTP. Besonders deutlich wird dies zwischen den *Crosslinks* der Seitenketten von Lys106 und Lys114, sowie Lys98 und Lys114. Dieser Bereich ist an der ssRNA-Bindung beteiligt und sollte keinen Einfluss auf die NTP-Bindung haben. In der MD-Simulation mit GTP zeigten sich allerdings genau für diesen Bereich große Unterschiede zur Kristallstruktur (vgl. Abbildung 26). Diese verstärkte Bewegung könnte zu einer größeren Exponiertheit der Lysinseitenketten führen und somit die Zugänglichkeit für die *Crosslink*-Reagenzien erhöhen. Zwei weitere interessante *Crosslinks* befinden sich zwischen den Seitenketten von Ser44 und Lys491, sowie zwischen Lys155 und Lys491. Die vorhergesagte Rotation der Daumen-Domäne bei Anwesenheit von NTP verkürzt den C_{α} -Abstand zwischen diesen Aminosäureresten und bringt Lys491 in eine verbesserte

Crosslink-Position. In der MD-Simulation verringert sich der C_{α} -Abstand zwischen Lys155 und Lys491 von 22,5 Å in der MD mit 2XI2 zu 11,4 Å in der Simulation von 2XI3 mit GTP (siehe Tabelle 5). Der C_{α} -Abstand zwischen Ser44 und Lys491 verkürzt sich in der Simulation von 2XI2 von 32,1 Å auf 18,6 Å in der Simulation von 2XI3 mit GTP.

Die Zugabe von ssRNA zum Reaktionsansatz führt bei der Variante 1b, BK zu unterschiedlichen *Crosslink*-Verhältnissen. Im Bereich der Finger-Domäne entspricht das Verhältnis des *Crosslinks* zwischen Lys98 und Lys114, sowie zwischen Lys106 und Lys114 dem Apoprotein-Experiment. Obwohl die ssRNA an dieser Stelle an das Protein bindet, scheint sie das *Crosslink*-Verhalten nicht zu beeinflussen. In den übrigen *Crosslinks* überwiegen die d_0 -verknüpften Fragmente, was eine Abnahme der ausgebildeten *Crosslinks* im Vergleich zum Apoprotein bedeutet. Besonders die zwischen den Aminosäuren Ser44 und Lys155, Ser44 und Lys491, sowie Lys155 und Lys491 verknüpften Fragmente weisen sehr viel höhere Anteile an d_0 - als an d_4 -Fragmenten auf. Diese Aminosäuren interagieren nicht mit der ssRNA. Auch die C_{α} -Abstandsmessungen der MD-Simulationen zeigen keine auffälligen Veränderungen in diesen Bereichen auf. Allerdings ist in der MD-Simulation von 1NB7 mit ssRNA ersichtlich, dass sich die Position des Motivs F ändert und das Protein-*backbone* in Richtung des RNA-Tunnels und nicht in Richtung des NTP-Eingangs zeigt. Diese Konformationsänderung führt zu einer verschlechterten Zugänglichkeit der *Crosslink*-Reagenz für Ser44, welches in der MD-Simulation über Wasserstoffbrücken mit den Amidbindungen von Thr41 und Gly153 interagiert. Die Analyse der Lösungsmittelzugänglichkeit (*solvent-accessible-surface distance*) mit dem Server XWalk¹⁵⁴ ergab für diese *Crosslinks* (Ser44 und Lys155, sowie Ser44 und Lys491) einen Abstand größer als 34 Å, und damit eine Reichweite außerhalb der *Crosslinker*-Distanz. Aus der Dynamik von 1NB7 mit ssRNA lässt sich kein ersichtlicher Grund für die Abnahme der *Crosslinks* mit Lys155 ableiten. Das Motiv F und die benachbarten Aminosäuren Arg158 und Gly153 gehen stabile Wechselwirkungen mit der ssRNA und Ser44 ein. Daraus resultiert eine Abschirmung und Unzugänglichkeit von Lys155 für den *Crosslinker*.

Der Versuchsansatz von NS5B der Variante 1b, BK und dem Inhibitor HCV-796 führte zu sehr unterschiedlichen quantitativen *Crosslink*-Ergebnissen. Die Vernetzungen nehmen zwischen den Seitenketten von Lys106 und Lys114, Lys98 und Lys114, sowie zwischen Lys155 und Thr389 zu. Teilweise entsprechen die *Crosslink*-Anteile dem des Apoprotein-Versuchs, wie zwischen den Seitenketten von Lys155 und Lys51. Weiterhin zeigen sich deutliche Abnahmen an *Crosslink*-Produkten zwischen Ser44 und Lys155, Ser44 und Lys491 und Lys155 und Lys491, vergleichbar mit den *Crosslink*-Experimenten von NS5B mit ssRNA. Der Inhibitor HCV-796 scheint die Bindung eines NTPs zu imitieren, sodass es im

Bereich der Finger-Domäne zu einer verstärkten Quervernetzung kommt (Lys98 und Lys114, sowie Lys106 und Lys114), welche vergleichbar mit dem *Crosslink*-Versuch mit den Nukleotiden ist. In der MD-Simulation von NS5B mit HCV-796 zeigte sich eine beginnende Daumen-Rotation der Polymerase, welche jedoch durch die sterisch ungünstige Position des Inhibitors nicht ausgeführt werden konnte, sodass NS5B im geschlossenen Zustand blieb. Für die *Crosslink*-Fragmente zwischen den Seitenketten von Lys155 und Thr389 ist eine Zunahme des d_4 -Anteils erkennbar. Räumlich benachbart zu Thr389 befindet sich Arg386, ein Interaktionspartner von HCV-796. Die Wechselwirkungen zwischen Arg386 und HCV-796 könnten zu einer Immobilisierung von Thr389 führen, woraus die erhöhte *Crosslink*-Effizienz resultiert. Die Abnahme der *Crosslinks* zwischen den Seitenketten von Ser44 und Lys155, Ser44 und Lys491 und Lys155 und Lys491 lassen sich aus der MD-Simulation nicht erklären und sind vergleichbar mit den *Crosslink*-Ergebnissen bei Zugabe von ssRNA. Das verstärkte Auftreten der *Crosslinks* innerhalb der Finger-Domäne (Lys98 und Lys114, sowie Lys106 und Lys114) kann somit nicht mit einer erhöhten Flexibilität in Verbindung gebracht werden, wie bei der MD-Simulationen mit GTP. Dazu ist die Assoziation zum C-terminalen Linker in der Simulation mit HCV-796 zu stark (vgl. Kapitel 4.3.1).

Das *Crosslink*-Experiment mit der NS5B-Variante 2a, JFH1 ergab weniger *Crosslinks* als mit 1b, BK. Diese befinden sich im Bereich der NTP-Bindestelle (Lys51, Lys154 und Lys155) und in der Daumen-Domäne (Lys491, Lys501, Lys531 und Lys533).

Das d_0/d_4 -Verhältnis fällt bei Anwesenheit von NTP im Vergleich zum Apoprotein-Versuch geringer aus, wie bereits bei dem Versuch mit der Variante 1b, BK. Am stärksten wird der Anteil von BS^2G-d_4 im *Crosslink* zwischen den Seitenketten Lys491 und Lys501 erhöht (Verhältnis d_0/d_4 von 0,32). In der MD-Simulation besitzt dieser Bereich in der Daumen-Domäne eine große dynamische Variabilität, auch wenn sich die Durchschnittsabstände der C_α -Atome von Lys491 und Lys501 in der Simulation nicht stark verändern (von 20,4 Å in der MD mit 2XI2 auf 20,7 Å in der MD von 2XI3 mit GTP). Auch der Überschuss von BS^2G-d_4 in den *Crosslinks* zwischen Lys501 und Lys531, sowie Lys501 und Lys533 deuten auf eine Konformationsänderung innerhalb der Daumen-Domäne hin. Zwischen dem Motiv F und der Daumen-Domäne ist bei Zugabe von Nukleotiden ebenfalls ein Anstieg von BS^2G-d_4 detektiert worden (*Crosslink* zwischen Lys155 und Lys491). In der MD-Simulation gelangt die Daumen-Domäne durch ihre Rotation in Richtung des NTP-Eingangs näher an das Motiv F, sodass sich der Abstand für die C_α -Atome von Lys155 und Lys491 in der MD-Simulation von durchschnittlich 22,5 Å in der MD mit 2XI2 auf 11,4 Å in der MD mit 2XI3 und GTP verkürzt, was in einer verstärkten Effizienz der Vernetzung resultieren könnte.

Der *Crosslink*-Versuch mit NS5B und ssRNA zeigte auch mit der Variante 2a, JFH1 sehr unterschiedliche d_0/d_4 -Verhältnisse. Eine deutliche Zunahme der *Crosslinks* bei Anwesenheit von ssRNA konnte innerhalb der Daumen-Domäne zwischen den Seitenketten Lys501 und Lys531, sowie zwischen Lys501 und Lys533 beobachtet werden, was für eine Konformationsänderung in diesem Bereich spricht. Aus der MD-Simulation mit ssRNA konnte für den Abschnitt zwischen den Aminosäuren 490-520 und 530-545 eine deutlich erhöhte Flexibilität des Proteinrückgrates festgestellt werden im Vergleich zum Apoprotein (vgl. Abbildung 25). Die Verteilung der d_0 - und der d_4 -Fragmente aus dem Motiv F und der Daumen-Domäne, repräsentiert durch die Quervernetzung zwischen den Seitenketten von Lys155 und Lys491, zeigt dagegen eine Abnahme an *Crosslinks* in Anwesenheit von ssRNA. Dieses Verhalten konnte bereits mit der Variante 1b, BK beobachtet werden. Zwar ist aus den MD-Simulationen keine große Abstandsveränderung zu erkennen (19,6 Å in der MD mit 1NB4 und 18,7 Å in der MD mit 1NB7 und ssRNA), so zeigte sich doch die Positionsänderung des Motivs F aus der MD-Simulation von 1NB7 mit ssRNA. Aus der Kristallstruktur ist bekannt, dass die ssRNA unter anderem mit Lys141 und Arg158 interagiert (vgl. Abbildung 27C), was eine verschlechterte Zugänglichkeit von Lys155 bezüglich der *Crosslink*-Reagenz bewirken kann.

Die Ergebnisse von NS5B der Variante 2a, JFH1 unter der Zugabe von HCV-796 zeigt ähnliche *Crosslink*-Fragmente und d_0/d_4 -Verhältnisse, verglichen mit dem Apoprotein-Versuch. Interessanterweise wurden Fragmente für die benachbarten Aminosäuren Lys154 und Lys155 mit dem *Crosslink*-Partner Lys491 identifiziert, welche sehr unterschiedliche BS^2G - d_0/d_4 -Verhältnisse aufzeigen ($d_0/d_4 = 0,23$ für Lys154 und $d_0/d_4 = 1,09$ für Lys155). Diese benachbarten Aminosäuren aus dem Motiv F zeigen mit ihren Aminosäureresten in entgegengesetzte Richtungen. Lys155 zeigt in das Innere der Polymerase in Richtung NTP-Eingang und Lys154 aus dem Protein heraus, wodurch es besser mit dem *Crosslinker* reagieren kann.

Auch wenn die quantitativen *Crosslink*-Experimente nur geringe Unterschiede aufweisen, konnten mit den HCV Genotypen 1b und 2a vergleichbare Ergebnisse erzielt werden. Die *Crosslink*-Experimente lieferten insbesondere für die Komplexierung von NTPs Hinweise auf mögliche Konformationsänderungen. Im Versuch von NS5B mit ssRNA zeigten sich hauptsächlich Veränderungen des d_0/d_4 -Verhältnisses zwischen Daumen- und Finger-Domäne. Am besten lässt sich dies mit der veränderten Position des Motivs F aus der MD-Simulation und der gesteigerten Flexibilität im Bereich des Linkers zwischen den Aminosäuren 530 und 545 erklären. Die Ergebnisse mit dem Inhibitor HCV-796 zeigen sehr unterschiedliche *Crosslink*-Verhältnisse. Besonders bei den Quervernetzungen mit dem Motiv

F gibt es vermehrte und verminderte Reaktionen mit BS²G im Vergleich zum Apoprotein. Die *Crosslinks* mit den Aminosäureresten innerhalb des Motivs F sind sehr variabel und verdeutlichen die Flexibilität dieses Proteinabschnitts, welcher auch in der MD-Simulation mit HCV-796 zu beobachten ist. Im Bereich der Finger-Domäne (in der Nähe von Motiv G) zeigt der Inhibitor dasselbe *Crosslink*-Verhalten wie im Versuch mit NTP.

5. ZUSAMMENFASSUNG UND AUSBLICK

Die bisherigen experimentellen Untersuchungen von Stefan Reich⁸³ an der RNA-abhängigen RNA-Polymerase von Hepatitis C wurden im Zuge dieser Arbeit mit Hilfe von computergestützten Berechnungen und Modellierungen auf molekularer Ebene analysiert. Ein besonderes Augenmerk lag dabei auf der konformationellen Analyse der Polymerase im prozessiven Zustand. Dies geschah anhand verschiedener Untersuchungsmethoden. Die Primär- und Sekundärstrukturanalysen der 189 Kristallstrukturen gaben einen guten Überblick bezüglich der konservierten Bereiche des Proteins und deren struktureller Ausprägung. Eine genaue Tertiärstrukturanalyse erfolgte anhand von RMSD-Berechnungen, Hauptkomponentenanalyse und THESEUS-Analyse. Dabei konnten spezifische Strukturunterschiede identifiziert werden. Diese befinden sich im $\Delta 1$ -Loop (einschließlich Helix A), im Motiv F, im Bereich der Daumen-Domäne und des Linkers.

Die anschließenden MD-Simulationen von NS5B mit Substraten und Inhibitoren wurden mit den bisher durchgeführten Experimenten verglichen. So konnten im Komplex mit ssRNA sowohl experimentell als auch in den Simulationen keine Konformationsänderungen in den prozessiven Zustand beobachtet werden. Auffällig in der MD-Simulation war jedoch die Positionsveränderung des Motives F in Richtung RNA-Tunnel. Die experimentelle Untersuchung der Bindung von Nukleotiden an NS5B zeigte eine Veränderung der Tertiärstruktur, welche im CD-Spektrum gemessen wurde.⁸³ In der MD-Simulation konnte eine signifikante Positionsänderung der Polymerase beobachtet werden, die in einer geöffneten Konformation resultierte. Ausschlaggebend war dabei die Rotation der Daumen-Domäne, wodurch sich der β -Loop und der Linker aus dem aktiven Zentrum des Proteins bewegen konnten. Grundlegend für diese Rotation war die Veränderung der Lage des $\Delta 1$ -Loops und der Helix A, welche zeitgleich mit der Daumen-Domänen-Bewegung einherging. Im Vergleich zur Simulation mit ssRNA konnte bei Anwesenheit von NTP eine veränderte Position des Motivs F identifiziert werden, welches sich aus Richtung des RNA-Tunnels zum NTP-Eingang verlagerte. Diese Positionsänderung wird durch attraktive Wechselwirkungen zwischen den Triphosphatgruppen und den positiv geladenen Seitenketten des Motivs F induziert. Auf diese Weise könnte das Protein die Anwesenheit eines passenden NTPs registrieren und entsprechende Mechanismen einleiten, welche die dsRNA-Produktion vorantreiben. In der MD-Simulation mit NTP konnte eine starke Beweglichkeit der RNA-Binderegion mit dem Motiv G beobachtet werden. Diese Region wurde schon in den Kristallstrukturuntersuchungen von Caillet-Saguy *et al.*⁶⁶ als Hinweis zur

Konformationsänderung von NS5B beschrieben. Um zu überprüfen, ob die beobachtete Konformationsänderung von NS5B in der Simulation mit GTP der prozessiven Form entspricht, wurde die Struktur aus der MD-Simulation mit einer dsRNA-Struktur erneut berechnet. Dies führte zu Wechselwirkungen zwischen dem Motiv G und der eintretenden ssRNA, wobei die Fluktuationen dieser Region vermindert wurden.

Neben Substraten von NS5B wurden auch allosterische Inhibitoren der Polymerase untersucht. Von besonderem Interesse waren hier Interaktionen mit dem Inhibitor HCV-796, die von Reich *et al.*⁸³ experimentell untersucht wurden. CD-Messungen lieferten Hinweise darauf, dass sie Komplexbildung zu einer Konformationsänderung führte, die vergleichbar mit der Veränderung bei Anwesenheit von Nukleotiden ist. Damit im Einklang stehen die im Rahmen der vorliegenden Dissertation durchgeführten MD-Simulationen, die auf vergleichbare Bewegungsmuster der Polymerase in Anwesenheit von Inhibitor bzw. NTP hinweisen.

Weitere Inhibitoren (NNI-2, NNI-3 und NNI-5) wurden moleküldynamisch untersucht, deren MD-Simulationen immer in einer geschlossenen Form der Polymerase resultierten. Die beiden NNI-2-Inhibitoren verhinderten die Rotation der Daumen-Domäne, sowie die Positionsänderung des Linkers. Die NNI-3-Inhibitoren O8E und O8F zeigten in den MD-Simulationen Wechselwirkungen mit den Aminosäuren Lys141, Arg158, Asp220 und Asp318 im Bereich des NTP-Eingangs und blockierten damit die NTP-Bindung. Der NNI-5-Inhibitor JTK1 interagiert mit der Finger- und der Daumen-Domäne und verhinderte so die Rotation der Daumen-Domäne und die Positionsänderung des Linkers.

Um ein Modell des Volllängenproteins zu erstellen, wurde der Membrananker modelliert. Bei der Zusammenführung der beiden Protein-Bestandteile zeigte sich, dass sich Membran und Protein überschneiden, wenn die Position des Linkers aus der Kristallstruktur nicht verändert wird. In diesem Modell ist außerdem der Zugang zum RNA-Tunnel oder zum NTP-Eingang erschwert. Daher wäre eine Position des Linkers wahrscheinlicher, die sich nur an der Daumen-Domäne anlagert und nicht in das aktive Zentrum hineinreicht. In diesem Fall wäre der Abstand zwischen Membran und Protein groß genug, sodass keine Behinderungen zwischen Membran und RNA-Tunnel oder NTP-Eingang auftreten, wie in dem zweiten Modell des Volllängen-Proteins gezeigt wurde.

Die quantitativen *Crosslink*-Experimente wurden in Zusammenarbeit mit der Abteilung Pharmazeutische Chemie & Bioanalytik durchgeführt und konnten die in den MD-Simulationen vorhergesagten Konformationsänderungen bei Anwesenheit von ssRNA, NTP und HCV-796 bestätigen. Die *Crosslink*-Verhältnisse zwischen BS²G-d₀ und BS²G-d₄ zeigten

allerdings nur Unterschiede mit einem Faktor von 1,5 bis 3. Ein Problem bei diesen Experimenten scheint die Flexibilität der Polymerase zu sein. Da sich die Polymerase bei der Substratbindung in einem Gleichgewichtszustand befindet, finden auch kontinuierlich Konformationswechsel statt. Eventuell kann NS5B auch einen größeren Konformationsraum einnehmen als bisher aus den Kristallstrukturen und den MD-Simulationen ersichtlich ist.

Die Durchführung und Auswertung der *Crosslink*-Versuche sollten weiter modifiziert werden. Anstatt die *Crosslink*-Reaktionen bei Raumtemperatur durchzuführen, könnten die Reaktionen auch auf Eis und über einen längeren Zeitraum ausgeführt werden, um potentielle Konformationsübergänge zu verlangsamen. Um die *Crosslink*-Ergebnisse zu verifizieren, müssten nochmals quantitative Experimente durchgeführt werden, in denen man die Markierung mit BS²G-d₀ und -d₄ umkehrt (*label swapping* oder *reverse labeled* genannt).^{212,213} Das Referenz-*Crosslink*-Experiment wird dabei mit deuteriertem *Crosslinker* (d₄) und die Experimente bei Substratzugabe mit nicht-deuteriertem *Crosslinker* (d₀) durchgeführt. Des Weiteren wäre eine automatische Validierung des d₀/d₄-Verhältnisses, beispielsweise mit dem Programm MaxQuant²¹⁴, effizienter als die Intensitäten für jeden *Crosslink* einzeln auszuwerten. Diese Modifikationen der *Crosslink*-Versuche waren jedoch aus Zeitgründen nicht mehr möglich.

