Biochemical and NMR-Spectroscopic Investigation of the RNA-Unwinding Mechanism of *E. coli* DEAD-Box Helicase RhlB

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich Biochemie, Chemie und Pharmazie
der Johann Wolfgang Goethe-Universität
in Frankfurt am Main
von

Heidi Zetzsche

aus Leipzig

Frankfurt am Main 2023

(D30)

vom Fachbereich Biochemie, Chemie und Pharmazie der

Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Dekan:	Prof. Dr. Clemens Glaubitz
Gutachter:	Dr. Boris Fürtig
	Prof. Dr. Harald Schwalbe

Datum der Disputation:

Für meine Familie

Table of Contents

Summary

The continuous and rapid degradation of messenger RNA (mRNA) is essential for both prokaryotic and eukaryotic organisms as it provides them the ability to quickly adapt to changing environmental conditions and metabolic needs[1]. In bacteria the task of degrading mRNA transcripts immediately after translation is carried out by the multi-enzyme complex degradosome. The degradosome of *E. coli* is composed of the following four essential proteins: endoribonuclease RNase E, 3'-to-5'-exoribonuclease PNPase, ATP-dependent RNA helicase RhlB, and the glycolytic enzyme enolase (Figure 1A)[2–4]. Within this assembly the predominantly unstructured C-terminal domain of RNase E serves as a binding platform for the other three protein partners as well as RNA substrates and anchors the complex to the cell membrane[5–8]. While endonuclease RNase E cleaves the long single-stranded transcript into shorter fragments the PNPase is responsible for the further break down of those fragments into single nucleotides from 3' to 5' end[3,5,9]. However, PNPase can only cleave single-stranded RNA and so stable secondary structure elements such as hairpins within the RNA impede the degradation. Here, helicase RhlB comes into play and unwinds the double stranded RNA section in an ATP dependent reaction to allow PNPase to continue with the exonucleolytic degradation[10].

As a member of the DEAD-Box helicase family, RhlB's domain structure is highly conserved (see Figure 1B, C). Its reaction mechanism is similarly conserved amongst DEAD-Box proteins and follows the following steps: The helicase binds the RNA double strand as well as ATP, transitions into a closed conformation and introduces a strong bent in one of the RNA strands that pries the strands apart, resulting in the dissociation of a single strand. Upon ATP hydrolysis the remaining strand is released together with ADP and P_i and the helicases conformation is reset (Figure 1D)[11–13]. Consequently, DEAD-Box helicases most effectively unwind short RNA duplexes with typically 10-15 bp and are therefore also essential to other cellular processes that require un- or refolding of RNA secondary structures such as ribosome biogenesis, transcription, translation, or pre-mRNA splicing[14–16].

Measurements have shown that, in comparison to other DEAD-Box helicases, RhlB's enzymatic activity – measured in ATP turnover rate and RNA unwinding activity - is poor and barely detectable, prompting the question as to how this protein can have such a significant effect on the efficiency of the degradosome[17,18]. However, as discovered first by Vanzo *et al.* in 1998, RNase E not only binds to RhlB but also strongly stimulates the helicases ATPase and RNA unwinding activity[6,18–20]. The interaction site for RNase E has been shown to be far away from the catalytic site. It is therefore presumed that RNase E stimulates the helicase allosterically[18], but the precise nature of this activation as well as the global or local structural effects of RNase E on the helicase still remain unclear.