Eine weitere Möglichkeit, die *Crosslink*-Rate des Experiments zu erhöhen und die Konformationsänderungen der Polymerase besser untersuchen zu können, ist die Mutation von geeigneten Aminosäuren von NS5B. Vorzugsweise könnten diese Mutationen im Bereich der Daumen-Domäne erfolgen, da hier wenige Lysinseitenketten in den *Crosslink*-Experimenten vernetzt wurden, aber die größten konformationellen Änderungen erfolgten. Um die Struktur und Aktivität der Polymerase weiterhin zu gewährleisten, sollte dabei eine konservative Mutation vorgenommen werden, bei der Aminosäuren mit ähnlichen biophysikalischen Eigenschaften ausgetauscht werden. Diese Mutationsstudie stellt eine weitere interessante Aufgabenstellung dar, die in zukünftigen Untersuchungen zu NS5B nicht außer Acht gelassen werden sollte.

LITERATURVERZEICHNIS

1. D. Harrus, N. Ahmed-El-Sayed, P. C. Simister, S. Miller, M. Triconnet, C. H. Hagedorn, K. Mahias, F. A. Rey, T. Astier-Gin, S. Bressanelli, J. Biol. Chem. **285** (43), 32906–32918 (2010).
doi 10.1074/jbc.M110.151316.
2. World Health Organization *Hepatitis C Fact sheet N°164*,
<<http://www.who.int/mediacentre/factsheets/fs164/en/>>, vom 05.08.2014.
3. H. E. Drummer, Front. Microbiol. **5** (329), 1–10 (2014). doi 10.3389/fmicb.2014.00329.
4. K. S. Abdelwahab, Z. N. Ahmed Said, World J. Gastroenterol. **22** (2), 862–873 (2016).
doi 10.3748/wjg.v22.i2.862.
5. P. Simmonds, A. Alberti, H. J. Alter, F. Bonino, D. W. Bradley, C. Brechot, J. T. Brouwer, S. W. Chan, K. Chayama, D. S. Chen, P. Simmonds, A. Alberti, H. J. Alter, F. Bonino, D. W. Bradley, C. Brechot, J. T. Brouwer, S.-W. Chan, K. Chayama, D.-S. Chen, Q.-L. Choo, M. Colombo, H. T. M. Cuypers, T. Date, G. M. Dusheiko, J. I. Esteban, O. Fay, S. J. Hadziyannis, J. Han, A. Hatzakis, E. C. Holmes, H. Hotta, M. Houghton, B. Irvine, M. Kohara, J. A. Kolberg, G. Kuo, J. Y. N. Lau, P. N. Lelie, G. Maertens, F. McOmish, T. Miyamura, M. Mizokami, A. Nomoto, A. M. Prince, H. W. Reesink, C. Rice, M. Roggendorf, S. W. Schalm, T. Shikata, K. Shimotohno, L. Stuyver, C. Trépo, A. Weiner, P. L. Yap, M. S. Urdea, Hepatology **19** (5), 1321–1324 (1994). doi 10.1002/hep.1840190538.
6. H. W. Reesink, *Hepatitis C virus*, 2nd, rev. and enl. ed (Karger, Basel, New York, 1998).
7. D. B. Smith, J. Bukh, C. Kuiken, A. S. Muerhoff, C. M. Rice, J. T. Stapleton, P. Simmonds, Hepatology **59** (1), 318–327 (2014). doi 10.1002/hep.26744.
8. P. Simmonds, J. Bukh, C. Combet, G. Deléage, N. Enomoto, S. Feinstone, P. Halfon, G. Inchauspé, C. Kuiken, G. Maertens, M. Mizokami, D. G. Murphy, H. Okamoto, J.-M. Pawlotsky, F. Penin, E. Sablon, T. Shin-I, L. J. Stuyver, H.-J. Thiel, S. Viazov, A. J. Weiner, A. Widell, Hepatology **42** (4), 962–973 (2005).
doi 10.1002/hep.20819.
9. P. Simmonds, J. Gen. Virol. **85** (Pt 11), 3173–3188 (2004). doi 10.1099/vir.0.80401-0.
10. P. Jackowiak, K. Kuls, L. Budzko, A. Mania, M. Figlerowicz, M. Figlerowicz, Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases **21**, 67–82 (2014). doi 10.1016/j.meegid.2013.10.021.
11. J. P. Messina, I. Humphreys, A. Flaxman, A. Brown, G. S. Cooke, O. G. Pybus, E. Barnes, Hepatology **61** (1), 77–87 (2015). doi 10.1002/hep.27259.
12. C. Gremion, A. Cerny, Rev. Med. Virol. **15** (4), 235–268 (2005). doi 10.1002/rmv.466.
13. B.D. Lindenbach, H.-J. Thiel, and C.M. Rice, eds. *Flaviviridae: The Viruses and Their Replication*, 5th ed. (Lippincott-Raven Publishers, Philadelphia, 2007).
14. T. Wakita, T. Pietschmann, T. Kato, T. Date, M. Miyamoto, Z. Zhao, K. Murthy, A. Habermann, H.-G. Krausslich, M. Mizokami, R. Bartenschlager, T. J. Liang, Nat. Med. **11** (7), 791–796 (2005).
doi 10.1038/nm1268.
15. L. Krekulová, V. Reháč, L. Riley, Folia Microbiol. **51** (6), 665–680 (2006).
16. T. Asselah, L. Rubbia-Brandt, P. Marcellin, F. Negro, Gut **55** (1), 123–130 (2006).
doi 10.1136/gut.2005.069757.
17. E. Blanchard, S. Belouzard, L. Goueslain, T. Wakita, J. Dubuisson, C. Wychowski, Y. Rouillé, J. Virol. **80** (14), 6964–6972 (2006). doi 10.1128/JVI.00024-06.
18. D. Moradpour, F. Penin, C. M. Rice, Nat. Rev. Micro. **5** (6), 453–463 (2007). doi 10.1038/nrmicro1645.

19. A. J. Weiner, H. M. Geysen, C. Christopherson, J. E. Hall, T. J. Mason, G. Saracco, F. Bonino, K. Crawford, C. D. Marion, K. A. Crawford, *PNAS* **89** (8), 3468–3472 (1992). doi 10.1073/pnas.89.8.3468.
20. B. M. Weiser, T. L. Tellinghuisen, *Drug Discov. Today Technol.* **9** (3), e195–e204 (2012). doi 10.1016/j.ddtec.2011.11.002.
21. T. v. Hahn, C. M. Rice, *J. Biol. Chem.* **283** (7), 3689–3693 (2008). doi 10.1074/jbc.R700024200.
22. A. Sakai, M. Claire, K. Faulk, S. Govindarajan, S. Emerson, R. Purcell, J. Bukh, A. Sakai, M. S. Claire, K. Faulk, S. Govindarajan, S. U. Emerson, R. H. Purcell, J. Bukh, *PNAS* **100** (20), 11646–11651 (2003). doi 10.1073/pnas.1834545100.
23. M. E. Gonzalez, L. Carrasco, *FEBS Lett.* **552** (1), 28–34 (2003). doi 10.1016/S0014-5793(03)00780-4.
24. C. Lin, B.D. Lindenbach, B. M. Prágai, D. W. McCourt, C. M. Rice, *J. Virol.* **68** (8), 5063–5073 (1994).
25. S. D. Griffin, L. P. Beales, D. S. Clarke, O. Worsfold, S. D. Evans, J. Jaeger, M. P. Harris, D. J. Rowlands, *FEBS Lett.* **535** (1-3), 34–38 (2003). doi 10.1016/S0014-5793(02)03851-6.
26. A. Grakoui, D. W. McCourt, C. Wychowski, S. M. Feinstone, C. M. Rice, *PNAS* **22** (90), 10583–10587 (1993).
27. P. Gallinari, D. Brennan, C. Nardi, M. Brunetti, L. Tomei, C. Steinkühler, R. De Francesco, *J. Virol.* **72** (8), 6758–6769 (1998).
28. K. A. Salam, A. Nobuyoshi, N. Akimitsu, *BioMed Res. Int.* **2013**, 1–9 (2013). doi 10.1155/2013/467869.
29. S. I. Asabe, Y. Tanji, S. Satoh, T. Kaneko, K. Kimura, K. Shimotohno, *J. Virol.* **71** (1), 790–796 (1997).
30. D. Egger, B. Wolk, R. Gosert, L. Bianchi, H. E. Blum, D. Moradpour, K. Bienz, *J. Virol.* **76** (12), 5974–5984 (2002). doi 10.1128/JVI.76.12.5974-5984.2002.
31. M. Lundin, H. Lindstrom, C. Gronwall, M. A. A. Persson, *J. Gen. Virol.* **87** (11), 3263–3272 (2006). doi 10.1099/vir.0.82211-0.
32. R. Gosert, D. Egger, V. Lohmann, R. Bartenschlager, H. E. Blum, K. Bienz, D. Moradpour, *J. Virol.* **77** (9), 5487–5492 (2003). doi 10.1128/JVI.77.9.5487-5492.2003.
33. T. Kaneko, Y. Tanji, S. Satoh, M. Hijikata, S. Asabe, K. Kimura, K. Shimotohno, *Biochemical and biophysical research communications* **205** (1), 320–326 (1994). doi 10.1006/bbrc.1994.2667.
34. T. L. Tellinghuisen, J. Marcotrigiano, A. E. Gorbalenya, C. M. Rice, *J. Biol. Chem.* **279** (47), 48576–48587 (2004). doi 10.1074/jbc.M407787200.
35. V. Brass, E. Bieck, R. Montserret, B. Wölk, J. A. Hellings, H. E. Blum, F. Penin, D. Moradpour, *J. Biol. Chem.* **277** (10), 8130–8139 (2002). doi 10.1074/jbc.M111289200.
36. L. Huang, J. Hwang, S. D. Sharma, M. R. S. Hargittai, Y. Chen, J. J. Arnold, K. D. Raney, C. E. Cameron, *J. Biol. Chem.* **280** (43), 36417–36428 (2005). doi 10.1074/jbc.M508175200.
37. D. Moradpour, M. J. Evans, R. Gosert, Z. Yuan, H. E. Blum, S. P. Goff, B. D. Lindenbach, C. M. Rice, *J. Virol.* **78** (14), 7400–7409 (2004). doi 10.1128/JVI.78.14.7400-7409.2004.
38. N. Appel, M. Zayas, S. Miller, J. Krijnse-Locker, T. Schaller, P. Friebe, S. Kallis, U. Engel, R. Bartenschlager, *PLOS Pathog.* **4** (3), e1000035 (2008). doi 10.1371/journal.ppat.1000035.
39. S. Behrens, L. Tomei, R. de Francesco, *EMBO J.* **15** (1), 12–22 (1996).
40. D. Moradpour, V. Brass, E. Bieck, P. Friebe, R. Gosert, H. E. Blum, R. Bartenschlager, F. Penin, V. Lohmann, *J. Virol.* **78** (23), 13278–13284 (2004). doi 10.1128/JVI.78.23.13278-13284.2004.
41. T. Suzuki, *Front. Microbiol.* **3** (38), 38 (2012). doi 10.3389/fmicb.2012.00038.
42. C. T. Jones, C. L. Murray, D. K. Eastman, J. Tassello, C. M. Rice, *J. Virol.* **81** (16), 8374–8383 (2007). doi 10.1128/JVI.00690-07.
43. Y. Miyanari, K. Atsuzawa, N. Usuda, K. Watashi, T. Hishiki, M. Zayas, R. Bartenschlager, T. Wakita, M. Hijikata, K. Shimotohno, *Nat. Cell. Biol.* **9** (9), 1089–1097 (2007). doi 10.1038/ncb1631.

44. Y. Ma, J. Yates, Y. Liang, S. M. Lemon, M. Yi, *J. Virol.* **82** (15), 7624–7639 (2008). doi 10.1128/JVI.00724-08.
45. T. L. Tellinghuisen, K. L. Foss, J. Treadaway, *PLOS Pathog.* **4** (3), e1000032 (2008). doi 10.1371/journal.ppat.1000032.
46. J. M. Werner, J. Vermehren, *Transplantation* **100** (2), 260–261 (2016). doi 10.1097/TP.0000000000001092.
47. European Association for the Study of the Liver, *J. Hepatol. (Journal of Hepatology)* **63** (1), 199–236 (2015). doi 10.1016/j.jhep.2015.03.025.
48. American Association for the Study of Liver Diseases, *Hepatology* **62** (3), 932–954 (2015). doi 10.1002/hep.27950.
49. T. Asselah, N. Boyer, D. Saadoun, M. Martinot-Peignoux, P. Marcellin, *Liver Int.* **36 Suppl 1**, 47–57 (2016). doi 10.1111/liv.13027.
50. V. Lohmann, F. Korner, U. Herian, R. Bartenschlager, *J. Virol.* **71** (11), 8416–8428 (1997).
51. H. Ago, T. Adachi, A. Yoshida, M. Yamamoto, N. Habuka, K. Yatsunami, M. Miyano, *Structure* **7** (11), 1417–1426 (1999).
52. S. Bressanelli, L. Tomei, A. Roussel, I. Incitti, R. L. Vitale, M. Mathieu, R. de Francesco, F. A. Rey, *PNAS* **96** (23), 13034–13039 (1999).
53. C. A. Lesburg, M. B. Cable, E. Ferrari, Z. Hong, A. F. Mannarino, P. C. Weber, *Nat. Struct. Biol.* **6** (10), 937–943 (1999). doi 10.1038/13305.
54. S. Bressanelli, L. Tomei, F. A. Rey, R. de Francesco, *J. Virol.* **76** (7), 3482–3492 (2002).
55. S. Chinnaswamy, A. Murali, P. Li, K. Fujisaki, C. C. Kao, *J. Virol.* **84** (12), 5923–5935 (2010). doi 10.1128/JVI.02446-09.
56. Y.-C. Kim, W. K. Russell, C. T. Ranjith-Kumar, M. Thomson, D. H. Russell, C. C. Kao, *J. Biol. Chem.* **280** (45), 38011–38019 (2005). doi 10.1074/jbc.M508145200.
57. D. O'Farrell, R. Trowbridge, D. Rowlands, J. Jager, *J. Mol. Biol.* **326** (4), 1025–1035 (2003).
58. N. Verdaguer, C. Ferrer-Orta, *Structure* **20** (9), 1448–1450 (2012). doi 10.1016/j.str.2012.08.014.
59. B. K. Biswal, M. M. Cherney, M. Wang, L. Chan, C. G. Yannopoulos, D. Bilimoria, O. Nicolas, J. Bedard, M. N. G. James, *J. Biol. Chem.* **280** (18), 18202–18210 (2005). doi 10.1074/jbc.M413410200.
60. P. Labonté, V. Axelrod, A. Agarwal, A. Aulabaugh, A. Amin, P. Mak, *J. Biol. Chem.* **277** (41), 38838–38846 (2002). doi 10.1074/jbc.M204657200.
61. A. A. van Dijk, E. V. Makeyev, D. H. Bamford, *J. Gen. Virol.* **85** (5), 1077–1093 (2004). doi 10.1099/vir.0.19731-0.
62. P. Simister, M. Schmitt, M. Geitmann, O. Wicht, U. H. Danielson, R. Klein, S. Bressanelli, V. Lohmann, *J. Virol.* **83** (22), 11926–11939 (2009). doi 10.1128/JVI.01008-09.
63. D. Garriga, C. Ferrer-Orta, J. Querol-Audi, B. Oliva, N. Verdaguer, *J. Mol. Biol.* **425** (13), 2279–2287 (2013). doi 10.1016/j.jmb.2013.03.034.
64. A. Nikonov, E. Juronen, M. Ustav, *J. Biol. Chem.* **283** (35), 24089–24102 (2008). doi 10.1074/jbc.M803422200.
65. E. Sesmero, I. F. Thorpe, *Viruses* **7** (7), 3974–3994 (2015). doi 10.3390/v7072808.
66. C. Caillet-Saguy, P. C. Simister, S. Bressanelli, *J. Mol. Biol.* **414** (3), 370–384 (2011). doi 10.1016/j.jmb.2011.10.001.
67. S. Reich, R. P. Golbik, R. Geissler, H. Lilie, S.-E. Behrens, *J. Biol. Chem.* **285** (18), 13685–13693 (2010). doi 10.1074/jbc.M109.082206.
68. Abdelrahman S. Mayhoub, *Bioorgan. Med. Chem.* **20** (10), 3150–3161 (2012). doi 10.1016/j.bmc.2012.03.049.

69. S. Di Marco, C. Volpari, L. Tomei, S. Altamura, S. Harper, F. Narjes, U. Koch, M. Rowley, R. de Francesco, G. Migliaccio, A. Carfi, *J. Biol. Chem.* **280** (33), 29765–29770 (2005). doi 10.1074/jbc.M505423200.
70. L. Tomei, S. Altamura, L. Bartholomew, A. Biroccio, A. Ceccacci, L. Pacini, F. Narjes, N. Gennari, M. Bisbocci, I. Incitti, L. Orsatti, S. Harper, I. Stansfield, M. Rowley, R. de Francesco, G. Migliaccio, *J. Virol.* **77** (24), 13225–13231 (2003). doi 10.1128/JVI.77.24.13225-13231.2003.
71. G. McKercher, P. L. Beaulieu, D. Lamarre, S. LaPlante, S. Lefebvre, C. Pellerin, L. Thauvette, G. Kukulj, *Nucleic Acids Res.* **32** (2), 422–431 (2004). doi 10.1093/nar/gkh160.
72. M. Wang, K. K.-S. Ng, M. M. Cherney, L. Chan, C. G. Yannopoulos, J. Bedard, N. Morin, N. Nguyen-Ba, M. H. Alaoui-Ismaili, R. C. Bethell, M. N. G. James, *J. Biol. Chem.* **278** (11), 9489–9495 (2003). doi 10.1074/jbc.M209397200.
73. R. A. Love, H. E. Parge, X. Yu, M. J. Hickey, W. Diehl, J. Gao, H. Wriggers, A. Ekker, L. Wang, J. A. Thomson, P. S. Dragovich, S. A. Fuhrman, *J. Virol.* **77** (13), 7575–7581 (2003). doi 10.1128/JVI.77.13.7575-7581.2003.
74. M. J. Sofia, W. Chang, P. A. Furman, R. T. Mosley, B. S. Ross, *J. Med. Chem.* **55** (6), 2481–2531 (2012). doi 10.1021/jm201384j.
75. P. L. Beaulieu, Y. S. Tsantrizos, *Curr. Opin. Investig. Drugs.* **5** (8), 838–850 (2004).
76. L. Tomei, S. Altamura, L. Bartholomew, M. Bisbocci, C. Bailey, M. Bosserman, A. Cellucci, E. Forte, I. Incitti, L. Orsatti, U. Koch, R. de Francesco, D. B. Olsen, S. S. Carroll, G. Migliaccio, *J. Virol.* **78** (2), 938–946 (2004).
77. J. Q. Hang, Y. Yang, S. F. Harris, V. Leveque, H. J. Whittington, S. Rajyaguru, G. Ao-Ieong, M. F. McCown, A. Wong, A. M. Giannetti, S. Le Pogam, F. Talamás, N. Cammack, I. Nájera, K. Klumpp, *J. Biol. Chem.* **284** (23), 15517–15529 (2009). doi 10.1074/jbc.M808889200.
78. I. Ando, T. Adachi, N. Ogura, Y. Toyonaga, K. Sugimoto, H. Abe, M. Kamada, T. Noguchi, *Antimicrob. Agents Chemother.* **56** (8), 4250–4256 (2012). doi 10.1128/AAC.00312-12.
79. C. C. Cheng, G. W. Shipps, Z. Yang, N. Kawahata, C. A. Lesburg, J. S. Duca, J. Bandouveres, J. D. Bracken, C.-k. Jiang, S. Agrawal, E. Ferrari, H.-C. Huang, *Bioorg. Med. Chem. Lett.* **20** (7), 2119–2124 (2010). doi 10.1016/j.bmcl.2010.02.054.
80. R. G. Gentles, S. Sheriff, B. R. Beno, C. Wan, K. Kish, M. Ding, X. Zheng, L. Chupak, M. A. Poss, M. R. Witmer, P. Morin, Y.-K. Wang, K. Rigat, J. Lemm, S. Voss, M. Liu, L. Pelosi, S. B. Roberts, M. Gao, J. F. Kadow, *Bioorg. Med. Chem. Lett.* **21** (8), 2212–2215 (2011). doi 10.1016/j.bmcl.2011.03.011.
81. C. Ferrer-Orta, A. Arias, C. Escarmis, N. Verdaguier, *Curr. Opin. Struc. Biol.* **16** (1), 27–34 (2006). doi 10.1016/j.sbi.2005.12.002.
82. J. Lescar, B. Canard, *Curr. Opin. Struc. Biol.* **19** (6), 759–767 (2009). doi 10.1016/j.sbi.2009.10.011.
83. S. Reich, *Biophysikalische Untersuchungen zur Funktion der Hepatitis C viralen RNA-abhängigen RNA-Polymerase.*, Dissertation, Martin-Luther-Universität Halle-Wittenberg, 2012.
84. D. L. Theobald, P. A. Steindel, *Bioinformatics* **28** (15), 1972–1979 (2012). doi 10.1093/bioinformatics/bts243.
85. A. P. Dempster, N. M. Laird, D. B. Rubin, *J. R. Stat. Soc. Series B Methodol.* **39** (1), 1–38 (1977).
86. J. Aldrich, *Stat. Sci.* **12** (3), 162–176 (1997). doi 10.1214/ss/1030037906.
87. K. Backhaus, B. Erichson, W. Plinke, R. Weiber, *Multivariate Analysemethoden* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2016).
88. A. R. Leach, *Molecular modelling. Principles and applications*, 2nd ed. (Prentice Hall, Harlow, England, New York, 2001).

89. K. I. Ramachandran, G. Deepa, K. Namboori, *Computational chemistry and molecular modeling. Principles and applications* (Springer, Berlin, 2008).
90. G. Náray-Szabó, *Protein modelling* (Springer International Publishing Switzerland, 2014).
91. P. Debye, *Math. Ann.* **67** (4), 535–558 (1909). doi 10.1007/BF01450097.
92. Magnus Hestenes, E. Stiefe, M. Hestenes, *J. Res. Natl. Bur. Stand.* **49** (6) (1952).
93. S. A. Adcock, J. A. McCammon, *Chem. Rev.* **106** (5), 1589–1615 (2006). doi 10.1021/cr040426m.
94. A. Kukol, *Molecular modeling of proteins* (Humana Press, Totowa, N.J., 2008).
95. M. A. González, *Collection SFN* **12**, 169–200 (2011). doi 10.1051/sfn/201112009.
96. M. P. Allen, D. J. Tildesley, *Computer simulation of liquids* (Clarendon Press; Oxford University Press, Oxford [England], New York, 1987).
97. L. Verlet, *Phys. Rev.* **159** (1), 98–103 (1967). doi 10.1103/PhysRev.159.98.
98. W. C. Swope, *J. Chem. Phys.* **76** (1), 648(Aendix) (1982). doi 10.1063/1.442716.
99. D. Beeman, *J. Comput. Phys.* **20** (2), 130–139 (1976). doi 10.1016/0021-9991(76)90059-0.
100. W. van Gunsteren, H. Berendsen, *Mol. Phys.* **34** (5), 1311–1327 (1977). doi 10.1080/00268977700102571.
101. Jean-Paul Ryckaert, Giovanni Ciccotti, Herman J.C Berendsen, *J. Comput. Phys.* (3), 327–341 (1977). doi 10.1016/0021-9991(77)90098-5.
102. H. C. Andersen, *J. Comput. Phys.*, 24–34 (1983).
103. B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, *J. Comput. Chem.* **18** (12), 1463–1472 (1997). doi 10.1002/(SICI)1096-987X(199709)18:12<1463:AID-JCC4>3.0.CO;2-H.
104. H. J. C. Berendsen, Postma, J. P. M., van Gunsteren, W. F., A. DiNola, J. R. Haak, *J. Chem. Phys.* **81** (8), 3684–3690 (1984). doi 10.1063/1.448118.
105. H. C. Andersen, *J. Chem. Phys.* **72** (4), 2384 (1980). doi 10.1063/1.439486.
106. D. J. Evans, B. L. Holian, *J. Chem. Phys.* **83** (8), 4069–4074 (1985). doi 10.1063/1.449071.
107. T. Schlick, *Molecular modeling and simulation. An interdisciplinary guide* (Springer-Verlag, New York, 2002).
108. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **79** (2), 926 (1983). doi 10.1063/1.445869.
109. M. W. Mahoney, W. L. Jorgensen, *J. Chem. Phys.* **112**, 8910–8922 (2000). doi 10.1063/1.481505.
110. B. Pullman, *Intermolecular Forces. Proceedings of the Fourteenth Jerusalem Symposium on Quantum Chemistry and Biochemistry Held in Jerusalem, Israel, April 13-16, 1981* (Springer Netherlands, Dordrecht, 2010).
111. H. J. C. Berendsen, J. R. Grigera, T. P. Straatsma, *J. Phys. Chem.* **91** (24), 6269–6271 (1987). doi 10.1021/j100308a038.
112. W. C. Still, A. Tempczyk, R. C. Hawley, T. Hendrickson, *J. Am. Chem. Soc.* **112** (16), 6127–6129 (1990). doi 10.1021/ja00172a038.
113. D. Qiu, P. S. Shenkin, F. P. Hollinger, W. C. Still, *J. Phys. Chem. A* **101** (16), 3005–3014 (1997). doi 10.1021/jp961992r.
114. D. A. Case, T. A. Darden, T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Götz, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, M.-J. Hsieh, G. Cui, Roe, D. H. Mathews, M. G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, P. A. Kollman, *AMBER 12* (University of California, San Francisco, 2012).

115. S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess, E. Lindahl, *Bioinformatics* **29** (7), 845–854 (2013).
doi 10.1093/bioinformatics/btt055.
116. I. T. Jolliffe, *Principal component analysis*, 2nd ed (Springer, New York, 2002).
117. B. Hess, *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics* **62** (6 Pt B), 8438–8448 (2000).
118. R. M. Levy, A. R. Srinivasan, W. K. Olson, J. A. McCammon, *Biopolymers* **23** (6), 1099–1112 (1984).
doi 10.1002/bip.360230610.
119. A. Kitao, N. Go, *Curr. Opin. Struc. Biol.* **9** (2), 164–169 (1999). doi 10.1016/S0959-440X(99)80023-2.
120. S. Wold, K. Esbensen, P. Geladi, *Chemometr. Intell. Lab.* **2** (1-3), 37–52 (1987). doi 10.1016/0169-7439(87)80084-9.
121. B. Hess, *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **65** (3), 31910 (2002).
doi 10.1103/PhysRevE.65.031910.
122. G. Brock, V. Pihur, S. Datta, S. Datta, *J. Stat. Softw.* **25** (4) (2008). doi 10.18637/jss.v025.i04.
123. P. J. Rousseeuw, *J. Comput. Appl. Math.* **20**, 53–65 (1987). doi 10.1016/0377-0427(87)90125-7.
124. J. C. Dunn, *J. Cybernetics* **4** (1), 95–104 (1974). doi 10.1080/01969727408546059.
125. B. Schilling, R. H. Row, B. W. Gibson, X. Guo, M. M. Young, *J. Am. Soc. Mass Spectrom.* **14** (8), 834–850 (2003). doi 10.1016/S1044-0305(03)00327-1.
126. A. Sinz, *Mass Spectrom. Rev.* **25** (4), 663–682 (2006). doi 10.1002/mas.20082.
127. M. Götze, J. Pettelkau, S. Schaks, K. Bosse, C. H. Ihling, F. Krauth, R. Fritzsche, U. Kühn, A. Sinz, *J. Am. Soc. Mass Spectrom.* **23** (1), 76–87 (2012). doi 10.1007/s13361-011-0261-2.
128. A. Sinz, *Expert Rev. Proteomics* **11** (6), 733–743 (2014). doi 10.1586/14789450.2014.960852.
129. S. Kalkhof, A. Sinz, *Anal. Bioanal. Chem.* **392** (1-2), 305–312 (2008).
130. S. Mädler, C. Bich, D. Touboul, R. Zenobi, *J. Mass Spectrom.* **44** (5), 694–706 (2009).
doi 10.1002/jms.1544.
131. J. Brunner, *Annu. Rev. Biochem.* **62** (1), 483–514 (1993). doi 10.1146/annurev.bi.62.070193.002411.
132. S. Kalkhof, A. Sinz, *Anal. Bioanal. Chem.* **392** (1-2), 305–312 (2008). doi 10.1007/s00216-008-2231-5.
133. W. D. Lehmann, H.-R. Schulten, *Chem. Unserer Zeit* **10** (5), 147–158 (1976).
doi 10.1002/ciuz.19760100504.
134. F. Lottspeich, *Bioanalytik*, 3rd ed. (Spektrum, Heidelberg, 2012).
135. M. Hesse, H. Meier, B. Zeeh, *Spektroskopische Methoden in der organischen Chemie*, 7th ed. (Thieme, Stuttgart, New York, 2005).
136. Thermo Fisher Scientific Inc., ed. *Orbitrap Fusion Tribrid Mass Spectrometer*. Product Specifications 63844. PS63844-EN 0216S (2016).
137. J. H. Gross, *Massenspektrometrie. Kapitel 4 "Massenspektrometer"*, 1., 2013. (Spektrum Akademischer Verlag, Heidelberg, Neckar, 2012).
138. H.-F. Wu, J. S. Brodbelt, *Int. J. Mass Spectrom. Ion Process.* **115** (1), 67–81 (1992). doi 10.1016/0168-1176(92)85032-U.
139. A. Makarov, E. Denisov, A. Kholomeev, W. Balschun, O. Lange, K. Strupat, S. Horning, *Anal. Chem.* **78** (7), 2113–2120 (2006). doi 10.1021/ac0518811.
140. M. Scigelova, A. Makarov, *Proteomics* **6 Suppl 2** (S2), 16–21 (2006). doi 10.1002/pmic.200600528.
141. A. Makarov, *Anal. Chem.* **72** (6), 1156–1162 (2000). doi 10.1021/ac991131p.
142. R. H. Perry, R. G. Cooks, R. J. Noll, *Mass Spectrom. Rev.* **27** (6), 661–699 (2008). doi 10.1002/mas.20186.
143. R. Matthiesen, *Mass spectrometry data analysis in proteomics* (Humana Press, Totowa, N.J., 2007).

144. Q. Hu, R. J. Noll, H. Li, A. Makarov, M. Hardman, R. Graham Cooks, *J. Mass Spectrom.* **40** (4), 430–443 (2005). doi 10.1002/jms.856.
145. R Core Team, *R: A language and environment for statistical computing.* (R Foundation for Statistical Computing., Vienna, Austria, 2014).
146. B. J. Grant, Rodrigues, Ana P C, K. M. ElSawy, J. A. McCammon, Caves, Leo S D, *Bioinformatics* **22** (21), 2695–2696 (2006). doi 10.1093/bioinformatics/btl461.
147. The PyMOL Molecular Graphics System, Version 1.5.0.4 (Schrödinger, LLC).
148. W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **14** (1), 33–38 (1996). doi 10.1016/0263-7855(96)00018-5.
149. A. Bakan, L. M. Meireles, I. Bahar, *Bioinformatics* **27** (11), 1575–1577 (2011). doi 10.1093/bioinformatics/btr168.
150. Molecular Operating Environment (MOE) 2013.08 (Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2016).
151. Tripos International, SYBYL 8.1 (1699 South Hanley Rd., St. Louis, Missouri, 63144, USA).
152. Thomas Williams, Colin Kelley, Russell Lang, Dave Kotz, John Campbell, Gershon Elber, Alexander Woo, James R. Van Zandt, Johannes Zellner, gnuplot 4.4: an interactive plotting program (2011).
153. A. M. Waterhouse, J. B. Procter, Martin, D. M. A., M. Clamp, G. J. Barton, *Bioinformatics* **25** (9), 1189–1191 (2009). doi 10.1093/bioinformatics/btp033.
154. A. Kahraman, L. Malmstrom, R. Aebersold, *Bioinformatics* **27** (15), 2163–2164 (2011). doi 10.1093/bioinformatics/btr348.
155. J. Kyte, R. F. Doolittle, *J. Mol. Biol.* **157** (1), 105–132 (1982). doi 10.1016/0022-2836(82)90515-0.
156. The UniProt Consortium, *Nucleic Acids Res.* **43** (D1), D204–D212 (2015). doi 10.1093/nar/gku989.
157. R. C. Edgar, *Nucleic Acids Res.* **32** (5), 1792–1797 (2004). doi 10.1093/nar/gkh340.
158. H. McWilliam, W. Li, M. Uludag, S. Squizzato, Y. M. Park, N. Buso, A. P. Cowley, R. Lopez, *Nucleic Acids Res.* **41** (W1), W597–W600 (2013). doi 10.1093/nar/gkt376.
159. D.A. Case, T.A. Darden, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Götz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T.Luchko, S. Gusarov, A. Kovalenko, and P.A. Kollman, *AMBER 12 Reference Manual* (University of California, San Francisco, 2012).
160. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **25** (9), 1157–1174 (2004). doi 10.1002/jcc.20035.
161. J. Wang, W. Wang, P. A. Kollman, D. A. Case, *J. Mol. Graph. Model.* **25** (2), 247–260 (2006). doi 10.1016/j.jmglm.2005.12.005.
162. K. L. Meagher, L. T. Redman, H. A. Carlson, *J. Comput. Chem.* **24** (9), 1016–1025 (2003). doi 10.1002/jcc.10262.
163. B. C. Davis, J. A. Brown, I. F. Thorpe, *Biophys. J.* **108** (7), 1785–1795 (2015). doi 10.1016/j.bpj.2015.03.005.
164. G. Pollastri, A. McLysaght, *Bioinformatics* **21** (8), 1719–1720 (2005). doi 10.1093/bioinformatics/bti203.
165. D. E. Kim, D. Chivian, D. Baker, *Nucleic Acids Res.* **32** (Web Server issue), W526–31 (2004). doi 10.1093/nar/gkh468.
166. J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, Y. Zhang, *Nat. Methods* **12** (1), 7–8 (2015). doi 10.1038/nmeth.3213.

167. S. Wu, Y. Zhang, *Nucleic Acids Res.* **35** (10), 3375–3382 (2007). doi 10.1093/nar/gkm251.
168. D. Xu, Y. Zhang, *Proteins* **80** (7), 1715–1735 (2012). doi 10.1002/prot.24065.
169. G. E. Tusnady, I. Simon, *J. Mol. Biol.* **283** (2), 489–506 (1998). doi 10.1006/jmbi.1998.2107.
170. A. Mitchell, H.-Y. Chang, L. Daugherty, M. Fraser, S. Hunter, R. Lopez, C. McAnulla, C. McMenamin, G. Nuka, S. Pesseat, A. Sangrador-Vegas, M. Scheremetjew, C. Rato, S.-Y. Yong, A. Bateman, M. Punta, T. K. Attwood, C. J. A. Sigrist, N. Redaschi, C. Rivore, I. Xenarios, D. Kahn, D. Guyot, P. Bork, I. Letunic, J. Gough, M. Oates, D. Haft, H. Huang, D. A. Natale, C. H. Wu, C. Orengo, I. Sillitoe, H. Mi, P. D. Thomas, R. D. Finn, *Nucleic Acids Res.* **43** (Database issue), D213–21 (2015). doi 10.1093/nar/gku1243.
171. M. Biasini, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. Kiefer, T. Gallo Cassarino, M. Bertoni, L. Bordoli, T. Schwede, *Nucleic Acids Res.* **42** (Web Server issue), W252–8 (2014). doi 10.1093/nar/gku340.
172. S. Wu, Y. Zhang, *Proteins* **72** (2), 547–556 (2008). doi 10.1002/prot.21945.
173. Y. Qi, X. Cheng, J. Lee, J. V. Vermaas, T. V. Pogorelov, E. Tajkhorshid, S. Park, J. B. Klauda, W. Im, *Biophys. J.* **109** (10), 2012–2022 (2015). doi 10.1016/j.bpj.2015.10.008.
174. S. C. Davison, E. D. Wills, *Biochem. J.* **140** (3), 461–468 (1974). doi 10.1042/bj1400461.
175. T. R. Schneider, *Acta Crystallogr. D Biol. Crystallogr.* **56** (Pt 6), 714–721 (2000).
176. T. R. Schneider, *Acta Crystallogr. D Biol. Crystallogr.* **58** (Pt 2), 195–208 (2002).
177. H. Okamoto, M. Kojima, S.-I. Okada, H. Yoshizawa, H. Iizuka, T. Tanaka, E. E. Muchmore, D. A. Peterson, Y. Ito, S. Mishiro, *Virology* **190** (2), 894–899 (1992). doi 10.1016/0042-6822(92)90933-G.
178. D. B. Smith, S. Pathirana, F. Davidson, E. Lawlor, J. Power, P. L. Yap, P. Simmonds, *J. Gen. Virol.* **78** (Pt 2), 321–328 (1997). doi 10.1099/0022-1317-78-2-321.
179. M. H. Powdrill, E. P. Tcheshnokov, R. A. Kozak, R. S. Russell, R. Martin, E. S. Svarovskaia, H. Mo, R. D. Kouyos, M. Götze, *PNAS* **108** (51), 20509–20513 (2011). doi 10.1073/pnas.1105797108.
180. R. M. Ribeiro, H. Li, S. Wang, M. B. Stoddard, G. H. Learn, B. T. Korber, T. Bhattacharya, J. Guedj, E. H. Parrish, B. H. Hahn, G. M. Shaw, A. S. Perelson, *PLOS Pathog.* **8** (8), e1002881 (2012). doi 10.1371/journal.ppat.1002881.
181. R. G. Gentles, M. Ding, J. A. Bender, C. P. Bergstrom, K. Grant-Young, P. Hewawasam, T. Hudyma, S. Martin, A. Nickel, A. Regueiro-Ren, Y. Tu, Z. Yang, K.-S. Yeung, X. Zheng, S. Chao, J.-H. Sun, B. R. Beno, D. M. Camac, C.-H. Chang, M. Gao, P. E. Morin, S. Sheriff, J. Tredup, J. Wan, M. R. Witmer, D. Xie, U. Hanumegowda, J. Knipe, K. Mosure, K. S. Santone, D. D. Parker, X. Zhuo, J. Lemm, M. Liu, L. Pelosi, K. Rigat, S. Voss, Y. Wang, Y.-K. Wang, R. J. Colonno, M. Gao, S. B. Roberts, Q. Gao, A. Ng, N. A. Meanwell, J. F. Kadow, *J. Med. Chem.* **57** (5), 1855–1879 (2014). doi 10.1021/jm4016894.
182. C. D. Manning, P. Raghavan, H. Schütze, *Introduction to information retrieval* (Cambridge University Press, New York, 2008).
183. A. Jacobo-Molina, J. Ding, R. G. Nanni, A. D. Clark Jr., X. Lu, C. Tantillo, R. L. Williams, G. Kamer, A. L. Ferris, P. Clark, *PNAS* **90** (13), 6320–6324 (1993).
184. H. Pelletier, M. Sawaya, A. Kumar, S. Wilson, J. Kraut, *Science* **264** (5167), 1891–1903 (1994). doi 10.1126/science.7516580.
185. H. Huang, *Science* **282** (5394), 1669–1675 (1998). doi 10.1126/science.282.5394.1669.
186. S. Doublié, S. Tabor, A. M. Long, C. C. Richardson, T. Ellenberger, *Nature* **391** (6664), 251–258 (1998). doi 10.1038/34593.
187. S. J. Johnson, J. S. Taylor, L. S. Beese, *PNAS* **100** (7), 3895–3900 (2003). doi 10.1073/pnas.0630532100.
188. S. Doublié, M. R. Sawaya, T. Ellenberger, *Structure* **7** (2), R31–R35 (1999). doi 10.1016/S0969-2126(99)80017-3.