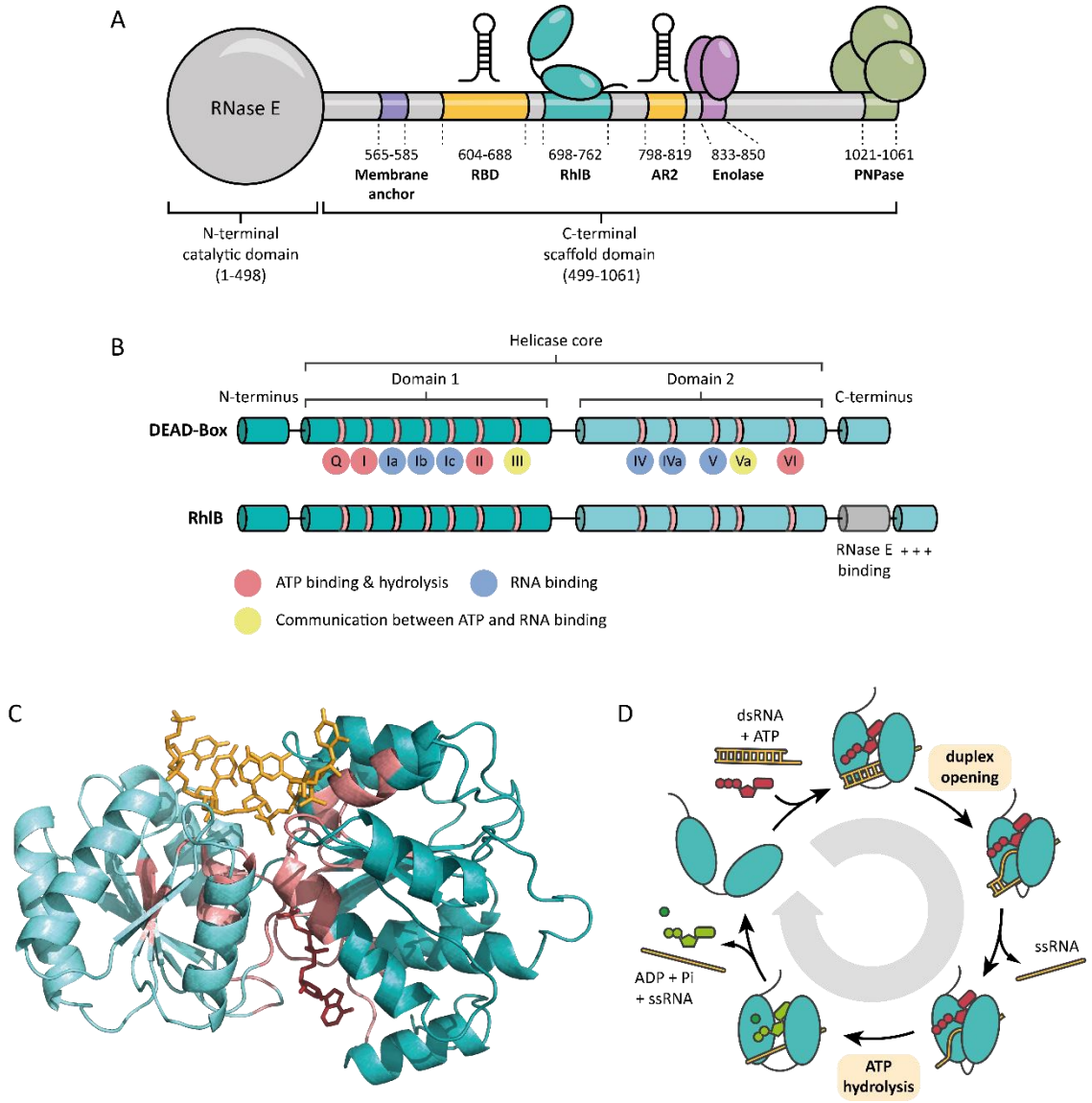

Figure 1. Structure and function of DEAD-Box helicase RhlB. A. Schematic representation of canonical degradosome complex composition including interaction sites between C-terminal scaffold domain of RNase E and complex partner RhlB, Enolase and PNPase. RNA binding sites RBD and AR2 are indicated in orange and membrane anchor in purple. B. Schematic representation of DEAD-Box helicase core domains and of domain architecture of RhlB including ancillary C-terminal RNase E binding site and positively charged C-terminal extension (CTE). Conserved sequence motifs are shown with colour coding corresponding to their primary function (red – ATP binding and hydrolysis, blue – RNA binding, yellow – communication between ATP and RNA binding). C. Homology model of RhlB based on crystal structure of DEAD-Box helicase Vasa[21]. Bound ssRNA (yellow) and AMP-PNP (red) are shown. Model does not include RNase E binding site and CTE of RhlB. D. Schematic pathway of DEAD-Box helicases ATP-dependent RNA unwinding reaction: Substrate free helicase in open conformation. RNA substrate (yellow) and ATP (red) are bound, and helicase adopts closed conformation. Helicase pries open RNA duplex and ssRNA is released. ATP is hydrolysed to ADP and P_i (green). Pi, ADP and remaining ssRNA are released from helicase as helicase is set back to open conformation.

A study by Chandran *et al.* were the first to assess the helicases RNA unwinding activity with specific short RNA duplex substrates that differed in their polarity of their appending single-stranded extension. They made the striking discovery that the extent of RNase E's stimulation varies quite substantially between RNA substrates and that the unwinding rate is significantly higher with 5'-tailed duplexes over 3'-tailed or blunt constructs[18].

This thesis presents new insights into the communication between the two *E. coli* degradosome components RhlB and RNase E by investigating the potential interplay of substrate selectivity and allosteric activation for the DEAD-Box helicase. The utilization of a variety of NMR spectroscopic techniques and oligonucleotide substrates with specific strand properties enables a unique RNA-centred approach to this insufficiently understood protein interaction.

Chapter 4 focuses on the preparation and biochemical as well as NMR spectroscopic characterization of RNA substrate constructs optimized for the subsequent investigation of RhlB's helicase reaction using NMR spectroscopy. Here, a set of short double-stranded RNA constructs that differ in their strand features, fulfil the thermodynamic requirements of a DEAD-Box helicase substrate, and provide an acceptable ^{1}H proton resonances assignment in the NMR spectrum were prepared. Seven individual RNA sequences were synthesized and purified to be combined into six heteroduplexes. The thermal stability and folding behaviour of all constructs were assessed and the ^{1}H imino proton and ^{13}C HSQC assignment obtained for suitable RNA duplexes. Thorough optimizations of measurement and buffer conditions were performed to obtain NMR spectra of the RNAs with sufficient sensitivity and resolution at conditions that also met the constraints of the helicase. Of the six screened heteroduplexes four were identified to be favourable for the subsequent investigations.

Chapter 5 examines the effect of two different RNase E fragments on individual aspects of RhlB's RNA unwinding reaction using the previously prepared RNA substrates with specific single-strand features. After successfully establishing purification protocols for both RhlB and the two investigated RNase E fragments (**chapter 5.1**), the overall structural and dynamic effects of the RhlB/RNase E interaction were characterized in **chapter 5.2**. Using different ^{15}N heteronuclear 2D NMR experiments and isotope-labelled proteins, it could be demonstrated that the RNase E fragment encompassing the RhlB binding site adopts an unstructured conformation reminiscent of an intrinsically disordered protein and that it maintains this conformation when bound to the helicase. It was further observed that RhlB adopts a stable fold with regions of disorder or flexibility, which aligns with the predicted protein structure, where RhlB's C-terminal extension is depicted as unstructured and flexible. The binding of RNase E appears to structurally stabilize the helicase in solution, as well-dispersed NMR spectra of RhlB could be recorded at significantly lower ion concentrations.

In **chapter 5.3** the influence of specific RNA substrates and the binding of two RNase E fragments on RhlB's ATP turnover rate was assessed using a spectrophotometric phosphate assay. The obtained rates clearly show that without its complex partner the helicase RhlB fails to sufficiently hydrolyse ATP independent of the strand features of the used RNA substrate. It could further be demonstrated that with addition of the smallest allosterically stimulating RNase E fragment RhlB's ATP turnover rates increased significantly for all RNA substrates, but most notably for a 5'-tailed duplex and a single strand. Those findings agree with the study from Chandran, who reported a similar

preference for 5'-tailed RNA in RNA unwinding reactions[18]. Upon addition of a larger RNase E fragment also including two adjacent RNA binding sites (RBD) the ATPase activity of the helicase is even further increased. Here, the RNA's strand features play an even larger role in the overall ATP turnover rate of RhlB, as the by far strongest activation is observed by the single strand and the 5'-tailed duplex. Although the single strand does not constitute a physiological substrate, it is the only other construct besides the 5'-tailed duplex that features a single-stranded 5' end. The RNA dependent secondary activation by the larger RNase E fragment also reveals that the binding to the adjacent RNA binding sites not only aids the helicase in its reaction but also appears to favour RNA single strands over blunt end duplexes.

To verify whether RNase E's allosteric activation can also be assessed under NMR spectroscopic conditions, **chapter 5.4** examines RhlB's ATP hydrolysis rate by time-resolved ^{31}P NMR experiments. Here RhlB's unwinding reaction is started by mixing the reaction components directly in the spectrometer and the kinetics of ATP turnover are measurement in real-time from ATP resonances in the ^{31}P spectrum. It could be demonstrated that with the smaller RNase E fragment stimulation of RhlB's ATPase activity is comparable to the spectrophotometric assay. In contrast, the larger, RBD carrying RNase E fragment failed to reach the same level of activation. This can presumably be ascribed to the altered protein and RNA concentrations required for NMR measurements that cause the RNase E fragment to compete with RhlB for bound RNA, thereby reducing the amount of available substrate for the helicase.