189. J. Deval, C. M. D'Abramo, Z. Zhao, S. McCormick, D. Coutsinos, S. Hess, M. Kvaratskhelia, M. Gotte, J. Biol. Chem. **282** (23), 16907–16916 (2007). doi 10.1074/jbc.M701973200.
190. A. Murayama, L. Weng, T. Date, D. Akazawa, X. Tian, T. Suzuki, T. Kato, Y. Tanaka, M. Mizokami, T. Wakita, T. Toyoda, PLOS Pathog. **6** (4), e1000885 (2010). doi 10.1371/journal.ppat.1000885.
191. M. Schmitt, N. Scrima, D. Radujkovic, C. Caillet-Saguy, P. C. Simister, P. Friebe, O. Wicht, R. Klein, R. Bartenschlager, V. Lohmann, S. Bressanelli, J. Virol. **85** (6), 2565–2581 (2011). doi 10.1128/JVI.02177-10.
192. M. Gotte, J. J. Feld, Nat. Rev. Gastroenterol. Hepatol. **13** (6), 338–351 (2016). doi 10.1038/nrgastro.2016.60.
193. P. Gong, M. G. Kortus, J. C. Nix, R. E. Davis, O. B. Peersen, PLOS ONE **8** (5), e60272 (2013). doi 10.1371/journal.pone.0060272.
194. S. Reich, M. Kovermann, H. Lilie, P. Knick, R. Geissler, R. P. Golbik, J. Balbach, S.-E. Behrens, Biochemistry **53** (44), 7002–7012 (2014). doi 10.1021/bi5006656.
195. A. Y. M. Howe, H. Cheng, S. Johann, S. Mullen, S. K. Chundururu, D. C. Young, J. Bard, R. Chopra, G. Krishnamurthy, T. Mansour, J. O'Connell, Antimicrob. Agents Chemother. **52** (9), 3327–3338 (2008). doi 10.1128/AAC.00238-08.
196. D. V. Kumar, R. Rai, K. A. Brameld, J. R. Somoza, R. Rajagopalan, J. W. Janc, Y. M. Xia, T. L. Ton, M. B. Shaghafi, H. Hu, I. Lehoux, N. To, W. B. Young, M. J. Green, Bioorg. Med. Chem. Lett. **21** (1), 82–87 (2011). doi 10.1016/j.bmcl.2010.11.068.
197. S. S. Antonysamy, B. Aubol, J. Blaney, M. F. Browner, A. M. Giannetti, S. F. Harris, N. Hebert, J. Hendle, S. Hopkins, E. Jefferson, C. Kissinger, V. Leveque, D. Marciano, E. McGee, I. Najera, B. Nolan, M. Tomimoto, E. Torres, T. Wright, Bioorg. Med. Chem. Lett. **18** (9), 2990–2995 (2008). doi 10.1016/j.bmcl.2008.03.056.
198. Advanced Chemistry Development, Inc., Advanced Chemistry Development (ACD/Labs) Software (Toronto, ON, Canada).
199. S. Yan, G. Larson, J. Z. Wu, T. Appleby, Y. Ding, R. Hamatake, Z. Hong, N. Yao, Bioorg. Med. Chem. Lett. **17** (1), 63–67 (2007). doi 10.1016/j.bmcl.2006.09.095.
200. S. E. Boyce, N. Tirunagari, A. Niedziela-Majka, J. Perry, M. Wong, E. Kan, L. Laggacan, O. Barauskas, M. Hung, M. Fenaux, T. Appleby, W. J. Watkins, U. Schmitz, R. Sakowicz, PLOS ONE **9** (1), e84808 (2014). doi 10.1371/journal.pone.0084808.
201. B. C. Davis, I. F. Thorpe, Proteins **81** (1), 40–52 (2013). doi 10.1002/prot.24154.
202. J. M. Ontoria, E. H. Rydberg, S. Di Marco, L. Tomei, B. Attenni, S. Malancona, J. I. Martin Hernando, N. Gennari, U. Koch, F. Narjes, M. Rowley, V. Summa, S. S. Carroll, D. B. Olsen, R. de Francesco, S. Altamura, G. Migliaccio, A. Carfi, J. Med. Chem. **52** (16), 5217–5227 (2009). doi 10.1021/jm900517t.
203. H. Li, J. Tatlock, A. Linton, J. Gonzalez, A. Borchardt, P. Dragovich, T. Jewell, T. Prins, R. Zhou, J. Blazel, H. Parge, R. Love, M. Hickey, C. Doan, S. Shi, R. Duggal, C. Lewis, S. Fuhrman, Bioorg. Med. Chem. Lett. **16** (18), 4834–4838 (2006). doi 10.1016/j.bmcl.2006.06.065.
204. G. N. Anilkumar, O. Selyutin, S. B. Rosenblum, Q. Zeng, Y. Jiang, T.-Y. Chan, H. Pu, L. Wang, F. Bennett, K. X. Chen, C. A. Lesburg, J. Duca, S. Gavalas, Y. Huang, P. Pinto, M. Sannigrahi, F. Velazquez, S. Venkatraman, B. Vibulbhan, S. Agrawal, E. Ferrari, C.-k. Jiang, H.-C. Huang, N.-Y. Shih, F. George Njoroge, J. A. Kozlowski, Bioorg. Med. Chem. Lett. **22** (1), 713–717 (2012). doi 10.1016/j.bmcl.2011.10.041.
205. G. E. Tusnady, I. Simon, Bioinformatics **17** (9), 849–850 (2001). doi 10.1093/bioinformatics/17.9.849.
206. K. Mitra, I. Ubarretxena-Belandia, T. Taguchi, G. Warren, D. M. Engelman, PNAS **101** (12), 4083–4088 (2004). doi 10.1073/pnas.0307332101.

207. S. Ishido, T. Fujita, H. Hotta, *Biochem. Biophys. Res. Commun.* **244** (1), 35–40 (1998).
doi 10.1006/bbrc.1998.8202.
208. Y. Shirota, H. Luo, W. Qin, S. Kaneko, T. Yamashita, K. Kobayashi, S. Murakami, *J. Biol. Chem.* **277** (13), 11149–11155 (2002). doi 10.1074/jbc.M111392200.
209. S. Piccininni, A. Varaklioti, M. Nardelli, B. Dave, K. D. Raney, J. E. G. McCarthy, *J. Biol. Chem.* **277** (47), 45670–45679 (2002). doi 10.1074/jbc.M204124200.
210. M. Q. Muller, F. Dreiocker, C. H. Ihling, M. Schafer, A. Sinz, *Anal. Chem.* **82** (16), 6958–6968 (2010).
doi 10.1021/ac101241t.
211. P. Lössl, K. Kölbl, D. Tänzler, D. Nannemann, C. H. Ihling, M. V. Keller, M. Schneider, F. Zaucke, J. Meiler, A. Sinz, *PLOS ONE* **9** (11), e112886 (2014). doi 10.1371/journal.pone.0112886.
212. Z. A. Chen, L. Fischer, J. Cox, J. Rappsilber, *Mol. Cell. Proteomics* **15** (8), 2769–2778 (2016).
doi 10.1074/mcp.M115.056481.
213. Z. Chen, L. Fischer, S. Tahir, J.-C. Bukowski-Wills, P. Barlow, J. Rappsilber, *Wellcome Open Res.* **1**, 5 (2016). doi 10.12688/wellcomeopenres.9896.1.
214. J. Cox, M. Mann, *Nature biotechnology* **26** (12), 1367–1372 (2008). doi 10.1038/nbt.1511.

Abbildungsverzeichnis

Abbildung 1: Aufbau des HCV Partikels nach Krevulová <i>et al.</i> ¹⁵	2
Abbildung 2: Das HCV Genom. Veränderte Darstellung nach Moradpour <i>et al.</i> ¹⁸ und Weiser <i>et al.</i> ²⁰	3
Abbildung 3: Mikroskopische Aufnahme des <i>membranous web</i> in Huh-7 Zellen, welche mit dem HCV Replikon infiziert wurden (verändert aus Gosert <i>et al.</i> ³²).	5
Abbildung 4: Modellvorschlag zur Morphogenese des HCV- Partikels von Suzuki <i>et al.</i> ⁴¹	6
Abbildung 5: Auflistung der zurzeit verfügbaren verschiedenen Medikamente gegen HCV aus Asselah <i>et al.</i> ⁴⁹ . .	8
Abbildung 6: Nicht-Struktur Protein 5B des Hepatitis C Virus (PDB Index: 2XI3 ¹).....	9
Abbildung 7: Die Bindestellen der Nicht-Nukleosid-Inhibitoren von NS5B.	11
Abbildung 8: Vereinfachende Darstellung einer MD als Fließdiagramm nach A.Kukol. ⁹⁴	18
Abbildung 9: Schematische Darstellung der Vorgehensweise der Integrationsmethoden. ⁹⁶	21
Abbildung 10: Reaktionsmechanismus des aminreaktiven <i>Crosslinkers</i> BS ² G-d ⁰ /d ⁴ . ¹³²	31
Abbildung 11: Schematische Darstellung des Orbitrap Fusion Tribrid Massenspektrometer (Thermo Fisher Scientific). ¹³⁶	32
Abbildung 12: Nomenklatur der Peptidfragment-Ionen. ¹⁴³	34
Abbildung 13: Berechnung des Interdomänen-Winkels und des RNA-Tunnel-Durchmessers.....	39
Abbildung 14: Vereinfachter Überblick zum quantitativen <i>Crosslink</i> -Experiment.....	42
Abbildung 15: Sequenzalignment der Genotypen von NS5B, von denen eine Kristallstruktur existiert.	47
Abbildung 16: Sekundärstrukturelemente von NS5B in den 189 Kristallstrukturen.	48
Abbildung 17: Darstellung der Struktur und der Elektronendichte von 4NLD mit dem Fokus auf Cys146.	50
Abbildung 18: Dendrogramm der RMSD-Analyse der Kristallstrukturen von NS5B.	51
Abbildung 19: <i>Cartoon</i> -Darstellung repräsentativer Kristallstrukturen der RMSD-Einteilung in 4 Gruppen.	52
Abbildung 20: PCA der Kristallstrukturen mit Bio3D in R.....	53
Abbildung 21: Darstellung der THESEUS-Analyse der 23 Apoprotein-Strukturen von NS5B aus der PDB.....	55
Abbildung 22: Überlagerung der 23 Apoprotein-Kristallstrukturen aus der THESEUS-Analyse.....	56
Abbildung 23: <i>Root mean square deviation</i> (RMSD)-Werte der <i>backbone</i> -Atome N, C _α und C der MD-Simulationen.....	58
Abbildung 24: Darstellung des RNA-Tunnel-Durchmessers aufgetragen gegen den Interdomänen-Winkel.	59
Abbildung 25: RMSF-Vergleich der MD-Durchschnittsstrukturen von 1NB4 und 1NB7.....	60
Abbildung 26: Vergleich der MD-Durchschnittsstrukturen aus 2XI2 und 2XI3.....	61
Abbildung 27: Positionsvergleich der ssRNA aus der MD-Simulation mit 1NB7.....	63
Abbildung 28: Nicht-bindende Wechselwirkungen von NS5B und GTP aus der MD-Simulation von 2XI3.	65
Abbildung 29: Sequenz des RNA-Konstrukts im NS5B-dsRNA-Modell.	66
Abbildung 30: Überblick über die Simulation von NS5B in der geöffneten Form mit dsRNA.	66
Abbildung 31: Vergleich der MD-Simulationen von 2XI3 und der geöffneten Konformation mit dsRNA.....	67
Abbildung 32: Schematische Darstellung der Wasserstoffbrücken zwischen NS5B und der Template-RNA.....	68
Abbildung 33: Detaillierte Darstellung des Δ1-Loops mit der Helix A und dem Motiv F.....	69
Abbildung 34: Die Position des Inhibitors HCV-796 im Protein und dessen Molekularstruktur.....	70
Abbildung 35: Überblick über die Simulation von NS5B mit HCV-796.....	71
Abbildung 36: Vergleich der Durchschnittsstruktur der Simulation von 3FQK mit der Kristallstruktur.....	72

Abbildung 37: Wasserstoffbrücken-Bindungen zwischen NS5B und dem Inhibitor HCV-796 in der Simulation mit 3FQK.....	73
Abbildung 38: Vergleich der Bindestelle von HCV-796 aus den Simulationen mit 2XI3 und 3FQK.	74
Abbildung 39: Bindestelle der verwendeten NNI-2-Inhibitoren im Protein und deren Molekularstruktur.	75
Abbildung 40: Überblick über die Simulationen von NS5B mit NNI-2.....	76
Abbildung 41: Vergleich der NNI-2-Durchschnittsstrukturen mit den Kristallstrukturen.	77
Abbildung 42: Repositionierung des Inhibitors KLI während der MD-Simulation.	79
Abbildung 43: Darstellung der NNI-2-Bindestelle im Vergleich der Simulation von 2XI3 und 3UDL.	80
Abbildung 44: Positionswechsel und nichtbindende Wechselwirkungen von SX3 während der MD-Simulation.	81
Abbildung 45: Bindestelle und Molekularstruktur der in den Simulationen verwendeten NNI-3.....	82
Abbildung 46: Überblick über die Simulationen von NS5B mit NNI-3.....	83
Abbildung 47: Vergleich der NNI-3-Durchschnittsstrukturen mit der Kristallstruktur.....	84
Abbildung 48 (vorherige Seite): RMSD-Werte und Wasserstoffbrücken-Wechselwirkungen zwischen NNI 3 und NS5B aus den Simulationen mit den Kristallstrukturen 3U4O und 3U4R.....	87
Abbildung 49: Vergleichende Darstellung der Liganden der Durchschnittsstrukturen aus den Simulationen von 1NB7, 2XI3, 3U4O und 3U4R.	88
Abbildung 50: Darstellung der Kristallstruktur 3VQS, der Bindestelle und der Molekülstruktur von JT1.....	89
Abbildung 51: Überblick über die Simulation mit der Kristallstruktur 3VQS.	89
Abbildung 52: RMSF-Werte aus dem Vergleich zwischen der Durchschnittsstruktur der MD-Simulation mit 3VQS und der Kristallstruktur.....	90
Abbildung 53: Positionswechsel von JT1 in der MD-Simulation mit 3VQS.	92
Abbildung 54: Position von JT1 im geschlossenen und geöffneten Protein.....	93
Abbildung 55: Auftragung der Eigenwerte der 10 ersten Eigenvektoren (<i>principle components</i>) der Hauptkomponentenanalyse.....	95
Abbildung 56: Darstellung der ersten zwei Hauptkomponenten (PC1 und PC2) für die Moleküldynamik-Simulationen von NS5B.	96
Abbildung 57: Sequenzalignment der bisher kristallisierten NS5B-Genotypen.....	97
Abbildung 58: Verschiedene Darstellungen des modellierten Membranankers.	98
Abbildung 59: Vergleichende Darstellung der Membrananker-Modelle und der NMR-Struktur.	99
Abbildung 60: Das Ergebnis des HMMM- <i>Builders</i> ist die Modellierung des Membranankers von NS5B in einer Lipiddoppelschicht.	100
Abbildung 61: Modellierung des gesamten NS5B-Proteins.	101
Abbildung 62: Aminosäuresequenzen der NS5B-Varianten 1b, BK und 2a, JFH1.	103
Abbildung 63: 10 %-iges SDS-Gel der quantitativen <i>Crosslink</i> -Versuche mit NS5B und BS ² G d ₀ /d ₄	104
Abbildung 64: MS- und MS ² -Spektren der <i>Crosslink</i> -Versuche mit BS ² G und NS5B.....	105
Abbildung 65: Darstellung der <i>Crosslinks</i> aus den Versuchen mit den NS5B-Varianten 1b, BK und 2a, JFH1.	106
Abbildung 66: Kreisdiagramme der <i>Crosslink</i> -Versuche mit NS5B Genotyp 1b, BK und Genotyp 2a, JFH1... ..	109
Abbildung 67: Alignment der anerkannten (<i>reviewed</i>) HCV NS5B Sequenzen der Uniprot Datenbank.....	XIX
Abbildung 68: RMSF-Werte der Kristallstrukturen aus den MD-Simulationen berechnet über den Zeitraum der jeweils letzten 100 ns.....	XX
Abbildung 69: Durchmesser des dsRNA-Ausgangs anhand einzelner Messpunkte der MD-Durchschnittsstruktur von 2XI3.....	XX

Tabellenverzeichnis

Tabelle 1: Liste der Einteilungen von NS5B nach Domänen und Motiven sowie deren Funktionen.....	10
Tabelle 2: Vergleich der Sequenzidentitäten der Genotypen und Subtypen der NS5B-Kristallstrukturen.....	46
Tabelle 3: RMSD-Matrix der Durchschnittsstrukturen der Moleküldynamik-Simulationen.....	94
Tabelle 4: Auflistung der identifizierten <i>Crosslinks</i> von NS5B mit BS ² G und verschiedenen Liganden.	108
Tabelle 5: Abstandsmessung der quervernetzten Aminosäuren aus den MD-Simulationen.	110
Tabelle 6: Auflistung der Dynamikeinstellungen zum jeweiligen Dynamikprozess.	XXIII
Tabelle 7: Auflistung der 189 PDB-Einträge und deren Zusatzinformationen.....	XXIV
Tabelle 8: Liste der NTP-Interaktionen mit NS5B aus der Literatur.....	XLII
Tabelle 9: Übersicht der verwendeten Kristallstrukturen von NS5B mit den jeweiligen Inhibitoren.	XLIII
Tabelle 10: Übersicht der nichtbindenden Wechselwirkungen zwischen NS5B und den Inhibitoren.....	XLIV
Tabelle 11: Übersicht der RMSD-Werte und der Wasserstoffbrücken Wechselwirkungen der Proteine und ihrer Inhibitoren über die Simulationszeit.....	XLVII
Tabelle 12: Tabellarische Auflistung der PCA-Eigenwerte der MD-Simulationen.....	XLIX
Tabelle 13: Identifizierte intramolekulare <i>Crosslinks</i> mit BS ² G und NS5B GT-1b, BK mit verschiedenen Substraten.	L
Tabelle 14: Identifizierte intramolekulare <i>Crosslinks</i> mit BS ² G und NS5B GT-2a, JFH1 mit verschiedenen Substraten.	LI

Abkürzungsverzeichnis

ACN	Acetonitril
bidest	Zweifach destilliertes Wasser
BS ² G	<i>Crosslink</i> -Detergenz Bis(sulfosuccinimidyl)glutarat
CD	Circulardichroismus
CID	<i>collision induced dissociation</i> kollisionsinduzierte Dissoziation
XL	<i>Crosslink</i>
DMSO	Dimethylsulfoxid
dsRNA	<i>double-stranded</i> RNA Doppelstrang-RNA
ER	Endoplasmatisches Retikulum
ESI	Elektrospray-Ionisierung
Gr.	Gruppe
GT	Genotyp/ Genotypen
HB	Wasserstoffbrücke(n)
HCV	Hepatitis C Virus
HCV-796	Inhibitor der Bindestelle NNI-4 5-Cyclopropyl-2-(4-fluorphenyl)-6-[(2-hydroxyethyl)(methylsulfonyl)amino]- N-methyl-1-benzofuran-3-carboxamid
HEPES	2-(4-(2-Hydroxyethyl)-1-piperazinyl)-ethansulfonsäure
HPLC	Flüssigkeitschromatograph
JT1	auch als JTK-853 bezeichnet Inhibitor der Bindestelle NNI-5 (2R)-4-(5-Cyclopropyl[1,3]thiazolo[4,5-d]pyrimidin-2-yl)-N-[3-fluoro-4-(trifluoromethoxy)benzyl]-1-[[4-(trifluoromethyl)phenyl]sulfonyl]piperazine- 2-carboxamid
KLI	Inhibitor der Bindestelle NNI-2 3-(5-Bbenzyl-1,2,4-oxadiazol-3-yl)-6-fluoro-1-[2-fluoro-4-(trifluoromethyl)benzyl]-7-(4-methylpiperazin-1-yl)quinolin-4(1H)-on
MD	Moleküldynamik
ML	<i>maximum likelihood</i>
NNI(s)	Nicht-Nukleosid Inhibitor(en)
NS5B	Nicht-Strukturprotein 5 B
NTP	Nukleosidtriphosphat/ Nukleotid

O8E	Inhibitor der Bindestelle NNI-3 1-[(2-Aminopyridin-4-yl)methyl]-5-chloro-3-(2-oxo-1,2-dihydropyridin-3-yl)- 1H-indol-2-carboxylsäure
O8F	Inhibitor der Bindestelle NNI-3 1-[(2-Aminopyridin-4-yl)methyl]-5-chloro-N-({3- [(methylsulfonyl)amino]phenyl)sulfonyl)- 3-(2-oxo-1,2-dihydropyridin-3-yl)- 1H-indole- 2-carboxamid
PCA	<i>principle component analysis</i> Hauptkomponentenanalyse
PDB	<i>proteindatabank</i> Protein-Datenbank
RdRP	<i>RNA-dependent RNA-polymerase</i> RNA-abhängige RNA-Polymerase
rmsd/RMSD	<i>root mean square deviation</i>
rmsf/RMSF	<i>root mean square fluctuation</i>
RT	Raumtemperatur
SSE	Sekundärstrukturelement
ssRNA	<i>single-stranded RNA</i> Einzelstrang-RNA
ST	Subtyp
SX3	Inhibitor der Bindestelle NNI-2 4-Bromo-2-[[[(3R,5S)-3,5-dimethylpiperidin- 1-yl]carbonyl]anilin
TFA	Trifluoressigsäure
Thomson	Th
unrev.	<i>unreviewed</i> nicht überprüft
vgl.	vergleiche
XL	<i>Crosslink</i>

Anhang

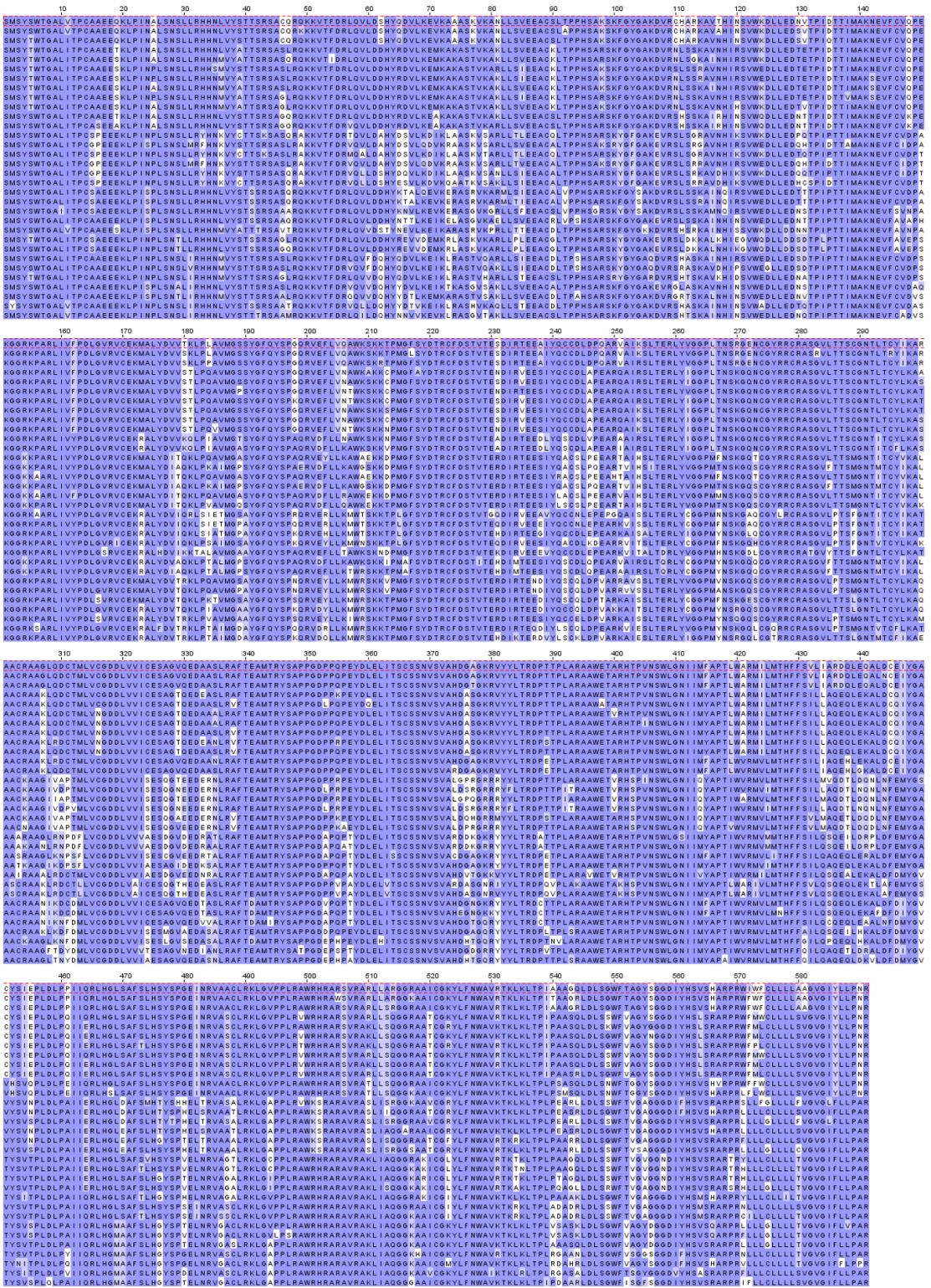


Abbildung 67: Alignment der anerkannten (reviewed) HCV NS5B Sequenzen der Uniprot Datenbank. Die Aminosäuren sind nach ihrem Konservierungsgrad lilafarbig gefärbt. Die jeweilige Uniprot-ID und der HCV-Subtyp sind links am Rand vermerkt.

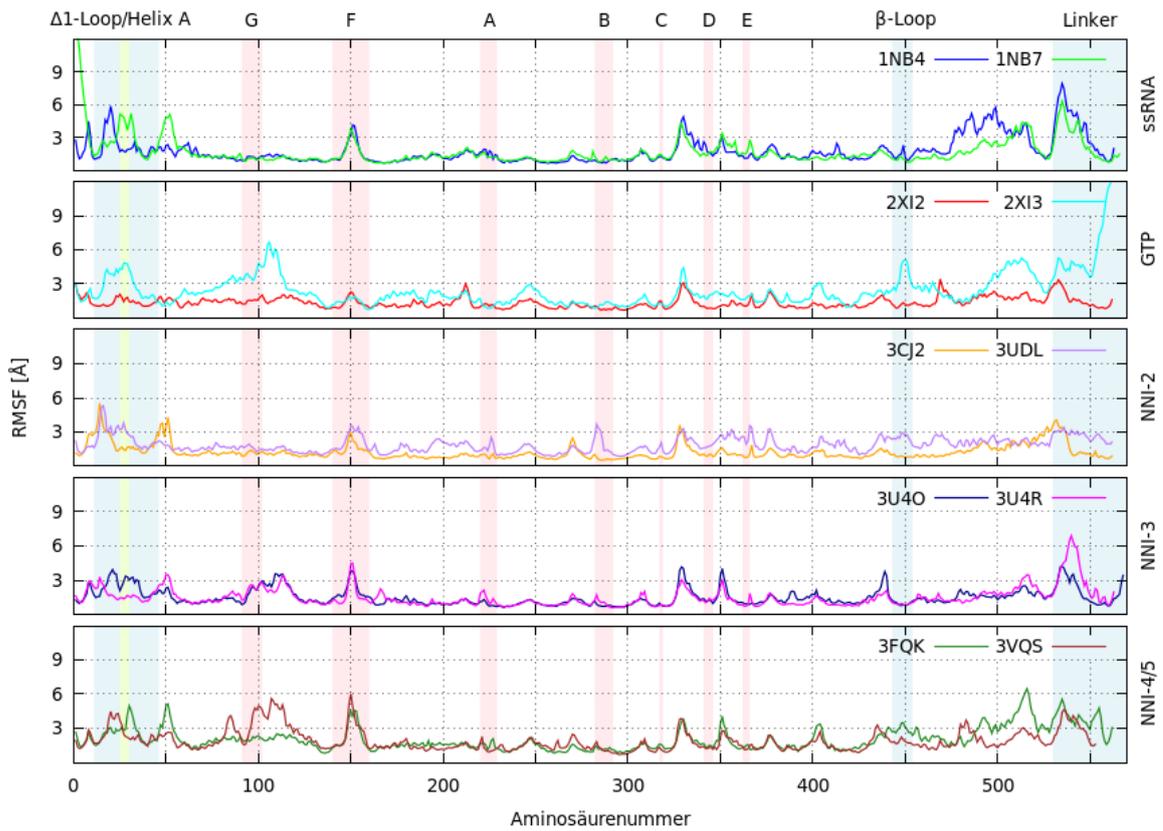


Abbildung 68: RMSF-Werte der Kristallstrukturen aus den MD-Simulationen berechnet über den Zeitraum der jeweils letzten 100 ns.

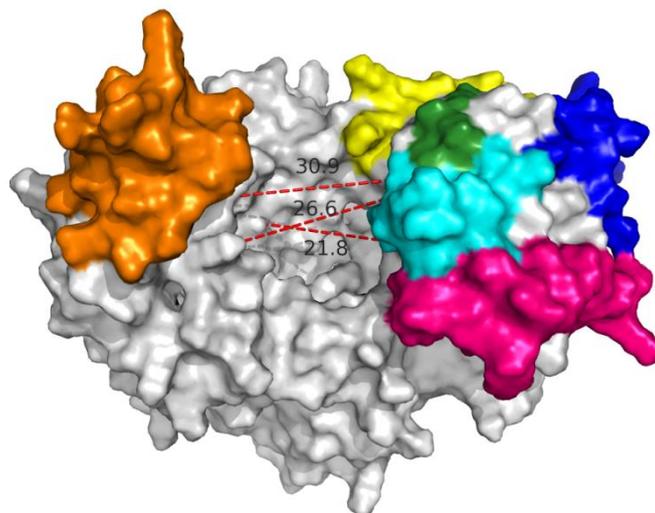


Abbildung 69: Durchmesser des dsRNA-Ausgangs anhand einzelner Messpunkte der MD-Durchschnittsstruktur von 2XI3. Die Messungen erfolgten zwischen den C_{α} -Atomen von 168 bis 406 mit 30,9 Å, 285 bis 407 mit 26,6 Å und zwischen 283 bis 411 mit 21,8 Å.

Skript I: R Skript für die RMSD und die Hauptkomponentenanalyse (PCA) der NS5B-Kristallstrukturen.

```

#### PCA mit R ####
library(muscle)
library(bio3d)
library(cIValid)

# Beliebige PDB-Datei von NS5B angeben für anschließende Blastsuche
pdb <- read.pdb("2XXD.pdb")
seq <- pdbseq(pdb)
# Blastsuche
blast <- blast.pdb(seq)
hits <- plot.blast(blast, cutoff=240)
# Herunterladen der PDB-Dateien und speichern im Ordner „raw_pdb“
files <- get.pdb(hits, path="raw_pdb")
pdbs <- pdbaln(files, exe="C:\\Users\\User\\Downloads\\muscle3.8.31.exe")
# Kernregion aller PDBs identifizieren
core <- core.find(pdbs)
core.inds <- print(core, vol=1.0)
# PDB mit Kernregionen herausschreiben
write.pdb(xyz=pdbs$xyz[1,core.inds$xyz], file="quick_core.pdb")
# Alignen der Kernregionen aller PDBs und speichern in "somedir"
xyz <- pdbfit(pdbs, core.inds, outpath="somedir")
# Berechnen der Sequenzidentität
pdbs$id <- substr(basename(pdbs$id), 1, 4)
seq_ident <- seqidentity(pdbs)
# Berechnen des RMSD
rd <- rmsd(xyz)
hist(rd, breaks=40, xlab="RMSD (Å)", main="Histogram of RMSD")
## Clusternummer bestimmen##
intern <- cIValid(rd, 2:20, cIMethod=c("hierarchical", "kmeans"), validation="internal")
plot(intern)
measures(intern)
# RMSD clustern mit 2 Gruppen (k=2) und Plotten
hc.rd <- hclust(as.dist(rd))
grps <- cutree(hc.rd, k=2)
cols <- c("darkgreen", "red")[grps]
pdbs$id <- substr(basename(pdbs$id), 1, 4)
hclustplot(hc.rd, col=cols, labels=pdbs$id, cex=0,7, ylab="RMSD (Å)", main="RMSD Cluster Dendrogramm", fillbox=FALSE)
# Identifizierung der Lücken (Gaps) im Alignment

```

```
gaps.res <- gap.inspect(pdb$ali)
gaps.pos <- gap.inspect(pdb$xyz)

### PCA ###
pc.xray <- pca.xyz(xyz[, gaps.pos$f.inds])
plot(pc.xray)
# Plot PC1 vs PC2
plot(pc.xray, pc.axes=1:2,)
# Identifizierung und Beschriftung der Komponenten
identify(pc.xray$z[,1], pc.xray$z[,2], labels=pdb$id)
### Clusternummer identifizieren ###
# für PC1 #
intern.pc1 <- clValid((pc.xray$z[,1:2]),2:20, clMethod=c("hierarchical","kmeans"), validation="internal")
optimalScores(intern.pc1)
plot(intern.pc1)
clust_PC1 <- data.frame(measures(intern.pc1))
write.table(clust_PC1,file='clust_PC1.dat', sep='\t')

# Clustern der PCA mit 4 Gruppen
hc_pc1 <- hclust(dist(pc.xray$z[,1:2]))
grps1 <- cutree(hc_pc1, k=4)
cols1 <- c("red", "orange", "blue", "darkgreen")[grps1]
hclustplot(hc.rd, col=cols1, labels=pdb$id, cex=0,7, ylab="RMSD (Å)", main="RMSD Cluster
Dendrogram", fillbox=FALSE)
plot(pc.xray, pc.axes=1:2, col=cols1)
# Identifizierung und Beschriftung der Komponenten
identify(pc.xray$z[,1], pc.xray$z[,2], labels=pdb$id)
pdb$id
# Dendrogramm erstellen
names(cols1) <- pdb$id
hclustplot(hc_pc1, colors=cols1, labels=pdb$id, ylab="Distance in PC Space", main="PC1-2", cex=0,7,
fillbox=FALSE)
```

Tabelle 6: Auflistung der Dynamikeinstellungen zum jeweiligen Dynamikprozess.

Prozess	AMBER-Einstellungen
Minimierung	imin=1, maxcyc=5000, ncyc=500, igb=5, saltcon=0,05, cut=1000, ntb=0
Aufheizen auf 300 K	imin=0, irect=0, ntx=1, dt=0,0005, nstlim=120000, ntc=2, ntf=2, ntt=1, tautp=0,5, tempi=300, temp0=0, ntp=5000, ntwx=5000, ntwr=5000, ntb=0 , igb=5, saltcon=0,05, cut=1000, rgbmax=16, ntr=1, restraint_wt=5.0, restraintmask='@N,CA,C' &end &wt type='TEMP0', istep1=0, istep2=100000, value1=0,0, value2=300, &end &wt type='END' &end
Äquilibrierung mit abnehmenden <i>restraints</i>	imin=0, irect=1, ntx=5, dt=0,001, nstlim=300000, ntc=2, ntf=2, ntt=1, tautp=0,5, temp0=300,0, ntp=5000, ntwx=5000, ntwr=5000, ntb=0, igb=5, saltcon=0,05, cut=1000, rgbmax=16, ntr=1, restraint_wt=2.0, restraintmask='@N,CA,C' &end &wt type='REST', istep1=0, istep2=300000, value1=2.0, value2=0, &end &wt type='END' &end
Äquilibrierung ohne <i>restraints</i>	imin=0, irect=1, ntx=5, nstlim=200000, dt=0,001, ntc=2, ntf=2, ntt=1, tautp=0,5, tempi=300,0, temp0=300,0, ntp=1000, ntwx=1000, ntwr=50000, ntb=0, igb=5, saltcon=0,05, cut=1000, rgbmax=16
Produktionslauf	imin=0, irect=1, ntx=5, dt=0,002, nstlim=5000000, ntc=2, ntf=2, ntt=1, tautp=0,5, temp0=300,0, ntp=5000, ntwx=5000, ntwr=50000, ntb=0, igb=5, saltcon=0,05, cut=1000, rgbmax=16

Tabelle 7: Auflistung der 189 PDB-Einträge und deren Zusatzinformationen

PDB	Genotyp	Uniprot Eintrag	Auflösung [Å]	Substituent	Aminosäureanzahl	QMEAN	Ramachandran Analyse [%]	fehlende Aminosäuren oder Seitenkette und Mutationen
1C2P	1b, BK	P26663	1,90		558	0,804	91,5	149-153, 564-570, Selenomethionin statt Methionin
1CSJ	1b, BK	P26663	2,80		530	0,801	88,3	33, 45, 100, 131, 151, 157, 164, 180, 260, 268, 316, 347, 403, 404, 436, 438, 448, 457, 469, 476, Selenomethionin statt Methionin
1GX5	1b, BK	P26663	1,70	GTP	530	0,666	92,2	531-536, Selenomethionin statt Methionin
1GX6	1b, BK	P26663	1,85	UTP	531	0,670	92,2	531-536, Selenomethionin statt Methionin
1NB4	1b, HCJ4	O92972	2,00		566	0,866	89,1	567-570
1NB6	1b, HCJ4	O92972	2,60	UTP	566	0,842	88,1	567-570
1NB7	1b, HCJ4	O92972	2,90	ssRNA 3'-UUUU-5'	566	0,844	81,8	567-570
1NHU	1b, BK	P26663	2,00	NNI-2, Propionsäure-Derivat	558	0,789	91,5	149-153, 564-570
1NHV	1b, BK	P26663	2,90	NNI-2, Thiophen-Derivat	558	0,767	85,4	149-153, 564-570
1OS5	1b, BK	P26663	2,20	NNI-2, Dihydropyran-Derivat	563	0,847	92,0	564-577
1QUV	1b, BK	P26663	2,50		553	0,799	86,1	544-546, 557-578
1YUY	2a, HCJ6	P26660	1,90		558	0,637	90,8	149-153, 564-570
1YV2	2a, HCJ6	P26660	2,50		548	0,790	89,7	549-570