Since the results of the ATPase assay clearly revealed an RNA substrate-dependent ATP turnover rate for RhlB, **chapter 5.5** explores whether differences in RNA binding affinity are the origin of this differential activation and whether those preferences are influenced by binding of RNase E. To simultaneously detect any changes in the RNAs base pairing or conformation, ^{1}H NMR titration experiments were conducted. First, RhlB was titrated to the individual RNA substrates, revealing that RhlB has in fact an inherent preference for the 5'-tailed duplex construct over the 3'-tailed and the blunt ended duplexes as it exhibits measurably higher affinity for the first. The analogue measurement series with the RhlB/RNase E complex shows that RNase E's allosteric binding selectively increases the affinity of RhlB for 5'- and 3'-tailed duplexes while decreasing the affinity for blunt end duplexes. These results present the first evidence that RNase E actively affects RhlB's RNA binding. The binding of neither RhlB nor RhlB/RNase E appears to alter the base pairing or RNA conformation of any construct, as only a homogenous peak broadening was observed for the imino proton resonances of the RNA substrates.

However, it was suspected that the alteration in RhlB's RNA affinity must be accompanied with a change in the helicase's interaction with the RNA. To assess this assumption, in **chapter 5.6** the titrations with RhlB and RhlB/RNase E were repeated with isotope-labelled 5'-tailed duplex using ^{13}C HSQC, thereby shifting the focus onto the nucleobase resonances of the RNA. While the binding of RhlB does not appear to alter the RNAs conformation, binding of RhlB in complex with RNase E introduced the formation of a

new RNA conformation. By referencing the ^{13}C HSQC spectra of the 3'-tailed duplex as well as the single strand, the newly appearing resonances could be assigned to an unpaired blunt end of the 5'-tailed construct. These data indicate that in presence of complex partner RNase E, RhlB introduces a partial strand opening of six nucleotides on the blunt end of the RNA substrate, as the remaining base pairs were still intact. The results further provide the first evidence of a DEAD-Box helicase that – upon allosteric activation of its interaction partner – alters the conformation of its substrate RNA even in absence of ATP.

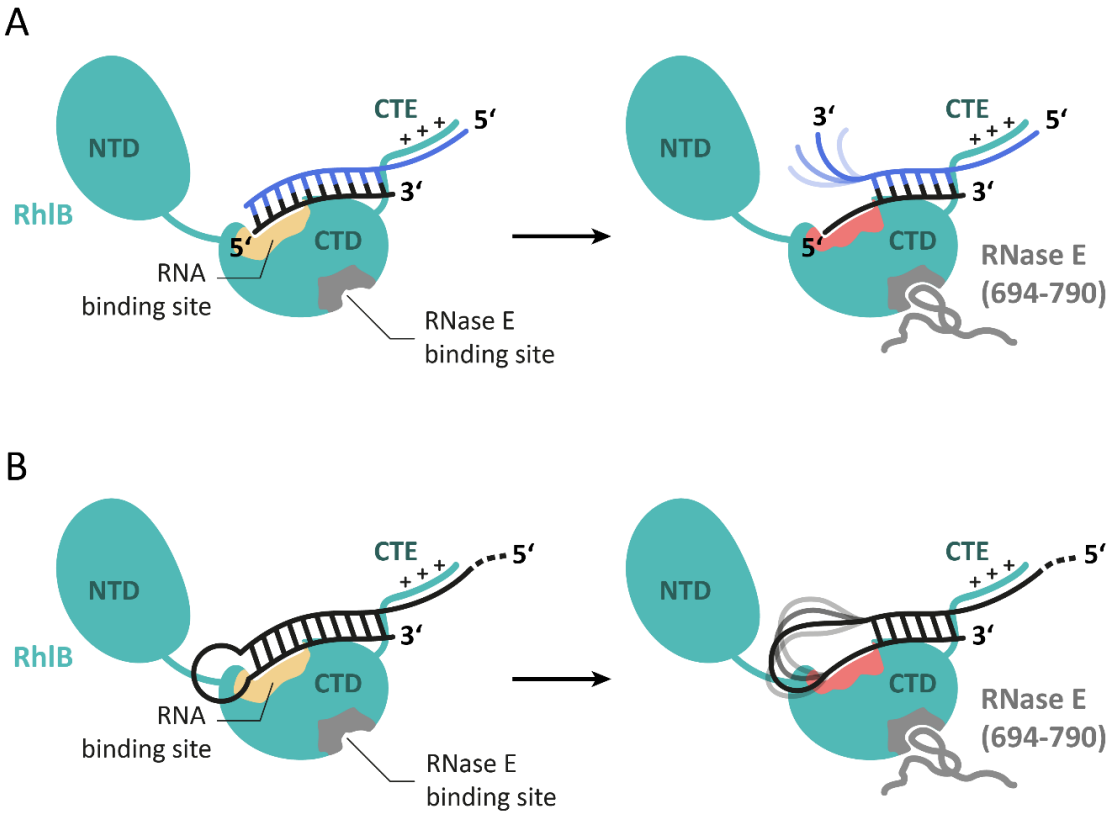

Figure 2. Proposed model of RhlB's interaction with a 5'-tailed substrate RNA in absence and presence of RNase E (694-790). A. Schematic representation of RhlB in open conformation with bound 5'-tailed RNA duplex J2Δ16-M2 in the absence of ATP. Binding sites for RNase E (grey) and RNA substrate (yellow) are highlighted within the C-terminal domain (CTD). The C-terminal extension (CTE) of RhlB interacts with the 5' single stranded extension of the RNA. Upon allosteric binding of RNase E (694-790) the interaction site is altered (red) in a way that the 5' terminal nucleotides of the shorter strand (black) are bound more tightly. This leads to a separation of the terminal base pairs and leaves the 3' end of the 21 Nt strand (blue) dynamic and flexible. B. Schematic representation of the same model in vivo: the bound double strand is part of a stem loop structure in a longer RNA transcript.

To summarize the obtained results, I propose an interaction model that combines RhlB's substrate preferences and RNase E's allosteric effects (See Figure 2). In this model the helicase's binding of the double-stranded section of the substrate in the RNA binding site is presumably aided by interaction of the flexible C-terminal extension with the 5' single strand. A similar behaviour has previously been reported for DEAD-Box helicase Mss116p, which also carries a flexible C-terminal extension[22]. RNA substrates that would lack the single-stranded extension with the right polarity are not coordinated correctly and so

their affinity is lower. However, in absence of the degradosome complex partner, RhlB's ATP turnover and unwinding activity are barely above the detection threshold and so any preferences for specific 5'-tailed RNA substrates are at this point unnoticed and without effect. Upon allosteric binding of RNase E affinity for the 5'-tailed duplex is increased. This is accompanied by a change in the RNA binding pocket that leads to a destabilization and partial opening of the duplex. I therefore propose that the allosteric stimulation of RNase E tightens the grip on the RNA in the binding pocket of RhlB, thereby increasing the affinity and that these altered interaction contacts force the bound part of the duplex into a conformation that destabilizes the base pairing.