PDB	Genotyp	Uniprot Eintrag	Auflösung [Å]	Substituent	Aminosäureanzahl	QMEAN	Ramachandran Analyse [%]	fehlende Aminosäuren oder Seitenkette und Mutationen
1YVF	<i>unreviewed</i>	O93077	2,50	NNI-3, Benzoyl-Derivat	564	0,821	87,6	564-576
1YVX	2a, HCJ6	P26660	2,00	NNI-2, Thiophen-Derivat	548	0,798	90,3	549-570
1YVZ	2a, HCJ6	P26660	2,20	NNI-2, Thiophen-Derivat	548	0,798	89,7	549-570
1Z4U	<i>unreviewed</i>	O93077	2,80	NNI-3, Benzoyl-Derivat	563	0,821	85,9	564-576
2AWZ	1b, BK	P26663	2,15	NNI-3, Thiazolidin-Derivat	557	0,825	92,1	541-580
2AX0	1b, BK	P26663	2,00	NNI-3, Thiazolidin-Derivat	554	0,827	91,6	541-580
2AX1	1b, BK	P26663	2,10	NNI-3, Thiazolidin-Derivat	554	0,815	92,2	541-580
2BRK	1b, BK	P26663	2,30	NNI-1, Indol-Derivat	512	0,797	91,6	22-35, 148-152, 532-536
2BRL	1b, BK	P26663	2,40	NNI-1, Indol-Derivat	512	0,785	90,7	22-35, 148-152, 532-536
2D3U	1b	Q99AU2	2,00	NNI-2, Thiophen-Derivat	558	0,788	92,0	149-153, 564-570
2D3Z	1b	Q99AU2	1,80	NNI-2, Thiophen-Derivat	560	0,788	91,5	151-153, 564-570
2D41	1b	Q99AU2	2,10	NNI-2, Thiophen-Derivat	558	0,778	90,7	149-153, 564-570
2DXS	1b, BK	P26663	2,20	NNI-1, Benzoxazepin-	517	0,786	92,2	15-36, 148-152, 545-552

PDB	Genotyp	Uniprot Eintrag	Auflösung [Å]	Substituent	Aminosäureanzahl	QMEAN	Ramachandran Analyse [%]	fehlende Aminosäuren oder Seitenkette und Mutationen
				Derivat				
2FVC	1b	Q99AU2	2,00	NNI-3, Chinolin-Derivat	560	0,843	90,8	-
2GC8	1b	Q99AU2	2,20	NNI-3, Prolin-Derivat	557	0,831	92,5	541-545, 563-570
2GIQ	1b, BK	P26663	1,65	NNI-3 (im Paper steht NNI-2; stimmt aber nicht), Benzothiadiazin	557	0,810	92,8	149-153, 561-568
2GIR	1b, BK	P26663	1,90	NNI-2, Thiophen-Derivat	557	0,803	93,4	149-153
2HAI	1b	Q99AU2	1,58	NNI-2, Isopropoxyphenyl-Derivat	562	0,807	93,0	563-576
2HWH	1b, BK	P26663	2,30	NNI-2, Thiazol-Derivat	559	0,762	89,4	149-153, 565-570
2HWI	1b, BK	P26663	2,00	NNI-2, Thiazol-Derivat	559	0,772	90,7	149-153, 565-570
2I1R	1b, BK	P26663	2,20	NNI-2, Thiazol-Derivat	559	0,757	90,1	149-153, 565-570
2IJN	1b	Q99AU2	2,20	NNI-3, Thioxopropanamid-Derivat	559	0,770	89,9	149-153, 565-570
2JC0	1b, BK	P26663	2,20	NNI-3, Pyrrolidin-Derivat	557	0,777	92,3	149-153, 563-570
2JC1	1b, BK	P26663	2,00	NNI-3, Pyrrolidin-Derivat	549	0,778	92,6	149-153, 542-549, 563-570
2O5D	1b, BK	P26663	2,20	NNI-2, Thiazol-	559	0,780	90,8	149-153, 565-570

PDB	Genotyp	Uniprot Eintrag	Auflösung [Å]	Substituent	Aminosäureanzahl	QMEAN	Ramachandran Analyse [%]	fehlende Aminosäuren oder Seitenkette und Mutationen
				Derivat				
2QE2	1b	Q99AU2	2,90	NNI-3, Benzoessäure-Derivat	521	0,738	86,5	1-29, 98-100, 130-132, 151, 347-348, 544 -546, 563-570
2QE5	1b	Q99AU2	2,60	NNI-3, Benzoessäure-Derivat	559	0,800	87,6	544-546, 563-578
2WCX	1b, BK	P26663	2,00	NNI-1, Pyrrol-Derivat	512	0,812	92,7	1, 23-36, 149-153, 533-537
2WHO	1b, BK	P26663	2,00	NNI-2, Chinolin-Derivat	531	0,785	91,3	532-536
2WRM	1b, BK	P26663	1,95	NNI-2, Chinolin-Derivat	531	0,815	90,9	532-537
2XHU	1b, HCJ4	O92972	2,29		562	0,805	75,8	151, 564-578
2XHV	1b, HCJ4	O92972	1,90		563	0,877	92,4	564-578
2XHW	1b, HCJ4	O92972	2,66		560	0,795	90,5	151-153, 564-578
2XI2	1a, H77	P27958	1,80		562	0,840	92,6	563-576
2XI3	1a, H77	P27958	1,70	GTP	562	0,856	92,5	563-576
2XWH	2a, HCJ6	P26660	1,80		561	0,634	92,9	151, 152
2XWY	1b, BK	P26663	2,53	NNI-1, Indol-Derivat	507	0,737	89,1	15-35, 150-152, 532-536
2XXD	2a, JFH1	Q99IB8	1,88		563	0,829	92,3	-
2XYM	2a, JFH1	Q99IB8	1,77		563	0,820	91,9	-
2YOJ	1b, HCJ4	O92972	1,76	NNI-3, Indol-Derivat	563	0,787	92,9	149-153, 569-585
2ZKU	1b, BK	P26663	1,95		562	0,752	91,5	150-153, 567-570

PDB	Genotyp	Uniprot Eintrag	Auflösung [Å]	Substituent	Aminosäureanzahl	QMEAN	Ramachandran Analyse [%]	fehlende Aminosäuren oder Seitenkette und Mutationen
3BR9	1b, BK	P26663	2,30	NNI-3, Benzothiadiazin-Derivat	557	0,791	91,4	149-153, 563-578
3BSA	1b, BK	P26663	2,30	NNI-3, Benzothiadiazin-Derivat	557	0,798	90,8	149-153, 563-578
3BSC	1b, BK	P26663	2,65	NNI-3, Benzothiadiazin-Derivat	557	0,799	91,0	149-153, 563-578
3CDE	1b, BK	P26663	2,10	NNI-3, Benzothiadiazin-Derivat	557	0,807	93,4	149-153, 563-578
3CIZ	1b, BK	P26663	1,87	NNI-2, Benzoessäure-Derivat	562	0,791	92,6	151, 152, 564-570
3CJ0	1b, BK	P26663	1,90	NNI-2, Pyridin-Derivat	562	0,774	92,5	151-152, 564-570
3CJ2	1b, BK	P26663	1,75	NNI-2, Piperidin-Derivat	557	0,798	92,6	149-153, 563-570
3CJ3	1b, BK	P26663	1,90	NNI-2, Pyrrolidin-Derivat	560	0,793	92,6	149-152, 564-570
3CJ4	1b, BK	P26663	2,07	NNI-2, Piperidin-Derivat	558	0,794	92,9	149-153, 563-570
3CJ5	1b, BK	P26663	1,92	NNI-2, Piperidin-Derivat	562	0,795	92,4	151-152, 564-570
3CO9	1b, BK	P26663	2,10	NNI-3, Benzothiadiazin-Derivat	562	0,866	92,6	563-578, R544Q
3CS0	1b, HCJ4	O92972	2,71	NNI-3, Benzodiazepin-	561	0,796	91,3	151-152, 564-578

PDB	Genotyp	Uniprot Eintrag	Auflösung [Å]	Substituent	Aminosäureanzahl	QMEAN	Ramachandran Analyse [%]	fehlende Aminosäuren oder Seitenkette und Mutationen
				Derivat				
3CVK	1b, BK	P26663	2,31	NNI-3, Benzothiadiazin-Derivat	557	0,787	91,3	149-153, 563-578
3CWJ	1b, BK	P26663	2,40	NNI-3, Benzothiazin-Derivat	557	0,788	91,2	149-153, 563-578
3D28	1b, BK	P26663	2,20	NNI-3, Benzisothiazol-Derivat	557	0,775	90,6	149-153, 563-578
3D5M	1b, BK	P26663	2,30	NNI-3, Benzisothiazol-Derivat	559	0,806	92,1	150-152, 563-578
3E51	1b, BK	P26663	1,90	NNI-3, Benzothiadiazin-Derivat	557	0,806	92,7	149-153, 563-578
3FQK	1b, BK	P26663	2,20	NNI-4, Benzofuran-Drivat	557	0,764	92,2	149-153, 563-570
3FQL	1b, con1	Q9WMX2	1,80	NNI-4, Benzofuran-Drivat	559	0,773	93,4	148-153, 566-570
3FRZ	1b, BK	P26663	1,86	NNI-2, Pyrimidin	562	0,854	92,8	563-576
3G86	1b, BK	P26663	2,20	NNI-3, Benzothiazin-Derivat	557	0,785	92,3	149-153, 563-570
3GNV	1b, HCJ4	O92972	2,75	NNI-3, Benzodiazepin-Derivat	563	0,845	91,8	564-578
3GNW	1b, HCJ4	O92972	2,39	NNI-3, Benzodiazepin-	560	0,791	92,1	150-152, 564-578

PDB	Genotyp	Uniprot Eintrag	Auflösung [Å]	Substituent	Aminosäureanzahl	QMEAN	Ramachandran Analyse [%]	fehlende Aminosäuren oder Seitenkette und Mutationen
				Derivat				
3GOL	1b,HCI4	O92972	2,85	NNI-3, Benzodiazepin- Derivat	557	0,770	89,2	150-152, 544-546, 564-578
3GSZ	2b, HC-J8	P26661	1,90		558	0,652	92,4	149-153
3GYN	1b, BK	P26663	2,15	NNI-3, Benzothiadiazin- Derivat	557	0,802	92,7	149-153, 563-578
3H2L	1b, BK	P26663	1,90	NNI-3, Benzothiadiazin- Derivat	557	0,784	92,3	149-153, 563-578
3H59	1b, BK	P26663	1,95	NNI-3, Benzothiazin- Derivat	557	0,784	92,1	149-153, 563-570
3H5S	1b, BK	P26663	2,10	NNI-3, Pyrrol- Derivat	558	0,791	92,7	149-153, 563-570
3H5U	1b, BK	P26663	2,00	NNI-3, Benzisothiazol- Derivat	559	0,786	92,7	149-152, 564-570
3H98	1b, BK	P26663	1,90	NNI-3, Benzothiadiazin- Derivat	550	0,784	91,8	149-152, 540-570
3HHK	1b	Q99AU2	1,70	NNI-3, Benzothiadiazin- Derivat	560	0,859	92,8	-
3HKW	1a, isolate 1	P26664	1,55	NNI-3, Benzodiazepin- Derivat	561	0,781	93,3	150-151, 564-578
3HKY	1b, HCI4	O92972	1,90	NNI-3, Benzodiazepin-	563	0,872	92,1	564-578

PDB	Genotyp	Uniprot Eintrag	Auflösung [Å]	Substituent	Aminosäureanzahl	QMEAN	Ramachandran Analyse [%]	fehlende Aminosäuren oder Seitenkette und Mutationen
				Derivat				
3HVO	2b, HC-J8	P26661	2,00	NNI-2, Chinolin-Derivat	559	0,664	92,6	150-153
3I5K	2a, JFH1	Q99IB8	2,20		566	0,819	92,5	-
3IGV	1b, BK	P26663	2,60	NNI-3, Benzothiadiazin-Derivat	557	0,804	92,0	149-153, 563-578
3LKH	1b, HCJ4	O92972	2,05	NNI-5, Benzamid-Derivat	554	0,751	92,3	149-153, 541-544, 564-576
3MF5	1b, BK	P26663	2,00	NNI-2, Thiophen-Derivat	557	0,807	92,3	149-153, 563-570
3MWV	1b, HCJ4	O92972	2,20		559	0,797	91,5	150-153, 564-576
3MWW	1b, HCJ4	O92972	2,80	NNI-1, Indol-Derivat	559	0,785	86,3	150-153, 564-576
3PHE	<i>unreviewed</i>	D0PY27	2,20	NNI-2, Dihydrochinolin-Derivat	558	0,766	92,2	149-153, 564-570
3Q0Z	1b, con1	Q9WMX2	2,29	NNI-1, Benzodiazepin-Derivat	551	0,837	92,8	15-36
3QGD	1b, con1	Q9WMX2	2,60	NNI-1, NNI-5, Benzodiazepin- und Piperazin-Derivat	538	0,844	92,5	15-42, 567-573
3QGE	1b, con1	Q9WMX2	3,00	NNI-1, NNI-5, Benzodiazepin- und Piperazin-Derivat	541	0,854	92,6	15-37, 565-573

PDB	Genotyp	Uniprot Eintrag	Auflösung [Å]	Substituent	Aminosäureanzahl	QMEAN	Ramachandran Analyse [%]	fehlende Aminosäuren oder Seitenkette und Mutationen
3QGF	1b, con1	Q9WMX2	2,45	NNI-1, NNI-5, Benzodiazepin- und Piperazin-Derivat	534	0,804	91,2	15-37, 565-573
3QGG	1b, con1	Q9WMX2	1,97	NNI-1, NNI-5, Benzodiazepin- und Piperazin-Derivat	531	0,772	90,9	15-37, 150, 556-573
3QGH	1a, isolate 1	P26664	2,14	NNI-5, Piperazin-Derivat	559	0,797	93,1	152-153, 562-570
3QGI	1a, isolate 1	P26664	1,80	NNI-5, Piperazin-Derivat	561	0,810	94,1	152-153, 564-570
3SKA	1b, HCJ4	O92972	1,73	NNI-3, Indol-Derivat	563	0,794	93,9	149-153, 569-576
3SKE	1b, HCJ4	O92972	1,97	NNI-3, Indol-Derivat	563	0,804	93,8	149-153, 569-576
3SKH	1b, HCJ4	O92972	1,97	NNI-3, Indol-Derivat	558	0,795	90,5	149-153, 564-576
3TYQ	1b, HCJ4	O92972	1,60	NNI-3, Indol-Derivat	563	0,786	93,3	149-153, 569-576
3TYV	1b, HCJ4	O92972	1,65	NNI-3, Indol-Derivat	563	0,789	93,4	149-153, 569-576
3U4O	1b, HCJ4	O92972	1,77	NNI-3, Indol-Derivat	563	0,800	93,5	149-153, 569-578
3U4R	1b, HCJ4	O92972	2,00	NNI-3, Indol-Derivat	558	0,806	92,2	149-153, 564-578
3UDL	1b, BK	P26663	2,17	NNI-2, Chinolin-Derivat	555	0,784	91,6	23-25, 149-153, 564-570

PDB	Genotyp	Uniprot Eintrag	Auflösung [Å]	Substituent	Aminosäureanzahl	QMEAN	Ramachandran Analyse [%]	fehlende Aminosäuren oder Seitenkette und Mutationen
3UPH	1b, HCJ4	O92972	2,00	NNI-3, Indol-Derivat	563	0,790	92,8	149-153, 569-576
3UPI	1b, HCJ4	O92972	2,00	NNI-3, Indol-Derivat	563	0,782	93,3	149-153, 569-585
3VQS	1b, BK	P26663	1,90	NNI-5, Piperazin-Derivat	553	0,851	93,5	554-578
4ADP	2a, HCJ6	P26660	1,90		564	0,812	92,3	565-577
4AEP	2a, JFH1	Q99IB8	1,80		563	0,828	92,5	564-577
4AEX	2a, JFH1	Q99IB8	2,41		563	0,829	92,4	564-577
4DRU	1b, HCJ4	O92972	2,10	NNI-1, Indol-Derivat	549	0,817	91,9	23-32, 149-152
4E76	2a, JFH1	Q99IB8	2,50		551	0,630	89,5	552-578, $\Delta 8 \beta$ -Loop Deletion
4E78	2a, JFH1	Q99IB8	2,90	dsRNA RNA- <i>template</i> 3'-ACCGG-5'	545	0,620	89,6	546-578, $\Delta 8 \beta$ -Loop Deletion
4E7A	2a, JFH1	Q99IB8	3,00	dsRNA RNA- <i>template</i> 3'-UGGCC-5'	543	0,658	88,9	544-578, $\Delta 8 \beta$ -Loop Deletion
4EAW	1b, HCJ4	O92972	2,00	NNI-3, Benzodiazaphosphinin-Derivat	559	0,807	93,1	150-152
4EO6	1b, BK	P26663	1,79	NNI-2, Thiophen-Derivat	557	0,800	92,1	149-153, 563-570
4EO8	1b, BK	P26663	1,80	NNI-2, Thiophen-Derivat	557	0,812	92,3	149-153, 563-570
4GMC	1b, HCJ4	O92972	2,70	NNI-1, Indol-	559	0,771	85,5	150-153, 564-576

PDB	Genotyp	Uniprot Eintrag	Auflösung [Å]	Substituent	Aminosäureanzahl	QMEAN	Ramachandran Analyse [%]	fehlende Aminosäuren oder Seitenkette und Mutationen
				Derivat				
4IH5	1b, BK	P26663	1,90	NNI-3, Pyrimidin-Derivat	558	0,808	92,8	149-153, 563-570
4IH6	1b, BK	P26663	2,20	NNI-3, Imidazolidin-Derivat	558	0,814	92,5	149-153, 563-570
4IH7	1b, BK	P26663	2,30	NNI-3, Pyridin-Derivat	558	0,772	91,6	149-153, 563-570
4IZ0	1b, HCJ4	O92972	2,22	NNI-2 und NNI-3, Sulfonamid-Derivat	559	0,784	90,1	150-153, 564-576
4J02	1b, HCJ4	O92972	2,00	NNI-2, Indol-Derivat	559	0,778	91,6	150-153, 564-576
4J04	1b, HCJ4	O92972	2,00	NNI-2, Benzoesäure-Derivat	559	0,758	91,4	150-153, 564-576
4J06	1b, HCJ4	O92972	2,00	NNI-2, Benzoesäure-Derivat	559	0,776	92,1	150-153, 564-576
4J08	1b, HCJ4	O92972	2,10	NNI-2, Benzoesäure-Derivat	559	0,773	91,0	150-153, 564-576
4J0A	1b, HCJ4	O92972	2,40	NNI-2, Benzoesäure-Derivat	559	0,757	88,0	150-153, 564-576
4JJS	1b, HCJ4	O92972	2,20	NNI-2, Benzoesäure-Derivat	559	0,790	90,7	150-153, 564-576
4JJU	1b, HCJ4	O92972	1,91	NNI-2,	559	0,779	92,2	150-153, 564-576

PDB	Genotyp	Uniprot Eintrag	Auflösung [Å]	Substituent	Aminosäureanzahl	QMEAN	Ramachandran Analyse [%]	fehlende Aminosäuren oder Seitenkette und Mutationen
				Chinazolin-Derivat				
4JTW	1b, HCJ4	O92972	3,00	NNI-2, Chinazolin-Derivat	560	0,765	88,4	150-152, 564-576
4JTY	1b, HCJ4	O92972	2,60	NNI-2, Dihydrochinazolin-Derivat	560	0,766	89,0	150-152, 564-576
4JTZ	1b, HCJ4	O92972	2,80	NNI-2, Dihydrochinazolin-Derivat	560	0,768	87,0	150-152, 564-576
4JU1	1b, HCJ4	O92972	2,90	NNI-2, Chinazolin-Derivat	560	0,761	88,8	150-153, 564-576
4JU2	1b, HCJ4	O92972	2,70	NNI-2, Chinazolin-Derivat	560	0,777	88,6	150-152, 564-576
4JU3	1b, HCJ4	O92972	2,00	NNI-2, Benzoesäure-Derivat	559	0,793	91,3	150-153, 564-576
4JU4	1b, HCJ4	O92972	2,40	NNI-2, Benzoesäure-Derivat	559	0,768	90,4	150-153, 564-576
4JU6	1b, HCJ4	O92972	2,20	NNI-2, Benzoesäure-Derivat	559	0,780	89,9	150-153, 564-576
4JU7	1b, HCJ4	O92972	2,20	NNI-2, Benzoesäure-Derivat	559	0,780	89,9	150-153, 564-576
4JVQ	1b, HCJ4	O92972	2,40	NNI-2, Benzoesäure-Derivat	560	0,777	91,4	150-152, 564-576
4JY0	1b, HCJ4	O92972	2,20	NNI-3, Pyrazin-	559	0,790	90,9	150-153, 564-576

PDB	Genotyp	Uniprot Eintrag	Auflösung [Å]	Substituent	Aminosäureanzahl	QMEAN	Ramachandran Analyse [%]	fehlende Aminosäuren oder Seitenkette und Mutationen
				Derivat				
4JY1	1b, HCJ4	O92972	2,60	NNI-2, Thiophen-Derivat	559	0,778	85,3	150-153, 564-576
4KAI	1b, BK	P26663	2,30	NNI-4, Benzofuran-Derivat	552	0,753	92,5	150-153, 541-548, 565-576, L47Q, F101Y, K114R
4KB7	1b, BK	P26663	1,85	NNI-4, Benzofuran-Derivat	546	0,741	92,6	149-153, 563-570, L47Q, F101Y, K114R, N316Y
4KBI	1b, BK	P26663	2,06	NNI-4, Benzofuran-Derivat	544	0,734	92,6	149-153, 541-554, 564-578, L47Q, F101Y, K114R, N316Y
4KE5	1b, BK	P26663	2,11	NNI-4, Benzofuran-Derivat	548	0,736	93,0	150-153, 541-551, 564-578, L47Q, F101Y, K114R, N316Y
4KHM	<i>unreviewed</i>	B1PPP0	2,45	NNI-4, Benzofuran-Derivat	562	0,833	93,6	564-579, F101Y, C110S, R113S, K114R
4KHR	<i>unreviewed</i>	B1PPP0	1,70	NNI-4, Benzofuran-Derivat	562	0,830	93,6	563-578, F101Y, C110S, R113S, K114R, C316Y
4MIA	1b, BK	P26663	2,80	NNI-3, Dihydropyridin-Derivat	547	0,752	90,0	149-153, 540-570
4MIB	1b, BK	P26663	2,30	NNI-3, Pyrrolidin-Derivat	557	0,794	92,5	149-153, 563-570
4MK7	1b, BK	P26663	2,80	NNI-3, Pyridin-Derivat	557	0,774	91,0	149-153, 563-570
4MK8	1b, BK	P26663	2,09	NNI-3, Dihydropyridin-	557	0,800	91,6	149-153, 563-570

PDB	Genotyp	Uniprot Eintrag	Auflösung [Å]	Substituent	Aminosäureanzahl	QMEAN	Ramachandran Analyse [%]	fehlende Aminosäuren oder Seitenkette und Mutationen
				Derivat				
4MK9	1b, BK	P26663	2,05	NNI-3, Dihydropyridin-Derivat	557	0,797	92,6	149-153, 563-570
4MKA	1b, BK	P26663	2,05	NNI-3, Chromen-Derivat	557	0,794	92,6	149-153, 563-570
4MKB	1b, BK	P26663	1,90	NNI-3, Dihydropyridazin-Derivat	557	0,798	92,4	149-153, 563-570
4MZ4	1b, HCJ4	O92972	1,63	NNI-3, Indol-Derivat	563	0,793	93,3	149-153, 569-576
4NLD	1b, con1	Q9WMX2	2,75	NNI-1, NNI-4, Benzazepin- und Benzofuran-Derivat	500	0,738	91,6	0-53, 269-273, 559-573
4OBC	2a, JFH1	Q99IB8	2,50	MES bindet im RNA-Eingangsbereich; noch nicht als Inhibitorstelle identifiziert.	554	0,789	91,8	554-578
4OOW	1b, HCJ4	O92972	2,57	Catechol, RNA-Tunnel; gegenüber von MES aus 4OBC; event. neue Inhibitorstelle	555	0,799	90,2	150-152, 542-545, 563-578
4RY4	1b, HCJ4	O92972	2,59		566	0,864	87,1	567-570, Y448F
4RY5	1b, HCJ4	O92972	2,71	UTP	565	0,844	88,6	566-570, W550N
4RY6	1b, HCJ4	O92972	2,52		565	0,824	88,8	566-570, W550A

PDB	Genotyp	Uniprot Eintrag	Auflösung [Å]	Substituent	Aminosäureanzahl	QMEAN	Ramachandran Analyse [%]	fehlende Aminosäuren oder Seitenkette und Mutationen
4RY7	1b, HCJ4	O92972	3,00		565	0,815	89,8	566-570, D559E
4TLR	1b, con1	NA	1,86	NNI-2, NNI-4, Thiophen- und Benzofuran-Derivat	564	0,819	93,6	565-578
4TN2	1b, BK	P26663	2,70	NNI-2, Thiophen-Derivat	561	0,776	91,0	149-153, 568-578
4TXF	Veröffentlichung zurückgezogen							
4TXS	<i>unreviewed</i>	D0PY27	2,78	NNI-3, (4-hydroxyphenyl)acetonitril	544	0,754	89,6	23-27, 149-153, 540-547, 563-566
4TY8	<i>unreviewed</i>	D0PY27	2,78	NNI-3, Chromen-Derivat	544	0,763	91,0	23-27, 149-153, 540-547, 563-566
4TY9	<i>unreviewed</i>	D0PY27	2,78	NNI-2, Pyridin	544	0,771	89,8	23-27, 149-153, 540-547, 563-566
4TYA	<i>unreviewed</i>	D0PY27	2,94	NNI-3, Benzoesäure-Derivat	544	0,751	90,4	23-27, 149-153, 540-547, 563-566
4TYB	<i>unreviewed</i>	D0PY27	2,93	NNI-3, Morphonlin-Derivat	544	0,773	90,0	23-27, 149-153, 541-547, 563-566
4WT9	2a, JFH1	Q99IB8	2,50		544	0,645	91,4	19, 43, 51, 69, 106, 151, 206, 212, 270, 330, 377, 379, 401, 441, 531, 535, 541, 543, 544, 546, 548, 552-578, Δ8 β-Loop Deletion , E86Q, E87Q, C223H, V321I
4WTA	2a, JFH1	Q99IB8	2,80	dsRNA RNA- <i>primer</i> 5'-CAAAUUU-	535	0,642	91,7	43, 50, 51, 73, 106, 151, 206, 212, 270, 330, 377, 379, 401, 440, 441, 465, 501, 531-533, 535, 543-578, Δ8 β-Loop Deletion , S15G, E86Q,

PDB	Genotyp	Uniprot Eintrag	Auflösung [Å]	Substituent	Aminosäureanzahl	QMEAN	Ramachandran Analyse [%]	fehlende Aminosäuren oder Seitenkette und Mutationen
				3'				E87Q, C223H, V321I
4WTC	2a, JFH1	Q99IB8	2,75	dsRNA RNA-primer 5'-AGAAAUUU -3'	534	0,619	91,4	43, 50, 51, 69, 73, 106, 125, 151, 206, 212, 270, 330, 377, 379, 381, 401, 440, 441, 465, 531-533, 535, 541-578, Δ8 β-Loop Deletion , S15G, E86Q, E87Q, C223H, V321I
4WTD	2a, JFH1	Q99IB8	2,70	dsRNA RNA-primer 5'-AUAAAUUU -3'	534	0,634	91,9	18, 43, 50, 51, 69, 73, 77, 106, 114, 124, 151, 206, 212, 270, 330, 352, 377, 379, 381, 401, 440, 441, 465, 501, 531-533, 535, 541-578, Δ8 β-Loop Deletion , S15G, E86Q, E87Q, C223H, V321I
4WTE	2a, JFH1	Q99IB8	2,90	dsRNA RNA-primer 5'-ACAAAUUU -3'	535	0,628	90,2	43, 50, 51, 69, 73, 77, 106, 114, 124, 131, 151, 202, 206, 212, 270, 330, 352, 377, 379, 381, 401, 440, 441, 465, 501, 531-533, 535, 541, 543-578, Δ8 β-Loop Deletion , S15G, E86Q, E87Q, C223H, V321I
4WTF	2a, JFH1	Q99IB8	2,65	dsRNA RNA-primer 5'-CAAAAUUU- -3'	536	0,635	92,8	18, 43, 47, 50, 51, 69, 73, 77, 106, 151, 206, 212, 270, 330, 377, 379, 401, 440, 441, 465, 501, 531, 533, 535, 541, 543-578, Δ8 β-Loop Deletion , S15G, E86Q, E87Q, C223H, V321I
4WTG	2a, JFH1	Q99IB8	2,90	dsRNA RNA-primer 5'-CAAAAUUU- -3'	535	0,637	89,7	43, 50, 51, 69, 73, 106, 124, 151, 206, 212, 270, 330, 377, 379, 401, 440-442, 465, 501, 531-533, 535, 543-578, Δ8 β-Loop Deletion , S15G, E86Q, E87Q, C223H,

PDB	Genotyp	Uniprot Eintrag	Auflösung [Å]	Substituent	Aminosäureanzahl	QMEAN	Ramachandran Analyse [%]	fehlende Aminosäuren oder Seitenkette und Mutationen
								V321I
4WTI	2a, JFH1	Q99IB8	2,80	dsRNA RNA- <i>primer</i> 5'-ACGG-3'	553	0,765	91,4	18, 19, 43, 50, 51, 69, 106, 151, 206, 212, 270, 330, 333, 377, 379, 401, 440, 446-448, 450, 501, 531, 533, 535, 543, 544, 552, 553-578, S15G, E86Q, E87Q, C223H, V321I
4WTJ	2a, JFH1	Q99IB8	2,20	dsRNA RNA- <i>primer</i> 5'-AUCC-3'	555	0,793	93,9	19, 51, 100, 106, 131, 151, 212, 270, 333, 377, 379, 401, 446, 531, 535, 543, 555-578, S15G, E86Q, E87Q, C223H, V321I
4WTK	2a, JFH1	Q99IB8	2,50	dsRNA RNA- <i>primer</i> 5'-AGCC-3'	553	0,767	91,8	14,43, 47, 50, 51, 69, 106, 124, 151, 206, 212, 270, 330, 333, 377, 379, 401, 446, 531, 532, 533, 535, 541, 543, 544, 552-578, S15G, E86Q, E87Q, C223H, V321I
4WTL	2a, JFH1	Q99IB8	2,00	dsRNA RNA- <i>primer</i> 5'-UAGG-3'	553	0,786	93,6	19, 43, 51, 100, 114, 151, 184, 377, 379, 401, 441, 531, 535, 541, 543, 544, 553-578, S15G, E86Q, E87Q, C223H, V321I
4WTM	2a, JFH1	Q99IB8	2,15	dsRNA RNA- <i>primer</i> 5'-UAGG-3'	553	0,785	92,2	19, 43, 50, 51, 69, 106, 124, 151, 206, 212, 270, 330, 333, 377, 379, 401, 446, 448, 531, 532, 535, 541, 543, 544, 559-578, S15G, E86Q, E87Q, C223H, V321I
5CZB	1b, HCJ4	O92972	1,96	NNI-2, Pyrazol-Derivat	555	0,779	93,5	150-152, 542-555
5TRH	1b-con1	Q9WMX2	2,70	NNI-1, NNI-3, Benzodiazepin und Benzoessäure-Derivat	574	0,839	90,6	15-37, 98, 131, 148, 150, 154, 535, 568
5TRK	1b-con1	Q9WMX2	2,06	NNI-1, NNI-3,	574	0,836	93,5	15-36, 57, 100, 131, 147, 148, 151,

PDB	Genotyp	Uniprot Eintrag	Auflösung [Å]	Substituent	Aminosäureanzahl	QMEAN	Ramachandran Analyse [%]	fehlende Aminosäuren oder Seitenkette und Mutationen
				Benzodiazepin und Benzen-Derivat				309, 355, 568
5TRI	1b-con1	Q9WMX2	2,30	NNI-1, NNI-3, Benzodiazepin und Indol-Derivat	574	0,849	91,5	15-37, 57, 100, 131, 147, 148, 151, 158, 355, 523, 565-573
5TRJ	1b-con1	Q9WMX2	2,57	NNI-1, NNI-3, Benzodiazepin und Benzoessäure-Derivat	574	0,821	91,6	15-36, 57, 100, 141, 147, 148, 150, 151, 307, 333, 341, 355, 440, 501, 533, 535, 568

Tabelle 8: Liste der NTP-Interaktionen mit NS5B aus der Literatur.

Bindestelle	Interagierende Proteinseitenketten	Referenz
katalytisches Zentrum	C site: Asp220, Asp318, Asp319 P site: Arg158, Ser367, Arg386, Thr390, Arg394 I site: Arg48, Lys51, Lys151, Lys155, Arg158, Leu159, Asp225	Bressanelli <i>et al.</i> ⁵⁴
UTP	Arg41, Arg48, Lys51, Lys141, Arg158, Arg222, Ser282, Thr283, Asn291, Ser556	O'Farrel <i>et al.</i> ⁵⁷
GTP	Arg158, 282, Arg367, Arg386, Glu446, Ser556, Gly557	Harrus <i>et al.</i> ¹

Tabelle 9: Übersicht der verwendeten Kristallstrukturen von NS5B mit den jeweiligen Inhibitoren.

PDB-Eintrag	Genotyp	Ligandenklasse	Ligandenname	IC ₅₀ [μM]	Liganden-Struktur
3UDL	1b, BK	NNI-2, Chinolon-Derivat	KLI	0,014	
3CJ2	1b, BK	NNI-2, Antranilsäureamid	SX3	17	
3U4O	1b, HC-J4	NNI-3, Indol-Derivat	O8E	0,053	
3U4R	1b, HC-J4	NNI-3, Indol-Derivat	O8F	0,039	
3FQK	1b, BK	NNI-4, Benzofuran-Derivat	HCV-796	0,081	
3VQS	1b, BK	NNI-5, Piperazin-Derivat	JT1 (JTK-853)	0,018	

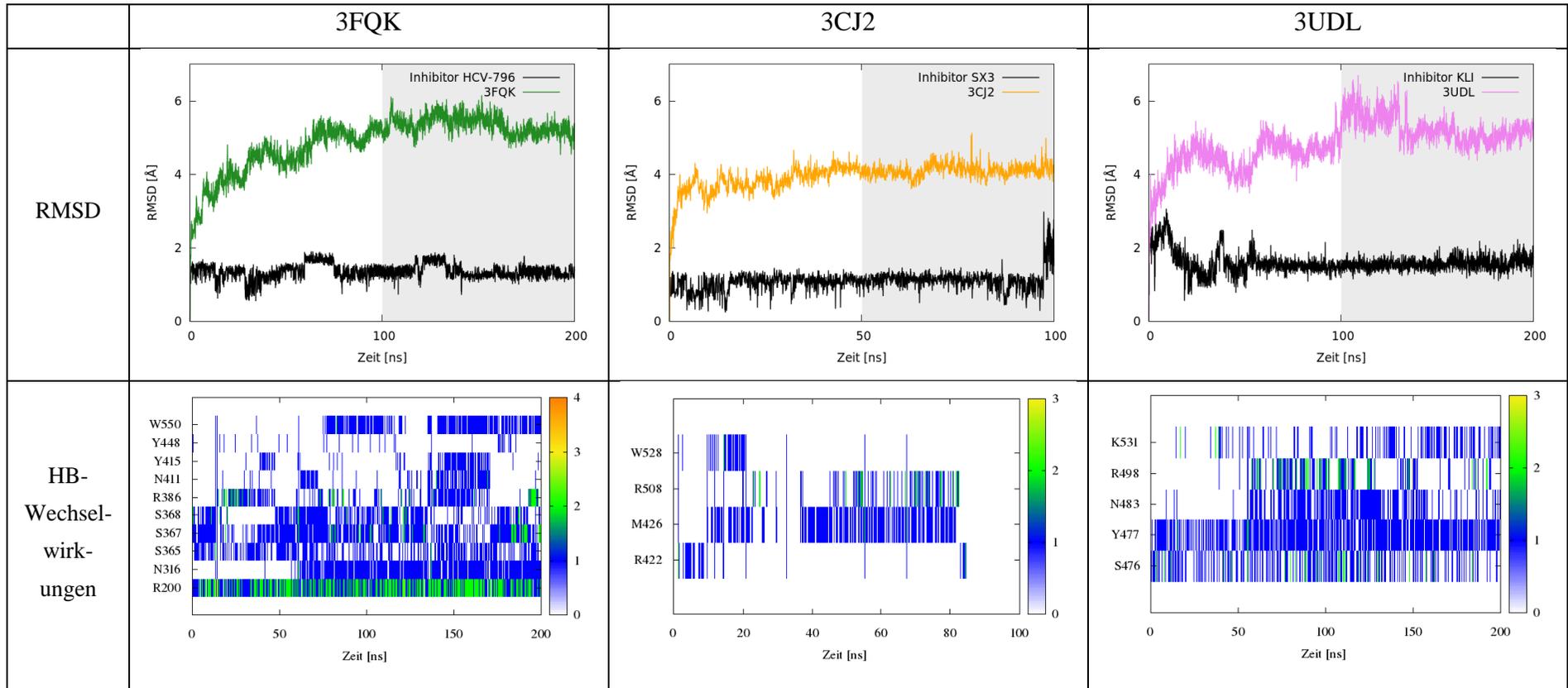
Tabelle 10: Übersicht der nichtbindenden Wechselwirkungen zwischen NS5B und den Inhibitoren.