Zusammenfassung

Der regelmäßige und rapide Abbau von Boten-RNA (mRNA) ist sowohl für prokaryotische als auch eukaryotische Organismen essenziell, da er den Zellen erlaubt, sich schnell an ändernde Umwelteinflüsse und metabolische Bedürfnisse anzupassen. In Bakterien wird die Aufgabe der Degradation von mRNA unmittelbar nach der Translation von dem Proteinkomplex Degradosom übernommen. Das Degradosom von *E. coli* setzt sich aus folgenden vier essentiellen Proteinen zusammen: Der Endoribonuklease RNase E, der 3'-nach-5'-Exoribonuclease PNPase, der ATP-abhängigen RNA Helikase RhlB und dem glykolytischen Enzyme Enolase (Figure 3A)[2–4]. Die vorwiegend unstrukturierte C-terminale Domäne (CTD) von RNase E dient dabei als Bindungs-Plattform für die übrigen drei Proteine sowie die zu bindende RNA. Darüber hinaus verankert sie den Proteinkomplex mit der inneren Zellmembran[5–8]. Während die Endonuklease RNase E zunächst die langen, einzelsträngigen mRNA-Transkripte in kleine Stücke zerteilt, ist die PNPase dafür zuständig, diese Fragmente vom 3'-Ende ausgehend zu einzelnen Nukleotiden abzubauen[3,5,9]. Jedoch kann die PNPase lediglich einzelsträngige RNA spalten, so dass stabile Sekundärstrukturelemente wie Haarnadelschlaufen im RNA-Strang den Abbau behindern. An dieser Stelle kommt die Helikase RhlB ins Spiel und öffnet in einer ATP-abhängigen Reaktion den doppelsträngigen RNA-Abschnitt, was der PNPase erlaubt die Degradation fortzusetzen[10].

Als Mitglied der Familie der DEAD-Box Helikasen ist das Herzstück von RhlBs Struktur hochkonserviert (Figure 3B, C). Der Reaktionsmechanismus ist innerhalb der DEAD-Box Helikase Familie ebenso unverändert und verläuft nach folgenden Schritten: Die Helikase bindet den RNA-Doppelstrang sowie ein ATP Molekül im Spalt zwischen den beiden Domänen und wechselt von einer offenen in eine geschlossene Konformation. Diese Konformationsänderung verursacht einen Knick im RNA-Duplex, der die beiden Stränge auseinander zwingt und die Dissoziation einer der beiden Einzelstränge zur Folge hat. Nach der anschließenden ATP-Hydrolyse werden der verbleibende RNA-Strang sowie ADP und P_i von der Helikase freigegeben und das Protein nimmt seine Ausgangs-Konformation wieder ein (Figure 3D) [11–13]. Folglich sind DEAD-Box Helikasen nur in der Lage, kurze RNA-Doppelstränge von typischerweise 10-15 bp wirksam zu trennen. Sie sind deswegen an allen zellulären Prozessen beteiligt, die die Um- oder Entfaltung von kurzen Sekundärstrukturelementen erfordern, wie zum Beispiel Ribosomen-Biogenese, Transkription, Translation oder prä-mRNA Spleißen[14–16].