HB-Akzeptor	HB-Donor	%-Anteil	Abstand	Winkel	
3FQK/HCV-796					
HCV-796-O3	R200-NH1	0,54	3,04	145,43	
HCV-796-O3	R200-NH2	0,77	2,91	153,59	
HCV-796-O2	N316-ND2	0,85	3,01	161,19	
HCV-796-O5	R386-NH12	0,33	2,96	150,70	
HCV-796-O5	N411-ND2	0,26	3,07	158,09	
S365-OG	HCV-796-N1	0,44	3,05	145,76	
HCV-796-O5	S367-N	0,50	2,96	153,54	
HCV-796-O5	S367-OG	0,26	2,92	157,89	
HCV-796-O5	S368-OG	0,36	2,91	158,50	
HCV-796-O4	W550-NE1	0,63	3,01	152,89	
3CJ2/SX3					
SX3-O1	R508-NE	0,42	3,05	149,53	
M426-O	SX3-N2	0,56	3,02	156,76	

HB-Akzeptor	HB-Donor	%-Anteil	Abstand	Winkel	
3UDL/KLI					
KLI-N3	S476-N	0,38	3,25	146,01	
KLI-O2	S476-N	0,09	3,32	150,86	
KLI-O1	Y477-N	0,79	3,01	152,23	
N483-OD1	KLI-N5	0,53	2,89	161,82	
KLI-N4	K531-NZ	0,38	3,26	155,44	
3U40/O8E					
O8E-N3	R158-NE	0,47	3,14	152,99	
O8E-O3	R386-NH1	0,93	2,85	151,86	
O8E-O3	R386-NH2	0,86	2,88	149,18	
O8E-O3	R394-NH1	0,83	3,06	155,21	
O8E-O2	R394-NH1	0,77	2,96	147,14	
O8E-O2	R394-NH2	0,93	2,84	154,26	
O8E-O1	Y448-N	0,84	3,08	160,04	

HB-Akzeptor	HB-Donor	%-Anteil	Abstand	Winkel	
3U4R/O8F					
D220-O	O8F-N4	0,78	3,02	155,44	
O8F-O5	D318-N	0,99	2,90	164,58	
O8F-O6	Q446-NE2	0,92	2,86	160,20	
O8F-O1	G449-N	0,85	2,98	152,36	
O8F-O6	S556-OG	0,88	2,69	154,09	
S556-OG	O8F-N5	0,77	3,17	153,26	
O8F-O3	G557-N	0,71	2,87	145,84	
3QVS/JT1 (JTK-853)					
S196-OG	JT1-N1	0,28	3,28	143,87	
JT1-O4	S196-N	0,99	2,81	163,28	
JT1-O4	S196-OG	0,86	2,82	152,00	
JT1-O3	Q199-NE2	0,99	2,83	158,54	
JT1-N3	Y452-OH	0,77	3,12	151,84	
JT1-N4	Y452-OH	0,16	3,24	144,32	
JT1-O2	R465-NH1/2	0,85	3,04	150,36	
JT1-N5	F551-N	0,79	3,21	162,89	
JT1-N4	F551-N	0,17	3,11	143,18	

Tabelle 11: Übersicht der RMSD-Werte und der Wasserstoffbrücken Wechselwirkungen der Proteine und ihrer Inhibitoren über die Simulationszeit. Die RMSD-Berechnungen erfolgten für das Proteinrückgrat bzw. für die schweren Atome des Inhibitors.



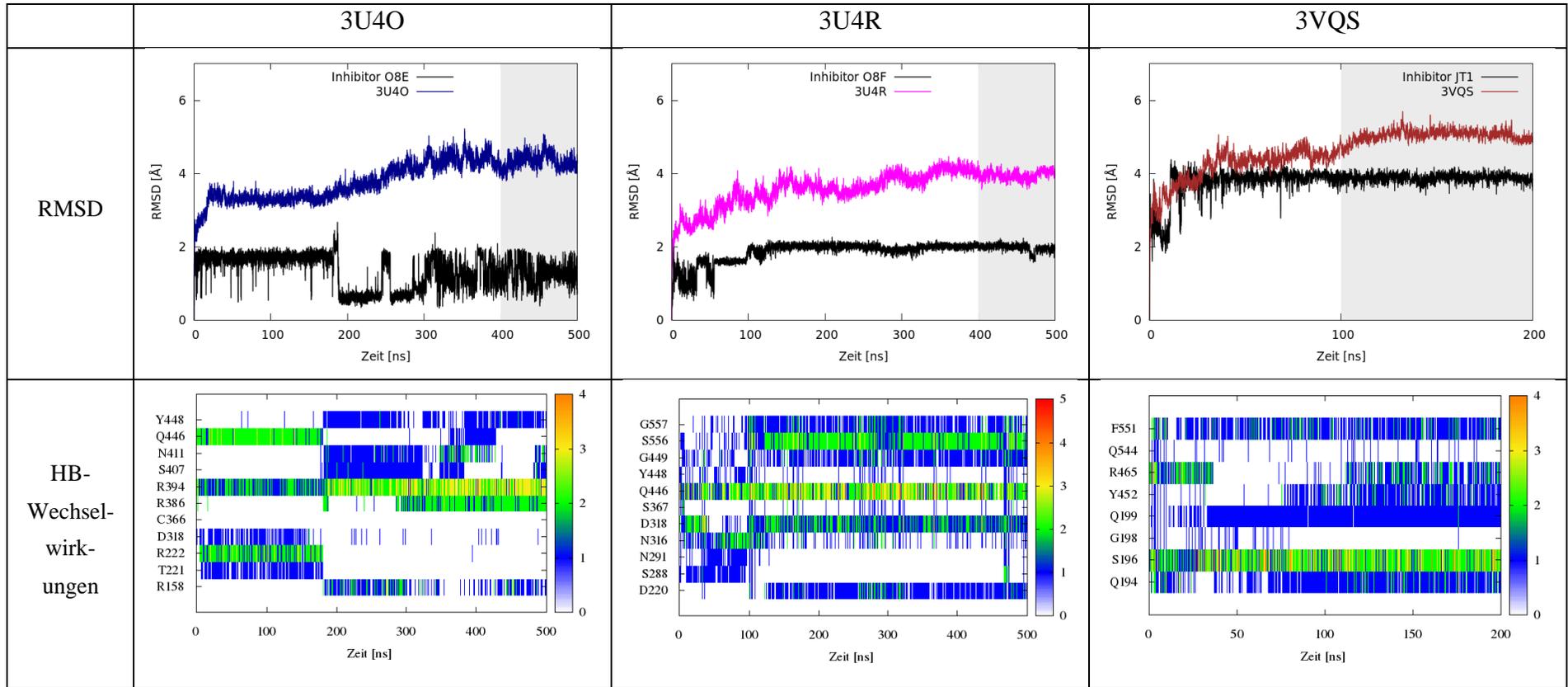


Tabelle 12: Tabellarische Auflistung der PCA-Eigenwerte der MD-Simulationen.

A) Die Berechnung erfolgte mit den Koordinaten der ersten 100 ns bzw. 50 ns für 3CJ2. **B)** Die Berechnung erfolgte mit den Koordinaten der letzten 100 ns bzw. 50 ns für 3CJ2.

A	1NB4	1NB7	2XI2	2XI3	3FQK	3CJ2	3UDL	3U4O	3U4R	3VQS
1	1086,23	1093,31	396,147	2284,9	1139,34	648,589	701,49	684,987	288,459	739,511
2	402,264	309,51	165,055	539,079	282,503	241,776	290,708	217,676	172,268	259,881
3	139,53	199,148	122,496	274,194	262,338	118,179	187,202	102,867	79,3495	257,149
4	104,509	142,607	66,9064	182,643	140,47	103,074	141,139	70,307	68,5897	176,589
5	77,148	80,1719	55,3892	121,429	109,523	63,6128	95,1672	55,062	42,8527	98,8407
6	48,766	69,7069	42,5587	90,0013	83,5259	50,4972	55,4926	48,7301	36,1589	74,775
7	44,8064	50,0844	37,7943	83,7158	79,0598	39,3407	49,7454	44,8229	29,9537	74,0103
8	42,1556	45,3815	31,0081	65,3025	70,8604	36,1789	44,3518	37,6506	27,8836	63,7917
9	39,168	39,182	22,9627	57,5356	56,0726	29,6881	42,8584	31,8488	25,541	54,7504
10	34,0814	35,9943	22,0922	56,3978	49,4205	24,3901	38,4167	26,5145	22,8835	46,6216

B	1NB4	1NB7	2XI2	2XI3	3FQK	3CJ2	3UDL	3U4O	3U4R	3VQS
1	520,324	394,53	145,465	1050	426,759	263,056	743,472	148,845	197,377	453,393
2	159,369	229,959	115,521	578,27	182,304	151,59	162,216	79,315	114,793	276,128
3	114,897	129,084	75,0901	222,013	130,078	75,6977	97,8651	59,9462	70,4298	125,394
4	77,0024	83,7102	56,5613	166,855	73,2497	56,7561	80,1382	46,0336	52,8971	67,2731
5	54,6765	56,5045	47,4726	156,155	53,3722	47,3987	62,2882	40,0251	35,9108	52,7482
6	42,0231	55,2911	39,7086	69,9507	49,4638	44,4635	35,585	34,172	31,5865	45,8313
7	35,397	46,0567	34,4844	62,817	42,6371	34,1092	34,2067	30,2265	27,2588	37,0868
8	33,0998	37,2311	27,2505	50,3627	37,4638	30,2995	32,8644	24,4228	23,9297	33,8486
9	30,6708	36,3696	23,3114	42,8669	31,83	25,7333	26,682	23,38	20,3687	26,0801
10	26,8008	30,984	20,6148	40,6479	29,9259	20,8284	26,3523	19,5151	15,995	23,923

Tabelle 13: Identifizierte intramolekulare *Crosslinks* mit BS²G und NS5B GT-1b, BK mit verschiedenen Substraten.

Die intrapeptidalen *Crosslinks* sind durch die Sequenzen und die quervernetzten Aminosäuren aufgelistet. Daneben wurde vermerkt, mit welchem d₀/d₄-Verhältnis der *Crosslink* identifiziert wurde.

GT-1b					NS5B		NTP		ssRNA		HCV-796				
m/z gemessen	[M + H] ⁺ berechnet	Ladung	Massenunterschied (ppm)		Peptidsequenz	Peptidsequenz	obere Bande	untere Bande	obere Bande	untere Bande	obere Bande	untere Bande			
472,083	2827,464	6	0,22	d ₀	K106 101[FGYGAKDVR]109	K114 110[NLSSKAVNHIHSVWK]124		0,60				0,99	0,51		
567,104	2831,489	5	-0,01	d ₄				0,70	0,44	0,17	0,55	0,77	0,60	0,24	
708,628	2831,489	4	-0,03						0,52	0,35		0,75	0,59	0,35	
553,96	1659,864	3	1,42	d ₄	K106 101[FGYGAKDVR]109	K114 110[NLSSK]114			0,55						
577,72	2884,573	5	0,1	d ₄	K98 91[LTPPHSAKSK]100	K114 110[NLSSKAVNHIHSVWK]124		0,48		0,27		0,44	0,25		
721,9	2884,573	4	1,1								0,26		0,32		
962,196	2884,573	3	-0,3								0,43				
357,209	1069,611	3	-0,05	d ₀	K155 155[KPAR]158	S44 44[SAGLR]48		0,77			2,35	1,60	2,17	1,36	
358,55	1073,636	3	0,11	d ₄						0,78	0,50				
318,385	1587,897	5	0,12	d ₀	K155 155[KPAR]158	K51 49[QKKVTFDR]56		1,04	1,01	0,99	1,65	1,44		0,89	
529,97		3	0,17	d ₄				0,95		0,89	1,52	1,42	1,34	1,04	
398,736	1591,922	4	-0,2						0,89	0,88	0,77	0,75	1,13	1,14	1,26
531,312		3	0,36									1,32			
333,691	1331,743	4	0,06	d ₀	K155 155[KPAR]158	K51 51[KVTFDR]56	0,85	0,55	0,61	0,70	0,77	0,70	0,92	0,53	
444,586		3	-0,37					1,13	0,69		0,64		0,67		0,41
666,376		2	0,56								0,69				1,05
445,927	1335,768	3	-0,53	d ₄				0,59		0,69		0,78			
668,388		2	0,16				0,64	0,75					0,93		
359,952	1436,786	4	0,06	d ₀	K155 155[KPAR]158	T389 387[DPTPLAR]394	0,87		0,66	0,69	1,46	1,44	0,48		
479,6		3	0,1					0,88				1,32			
360,958	1440,811	4	-0,26	d ₄											0,35
480,942		3	0,05					0,61	0,61		1,28	0,48	0,33		

GT-1b						NS5B		NTP		ssRNA		HCV-796			
m/z gemessen	[M + H] ⁺ berechnet	Ladung	Massenunterschied (ppm)		Peptidsequenz	Peptidsequenz	obere Bande	untere Bande							
720,909		2	-0,02										0,47	0,31	
362,229		4	0,14		K155 155[<u>K</u> PAR]158	K491 491[<u>K</u> LGVPPLR]498				0,56	1,37	1,54	2,91	1,55	
482,636	1445,895	3	-0,81	d ₀							1,79	1,40			1,54
723,451		2	-0,24												
363,235		4	-0,21		S44 44[<u>S</u> AGLR]48	K491 491[<u>K</u> LGVPPLR]498			0,57						
363,235	1449,92	4	-0,1	d ₄				0,67							
483,978		3	-0,11					0,67		0,53					
493,3	1477,885	3	-0,23	d ₀				1,16	0,67			1,50		3,51	
494,642	1481,91	3	0,47	d ₄						0,50					

Tabelle 14: Identifizierte intramolekulare *Crosslinks* mit BS²G und NS5B GT-2a, JFH1 mit verschiedenen Substraten.

Die intrapeptidalen *Crosslinks* sind durch die Sequenzen und die quervernetzten Aminosäuren aufgelistet. Daneben wurde vermerkt, mit welchem d₀/d₄-Verhältnis der *Crosslink* identifiziert wurde.

GT-2a						NS5B		NTP		ssRNA		HCV-796			
m/z gemessen	[M + H] ⁺ berechnet	Ladung	Massenunterschied (ppm)		Peptidsequenz	Peptidsequenz	obere Bande	untere Bande							
394,226		4	-0,16		K154 152[<u>G</u> G <u>K</u> PAR]158	K51 51[<u>K</u> VTFDR]56	2,41	2,08	0,60	1,75	1,41	1,45		2,07	
525,299	1573,881	3	0,2	d ₀				2,18			1,71		0,75		1,73
333,691		4	0,78		K155 155[<u>K</u> PAR]158	K51 51[<u>K</u> VTFDR]56	1,40	1,06	1,50	1,03	1,31	0,99	1,23	1,25	
444,586	1331,743	3	0,94	d ₀				1,37	1,73	1,47	1,07	1,31		1,23	1,03
666,376		2	0,81									1,49	1,11		
445,928	1335,768	3	0,39								1,04				
668,388		2	0,19	d ₄										1,25	

GT-2a							NS5B		NTP		ssRNA		HCV-796			
<i>m/z</i> gemessen	[M + H] ⁺ berechnet	Ladung	Massen- unterschied (ppm)		Peptidsequenz	Peptidsequenz	obere Bande	untere Bande	obere Bande	untere Bande	obere Bande	untere Bande	obere Bande	untere Bande		
510,963	1530,875	3	0,23	d ₀		K51 49[AK <u>K</u> VTFDR]56		1,61		1,21	1,56	1,15		0,90		
804,437	2411,293	3	1,11	d ₀							1,01	0,71				
604,586	2415,318	4	2,12	d ₄	K155 155[<u>K</u> PAR]158	K69 57[TQVLD ⁺ AHYDSV <u>L</u> <u>K</u> DIK]72							0,71			
805,776		3	-2,19												0,50	
415,756	1660,002	4	-0,48	d ₀	K154 152[GG <u>K</u> KPAR]158	K491 491[<u>K</u> LGAPPLR]498	0,70									
416,762	1664,027	4	-0,37	d ₄											0,23	
355,221	1417,864	4	-0,16	d ₀	K155 155[<u>K</u> PAR]158	K491 491[<u>K</u> LGAPPLR]498	0,91	1,16	0,66	0,83	1,25	1,59	1,09	0,94		
473,293		3	-0,05				1,07	1,03	0,68	0,75	1,36	1,62	1,44	0,92		
709,436		2	0,99									1,08				
711,449	1421,889	2	1,27	d ₄				0,53								
407,25	1625,979	4	-1,26	d ₄	K501 499[VW <u>K</u> SR]503	K491 491[<u>K</u> LGAPPLR]498				0,32						
542,664		3	-0,43									0,32				
510,782	2040,107	4	-0,03	d ₀	K501 499[VW <u>K</u> SR]503	K531 524[YLFNWAV <u>K</u> TK]533								1,70		
680,707		3	0,09						2,26	2,16		1,47	0,95		1,44	1,92
511,789	2044,132	4	0,82	d ₄							0,80	0,76	1,53			
682,049		3	1,09									0,73				
420,588	1259,747	3	0,85	d ₀			1,50	1,56	1,22	1,23						
630,378		2	0,74											0,53		
316,698	1263,772	4	-0,01	d ₄	K533 532[TK <u>L</u> K]535	K501 499[VW <u>K</u> SR]503	1,50	1,57	1,17	1,16	0,48	0,51				
421,929		3	0,55										0,46	0,52		
632,39		2	0,83										0,48			

Danksagung

Als erstes möchte ich Frau PD Dr. Iris Thondorf für die Überlassung des interessanten Themas und für die gewährte Freiheit bei dessen Bearbeitungen danken. Die stete Diskussionsbereitschaft sowie die Hinweise und Anregungen beim Anfertigen der Arbeit waren mir eine große Hilfe.

Weiterhin möchte ich Herrn PD Dr. Ralph Golbik für die umfangreiche Unterstützung und Hilfe bedanken, sowie für die Bereitstellung des Proteins. Nur so konnten wir versuchen die Brücke zwischen *in silico* und *in vitro* zuknüpfen.

Bei Frau Prof. Dr. Sinz bedanke ich mich für die Möglichkeit der Kooperation und die geduldigen Diskussionen über die *Crosslink*-Experimente von NS5B, die Bereitstellung der nötigen Chemikalien, Geräte und der Hilfe bei der Auswertung der experimentellen Daten. Ohne dem wäre ein gutes Stück meiner Arbeit nicht möglich gewesen.

Für das angenehme Arbeitsklima und die ausgesprochen große Hilfsbereitschaft aller aktuellen sowie ehemaligen Kollegen der Arbeitsgruppe Molecular Modelling möchte ich mich herzlich bedanken: Sabine Werner, Christiane Ehrt, Richard Bartelt, Anna-Maria Treutler, Sabine Drehkopf (geb. Hänsch), Katharina Jungnickel, Patrice Peterson, Malte Neudorf und Christoph Knorr.

Für die Förderung meiner Arbeit möchte ich dem Landesforschungsschwerpunkt Molekulare Biowissenschaften „Proteine und ihre Funktion in der Kontrolle zellulärer Prozesse“, sowie der Graduiertenförderung des Landes Sachsen-Anhalts danken.

Ein großer Dank gilt Gerd, meiner Familie, meinen Freunden und Kommilitonen, die mich auf diesem Weg begleitet und mich mit Zuversicht, guter Laune und Kaffee versorgt haben.

Lebenslauf

Name, Vorname Sarah Ludwig (geb. Schäfer)
Geburtsdatum 25.04.1986
Geburtsort Staßfurt
Familienstand verheiratet

07.2017 - 31.10.2018 Mitarbeiter Qualitätssicherung bei der Pharma Wernigerode GmbH

Ausbildung

02. - 07.2017 ATV-Seminar „Life Science Management“
11.2011 - 02.2017 wissenschaftlicher Mitarbeiter/Doktorand an der Martin-Luther-Universität Halle-Wittenberg
09.2011 Abschluss des Studiums: 1,7
07.2010 - 08.2011 Diplomarbeit
08.2008 - 02.2009 Auslandssemester an der „Lunds Universitet“ (Schweden), Teilnahme am Masterkurs „Medicinal Chemistry“
08.2007 Vordiplom: 2,4
10.2005 Studium der Biochemie an der Martin-Luther-Universität Halle-Wittenberg
06.2005 Abitur: 1,9
1996 - 2005 Dr. Frank-Gymnasium in Staßfurt

Stipendien

04.2015 - 09.2016 Graduiertenstipendium des Landes Sachsen-Anhalt
03.2012 - 12.2014 Förderung durch den Landesforschungsschwerpunkt Molekulare Biowissenschaften

Tagungen

02. 2013 Vortrag beim „Molecular Modelling Workshop“ in Erlangen
04. 2013 Posterpräsentation bei “Computer Simulation and Theory of Macromolecules” in Hünfeld
03. 2014 Posterpräsentation beim “Treffen des Wissenschaftlichen Beirats des Landesforschungsschwerpunkts Molekulare Biowissenschaften“ in Halle

Halle (Saale), 15. Januar 2019

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe angefertigt und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Die eingereichte schriftliche Fassung der Arbeit entspricht der auf dem elektronischen Speichermedium. Weiterhin versichere ich, dass die vorliegende Arbeit noch nicht als Abschlussarbeit an anderer Stelle eingereicht wurde.

Halle (Saale), 15. Januar 2019

Sarah Ludwig