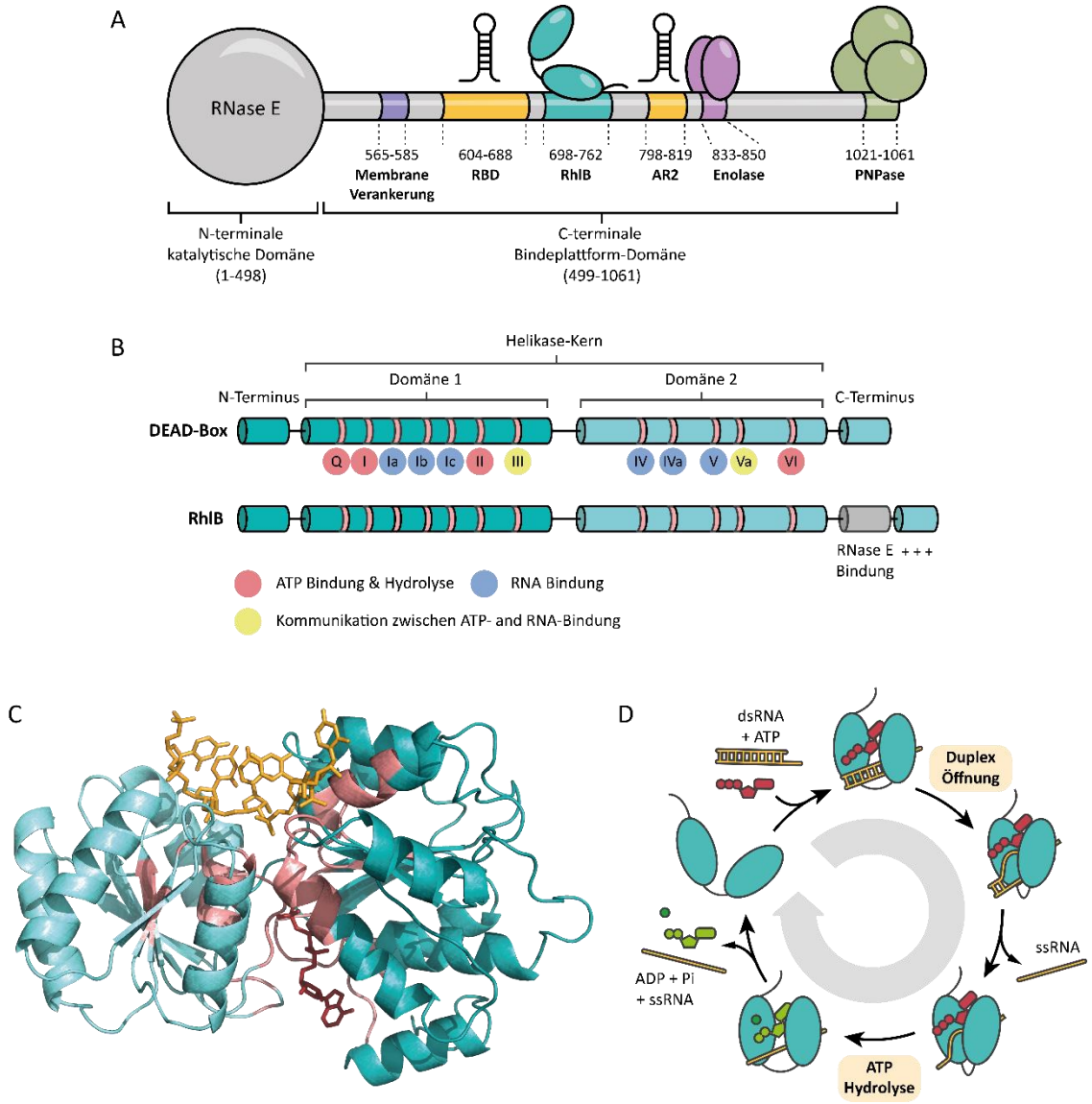

Figure 3. Struktur und Funktion der DEAD-Box Helikase RhlB. A. Schematische Darstellung der kanonischen Zusammensetzung des Degradosom-Komplexes inklusive der Interaktionen zwischen der C-terminalen Domäne von RNase E und den Komplex-Partnern RhlB, Enolase und PNPase. Die RNA-Bindestellen RBD und AR2 sind in orange angedeutet, der Membrananker in lila. B. Schematische Darstellung der Domänen-Anordnung allgemein für DEAD-Box Helikasen und im Speziellen für RhlB, welche eine zusätzliche Bindestelle für RNase E und eine positiv-geladene C-terminale Verlängerung aufweist. Konservierte Sequenzmotive sind hervorgehoben und entsprechend ihrer primären Funktion farblich markiert (rot – ATP-Bindung und -Hydrolyse, blau – RNA-Bindung, gelb – Kommunikation zwischen ATP- und RNA-Bindung). C. Homologie-Modell von RhlB basierend auf der Kristallstruktur der DEAD-Box-Helikase Vasa[21]. Die gebundenen Liganden sind in Gelb (RNA-Einzelstrang) und rot (AMP-PNP) hervorgehoben. Das Modell enthält nicht die RNase E Bindestelle sowie die C-terminale Verlängerung von RhlB. D. Reaktionsmechanismus der ATP-abhängigen RNA-Entwindungsreaktion von DEAD-Box Helikasen: Die freie Helikase befindet sich in einer offenen Konformation. RNA (gelb) und ATP (rot) werden gebunden und die Helikase wechselt in eine geschlossene Konformation. Die Helikase zwingt beide RNA-Stränge auseinander und RNA-Einzelstrang wird aus dem Komplex freigegeben. ATP wird zu ADP und P_i hydrolysiert (grün). P_i, ADP und der verbleibende Einzelstrang werden von der Helikase entlassen und die Helikase kehrt in ihre Ausgangs-Konformation zurück.

Die enzymatische Aktivität einer DEAD-Box Helikase wird üblicherweise anhand seiner ATP-Hydrolyse-Rate oder der RNA-Entwindungs-Rate bestimmt. Untersuchungen haben gezeigt, dass RhlB im Vergleich zu anderen DEAD-Box Helikasen nur eine sehr geringe und kaum messbare enzymatische Aktivität aufweist[17,18]. Dies wirft die Frage auf, wie dieses

Protein dennoch einen derart signifikanten Einfluss auf die Leistung des Degradosoms haben kann. Wie sich erstmals in einer Studie von Vanzo *et al.* herausstellte, bindet RNase E nicht nur an RhlB, sondern diese Bindung hebt auch sowohl die ATP-Hydrolyse-Rate als auch die RNA-Entwindungsrate von RhlB signifikant an[6,18–20]. Da sich diese Bindestelle für RNase E weit entfernt vom katalytischen Zentrum der Helikase befindet, wird davon ausgegangen, dass RNase Es Aktivierung allosterisch erfolgen muss[18]. Die konkrete Art dieser Aktivierung durch RNase E sowie die damit verbundenen lokalen und umfassenden strukturellen Änderungen in RhlB sind allerdings zum Großteil noch ungeklärt.

Chandran *et al.* untersuchte erstmals die RNA-Entwindungs-Rate von RhlB unter Verwendung kurzer RNA-Duplexe, die sich im Speziellen in der Polarität ihrer zusätzlichen Einzelstrang-Verlängerung unterschieden. Sie konnten die verblüffende Entdeckung machen, dass der Grad der Aktivierung durch RNase E bei der Verwendung von RNA-Duplexen mit einzelsträngigem 5‘-Überhang beträchtlich höher war bei vergleichbaren Konstrukten mit 3‘-Polarität oder Konstrukten ohne Einzelstrang[18].

Diese Arbeit liefert neue Erkenntnisse in Bezug auf die Kommunikation zwischen den Degradosom-Komponenten RhlB und RNase E aus *E. coli*, indem das potenzielle Wechselspiel zwischen RhlBs RNA-Selektivität und der allosterischen Aktivierung durch RNase E untersucht wurde. Der vielseitige Einsatz NMR-spektroskopischer Techniken sowie die Verwendung kurzer RNA-Substrate mit spezifischen Strang-Eigenschaften ermöglicht es, mit einen ungewöhnlichen, RNA-zentrierten Ansatz an diese unzureichend verstandene Protein-Interaktion heranzugehen.

Kapitel 4 befasst sich mit der Herstellung und der biochemischen sowie NMR-spektroskopischen Charakterisierung von RNA-Konstrukten, die für die nachfolgenden Untersuchung von RhlBs Reaktionsmechanismus geeignet sind. Hierfür wurde eine Reihe kurzer doppelsträngiger RNA-Konstrukte hergestellt, die sich nicht nur in ihren Einzelstrang-Merkmalen unterscheiden, sondern auch die thermodynamischen Anforderungen eines DEAD-Box Helikase Substrats erfüllen, und gleichzeitig eine ausreichende ^{1}H Protonen-Zuordnung im NMR-Spektrum erlauben. Sieben individuelle RNA-Sequenzen wurden synthetisiert, aufgereinigt und zu sechs Heteroduplexen zusammengefügt. Die thermale Stabilität und das Faltungsverhalten aller Konstrukte wurde bestimmt und die ^{1}H Imino-protonen- sowie ^{13}C HSQC-Zuordnungen für alle vielversprechenden Konstrukte bestimmt. Umfangreiche Optimierungen der Mess- und Pufferbedingungen waren notwendig, um NMR-Spektren mit zufriedenstellender Sensitivität und Auflösung zu erhalten, wobei gleichzeitig auf die Messbedingungen der Helikase Rücksicht genommen werden musste. Von den sechs untersuchten Heteroduplexen wurden vier als geeignet befunden.

In **Kapitel 5** wird der Einfluss zweier RNase E Fragmente auf verschiedene Aspekte von RhlBs Reaktionsmechanismus untersucht, wobei die zuvor hergestellten RNA-Konstrukte mit unterschiedlichen Strang-Eigenschaften verwendet werden. Nach der erfolgreichen Aufreinigung sowohl von RhlB als auch beider RNase E Fragmente (**Kapitel 5.1**), wurden

die allgemeinen strukturellen sowie dynamischen Effekte der RhlB/RNase E Interaktion auf beide Proteine in **Kapitel 5.2** charakterisiert. Mittel verschiedener ^{15}N heteronuklearer 2D NMR-Experimente an Isotopen-markierten Proteinen konnte gezeigt werden, dass das kleinere RNase E Fragment, welches lediglich die RhlB-Bindestelle umfasst, unstrukturiert vorliegt. Diese ungefaltete Struktur, welche typisch für intrinsisch unstrukturierte Proteine ist, wird selbst dann beibehalten, wenn das RNase E Fragment an die Helikase bindet. Darüber hinaus konnte beobachtet werden, dass RhlB eine stabile Faltung annimmt, jedoch auch unstrukturierte oder flexible Regionen aufweist. Diese Beobachtungen stimmen mit der Strukturvorhersage für RhlB überein, in der RhlBs C-terminale Verlängerung vollständig ungefaltet und flexibel ist. Die Bindung von RNase E scheint die Helikase zu stabilisieren, da von dem Komplex wohlaufgelöste Spektren bei erheblich niedrigeren Salzkonzentrationen aufgenommen werden können.

In **Kapitel 5.3** wurde der Einfluss spezifischer RNA-Substrate sowie die Bindung der RNase E Fragmente auf RhlBs ATP-Umsatzrate mit Hilfe eines photometrischen Phosphat-Assays beurteilt. Die gemessenen Raten zeigen deutlich, dass RhlB in Abwesenheit des Komplex-Partners nicht in der Lage ist, signifikante Mengen an ATP umzusetzen, unabhängig davon, welches RNA-Konstrukt eingesetzt wird. Darüber hinaus konnte demonstriert werden, dass die Zugabe des kleineren RNase E Fragments die ATP-Hydrolyse-Rate der Helikase signifikant erhöht, wobei die größte Aktivierung für den RNA-Duplex mit 5‘-Einzelstrang sowie ein einzelsträngiges Substrat zu beobachten ist. Diese Ergebnisse stimmen mit den Chandrans Studien überein, welcher ähnliche Präferenzen für Doppelstränge mit 5‘-Überhang in Entwindungs-Experimenten beobachten konnte[18]. Die Zugabe des größeren RNase E Fragments, welches zusätzlich zwei angrenzende RNA-Bindestellen umfasst, erhöht die ATPase-Aktivität der Helikase noch weiter. Hierbei wird der Einfluss der Einzelstrang-Überhange noch deutlicher, da die mit Abstand stärkste Aktivierung für den RNA-Einzelstrang sowie den Duplex mit 5‘-Überhang zu beobachten ist. Obgleich der Einzelstrang nicht das physiologische Substrat darstellt, ist er das einzige Konstrukt neben dem Duplex mit 5‘-Überhang, welches ein ungepaartes 5‘-Ende aufweist. Die RNA-abhängige sekundäre Aktivierung durch das größere RNase E Fragment macht darüber hinaus deutlich, dass die angrenzenden RNA-Bindestellen in dem Fragment nicht nur RhlB in der in seinem ATP-Umsatz unterstützen, sondern auch die Bindung von RNAs mit Einzelsträngen gegenüber jenen ohne Überhang vorziehen.

Um zu überprüfen, ob die allosterische Aktivierung durch RNase E auch unter NMR-spektroskopischen Bedingungen untersucht werden kann, befasst sich **Kapitel 5.4** mit der Messung von RhlBs ATP-Hydrolyse-Rate mittels Zeit-aufgelöster ^{31}P NMR-Experimente. Hierbei wird die RNA-Entwindungs-Reaktion durch Mischung der Reaktionspartner direkt im NMR-Spektrometer gestartet und die Kinetiken des ATP-Umsatzes in Echtzeit über die ATP-Signale im ^{31}P-Spektrum bestimmt. So konnte gezeigt werden, dass mit dem kleineren RNase E Fragment eine allosterische Aktivierung erfolgt, die mit den zuvor durchgeführten Phosphat-Assay vergleichbar ist. Im Gegensatz dazu erreichte das größere RNase E Fragment, welches die angrenzenden RNA-Bindestellen enthält, nicht die zuvor

erreichten ATP-Umsatz-Raten. Dies kann mit großer Wahrscheinlichkeit der geänderten Protein- und RNA-Konzentrationen zugeschrieben werden, welche für die NMR-spektroskopischen Untersuchungen notwendig sind. Dadurch konkurrieren sowohl RhlB als auch RNase E um die verfügbare RNA, wodurch weniger freie RNA für die Helikase zur Verfügung steht.

Da die Ergebnisse des ATPase Assays deutlich eine RNA-Abhängigkeit bei der ATP-Umsatz-Rate der Helikase zeigen, wird in Kapitel 5.5 untersucht, ob diese Unterschiede ihren Ursprung in den Affinitäten für die verschiedenen RNA-Substrate haben. Zugleich soll ermittelt werden, ob RNase E diese Affinitäten von RhlB beeinflusst. Um im gleichen Zuge zu überprüfen, ob die Bindung der RNA an RhlB die RNA-Konformation oder Basenpaarung ändert, werden ^{1}H NMR-Titrationsexperimente durchgeführt. Zunächst wurde nur RhlB zu den unterschiedlichen RNA-Substraten titriert, wodurch aufgedeckt werden konnte, dass RhlB in der Tat eine inhärente Präferenz für Duplexe mit 5‘-Überhang gegenüber Konstrukten mit 3‘-Überhang oder stumpfen Enden besitzt, was sich in einer erhöhten Affinität zeigt. Die analoge Messreihe mit dem RhlB/RNase E Komplex zeigt, dass RNase Es allosterische Bindung selektiv die Affinität gegenüber Konstrukten mit Einzelstrang-Überhang erhöht, während die Affinität zu RNA Duplexen ohne Überhang sogar verringert wird. Diese Ergebnisse liefern den ersten Nachweis, dass RNase E aktiv Einfluss auf RhlBs RNA-Bindung nimmt. Weder die Bindung der RNA and RhlB noch an den RhlB/RNase E Komplex scheint die Basenpaarung oder Konformation der RNA-Substrate zu beeinflussen, da lediglich eine homogene Peak-Verbreitung aller Imino-Protonen-Signale beobachtet werden konnte.

Nichtsdestotrotz wurde vermutet, dass die Änderung von RhlBs RNA-Affinität mit einer Veränderung der RNA-Bindung selbst einher gehen muss. Um diese Vermutung zu überprüfen, wurden die Titrationen in **Kapitel 5.6** sowohl mit RhlB als auch dem RhlB/RNase E Komplex mit einem Isotopen-markierten Duplex mit 5‘-Überhang wiederholt. Hierbei wurden ^{13}C HSQC-Spektren aufgenommen, welche es erlauben, den Fokus von Imino-Protonen zu den Resonanzen der Nukleobasen zu lenken. Während die Bindung von RhlB auch im ^{13}C HSQC-Spektrum keinen Einfluss auf die RNA-Konformation zu haben scheint, verursacht die Bindung des Protein-Komplexes die Bildung einer neuen RNA-Konformation. Die neu-entstandenen Signale konnten mit den ^{13}C HSQC-Spektren des Einzelstrangs sowie des Duplexes mit 3‘-Überhangs verglichen werden, was es ermöglicht, die neuen Signale einer ungepaarten Konformation des stumpfen Endes zuzuordnen. Diese Daten deuten darauf hin, dass RhlB in Anwesenheit des Komplex-Partners RNase E eine partielle Strang-Öffnung von sechs Nukleotiden Länge hervorruft, während die übrigen Basenpaare weiterhin intakt bleiben. Diese Resultate stellen somit den ersten Nachweis dafür dar, dass eine DEAD-Box Helikase durch allosterische Aktivierung mit einem Interaktionspartner die Konformation ihrer Substrat-RNA in Abwesenheit von ATP verändert.

Um die gewonnenen Resultate zusammenzufassen, schlagen wir ein Interaktions-Modell vor, welches RhlBs RNA-Präferenzen und die allosterische Aktivierung durch RNase E verbindet (Figure 4).

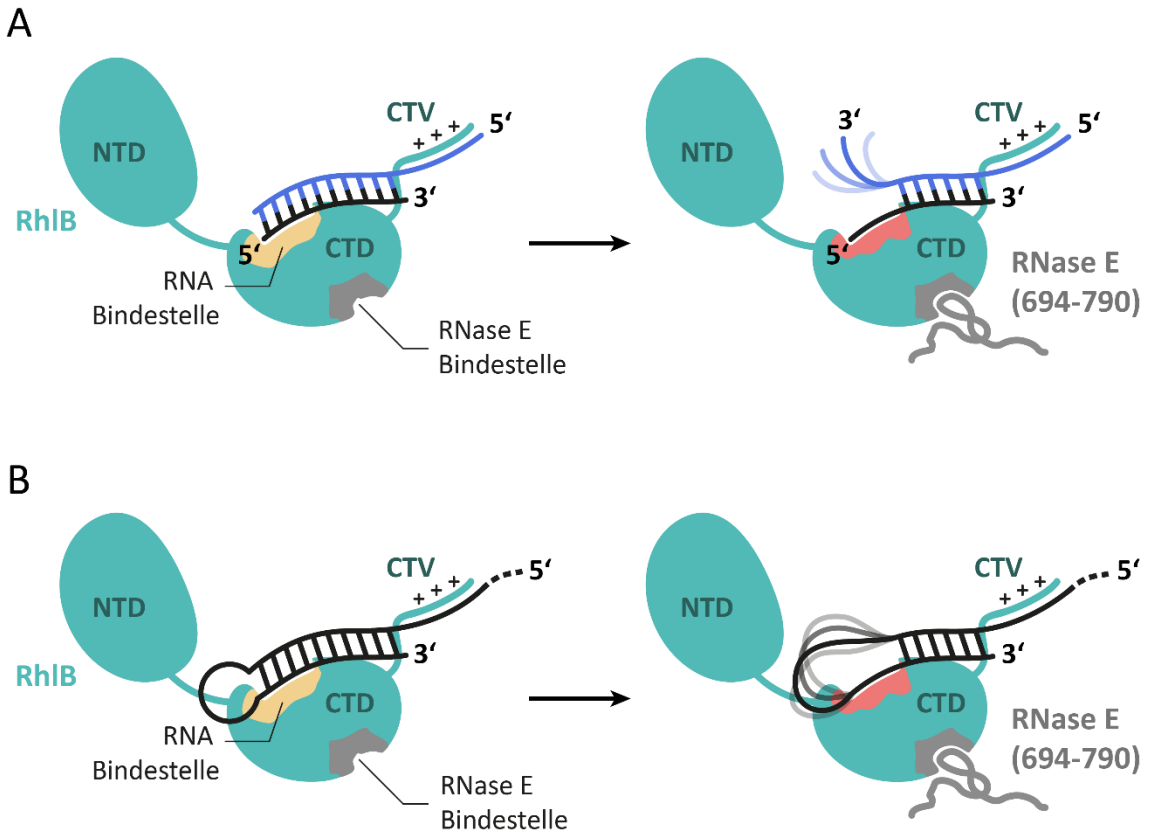

Figure 4. Vorgeschlagenes Modell der Interaktion zwischen RhlB und einem RNA-Duplex mit 5'-Überhang in Anwesenheit und Abwesenheit von RNase E. A. Schematische Darstellung von RhlB in offener Konformation mit gebundenem RNA-Duplex mit 5'-Einzelstrang (J2Δ16-M2) in Abwesenheit von ATP. Die Bindestelle für RNase E (grau) und für das RNA-Substrat (gelb) innerhalb der C-terminalen Domäne (CTD) von RhlB sind entsprechend farblich hervorgehoben. Die C-terminale Verlängerung (CTV) von RhlB interagiert mit dem 5'-Einzelstrang der RNA. Durch die allosterische Bindung von RNase E wird RhlBs RNA-Bindestelle so verändert (rot), dass die 5'-terminalen Nukleotide des kürzen RNA-Stranges (schwarz) fester gebunden werden. Dies führt zu einer Öffnung der terminalen Basenpaare und resultiert in einem einzelsträngigen und flexiblen 3'-Ende des längeren (blauen) Strangs. B. Schematische Darstellung derselben vorgeschlagenen Interaktion in vivo: Das RNA-Substrat ist hier Teil eines längeren mRNA-Stranges und der gebundene Duplex ist hier Teil einer Haarnadelschlaufe.

In diesem Modell wird die Bindung des doppelsträngigen Bereichs der Substrat-RNA in der RNA-Bindestelle der Helikase durch zusätzliche Wechselwirkungen des 5'-Überhangs mit der flexiblen C-terminalen Verlängerung unterstützt. Vergleichbares wurde bereits für die DEAD-Box-Helikase Mss116p gezeigt, welche ebenfalls eine C-terminale Verlängerung aufweist[22]. RNA-Konstrukte, denen der Einzelstrang-Überhang mit der korrekten Polarität fehlt, würden danach nicht korrekt gebunden werden und ihre Bindungs-Affinität wäre so geringer. Dennoch sind RhlBs ATP-Umsatz- und RNA-Entwindungs-Raten in Abwesenheit des Degradosom-Komplexpartners kaum detektierbar und so bleiben etwaige RNA-Präferenzen unbemerkt und ohne Konsequenz. Mit der allosterischen Bindung durch RNase E erhöht sich die Affinität für den Duplex mit 5'-Überhang. Gleichzeitig wird die RNA-Bindestelle der Helikase so verändert, dass der gebundene

Duplex destabilisiert und eine partielle Strang-Öffnung hervorgerufen wird. Ich stelle deswegen die Theorie auf, dass die allosterische Stimulierung durch RNase E die Bindung der RNA in der Bindetasche RhlBs verstärkt, wodurch sich die Affinität erhöht, und dass es diese veränderte Bindung ist, die den gebundenen Teil des Duplex in eine Basenpaar-destabilisierende Konformation drängt.

List of Abbreviations

%	Percentage
1D, 2D	One-Dimensional, Two-Dimensional
Å	Angstrom
A	Alanine
AA	Amino acid
ADP	Adenosine diphosphate
Ala	Alanine
AMP	Adenosine monophosphate
APS	Ammonium persulfate
AR2	Arginine-rich Domain 2
Arg	Arginine
Asp	Aspartate
ATP	Adenosine triphosphate
A.U.	Absorption units
bp	Base pair
BIS-TRIS	Bis-(2-hydroxy-ethyl)-amino-tris(hydroxymethyl)-methane
BEST	Band selective short transient
^{13}C	Carbon-13 (stable carbon isotope)
c	Concentration
°C	Degree Celsius
CD	Circular dichroism
cm	Centimetre
COSY	Correlation spectroscopy
CTD	Carboxy-terminal Domain
CTE	Carboxy-terminal Extension
C-terminal	Carboxy-terminal
CV	Column volume
Δ	Delta
D	Aspartate
d	Optical path length
Da	Dalton
ddH_2O	Deionised and distilled water
DEAE	Diethylaminoethanol
deg	Degree
DEPC	Diethyl pyrocarbonate
D.I.T.	Digital integration time
dmol	Decimole
DMSO	Dimethyl sulfoxide
DNA	Desoxyribonucleic acid
D_2O	Deuterium oxide (deuterated water)
ds	Double strand
dsRNA	Double-stranded RNA
DSS	Sodium trimethylsilylpropansulfonate
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
E	Extinction
E	Glutamate
ε	Extinction coefficient

E. coli	*Escherichia coli*
Et al.	*Et alia ("and others")*
FPLC	Fast protein liquid chromatography
ΔG, ΔG^0	Gibbs free energy, standard Gibbs free energy
g	Gram
g	Centrifugal force
GARP	Globally optimized alternating phase rectangular pulse
Glu	Glutamate
^1H	Hydrogen
H	Histidine
h	Hour
ΔH, ΔH^0	Enthalpy, standard enthalpy
HCl	Hydrogen chloride
HDV	Hepatitis delta virus
HDX-MS	Hydrogen-deuterium exchange analysis coupled with mass spectrometry
HH	Hammer head
His_6-tag	Poly-histidine tag
HPLC	High-performance liquid chromatography
HSQC	Heteronuclear single quantum correlation spectroscopy
Hz	Hertz
I	Intensity
INEPT	Insensitive nuclei enhancement by polarization transfer
IPTG	Isopropyl β-d-1-thiogalactopyranoside
J	Joule
K	Kelvin
K_A	Association constant
K_D	Dissociation constant
kbp	Kilobase pairs
kcal	Kilocalories
KCl	Potassium chloride
kDa	Kilodalton
K_2HPO_4/KH_2PO_4	Potassium monohydrogen phosphate/ Potassium dihydrogenphosphate
L	Litre
LB	Lysogeny broth
LiCl	Lithium chloride
lb	Line broadening
M	Molar Mass
mdeg	Millidegree
MESG	2-amino-6-mercapto-7-methylpurine riboside
mg	Milligram
$MgCl_2$	Magnesium chloride
$Mg(OAc)_2$	Magnesium acetate
MHz	Megahertz
min	Minute
mL	Millilitre
mM	Millimolar
mm	Millimetre
mmol	Millimole
mol	Mole
mRNA	Messenger RNA

ms	Millisecond
MW	Molecular weight
^{15}N	Nitrogen-15 (stable nitrogen isotope)
NaCl	Sodium Chloride
ncRNA	Non-coding RNA
Ni-NTA	Nitrilotriacetic acid
ng	Nanogram
NH_4Cl	Ammonium chloride
nM	Nanomolar
nm	Nanometre
NMR	Nuclear magnetic resonance
NOESY	Nuclear overhauser enhancement and exchange spectroscopy
Nt	Nucleotide
NTD	Amino-terminal Domain
(r)NTP	Nucleoside trisphosphate
N-terminal	Amino-terminal
OD	Optical density
^{31}P	Phosphorus-31
PAGE	Polyacrylamide gel electrophoresis
PCI	Phenol-chloroform-isoamyl alcohol
PCr	Creatine phosphate
PEI	Polyethylenimine
pg	Picogram
pH	Potential of hydrogen/ *potential Hydrogenii*
P_i	Orthophosphate
pI	Isoelectric point
pmol	Picomole
PNP	Purine nucleoside phosphorylase
PNPase	Polynucleotide phosphorylase
Poly(A)	Poly-Adenine
Poly(U)	Poly-Uridine
ppm	Parts per million
Pro	Proline
PSI	Pounds per square inch
R	Ideal gas constant
RBD	RNA binding domain
RecA	Recombinase A
RNA	Ribonucleic acid
RNase	Ribonuclease
rp	*Reversed phase*
rpm	Revolutions per minute
rRNA	Ribosomal RNA
s	Second
S	Svedberg unit
ΔS, ΔS^0	Entropy, standard entropy
SAXS	Small angle X-ray solution scattering
S. Cerevisiae	Saccharomyces cerevisiae
SDS	Sodium dodecyl sulphate
SF	Superfamily
siRNA	Small interfering RNA

sRNA	Small regulatory RNA
ss	Single strand
ssRNA	Single-stranded RNA
SSB	Sine bell shift
T	Temperature
T_m	Melting temperature
Θ	Ellipticity
TB	Terrific broth
TBE	TRIS/Borate/EDTA (-buffer)
TEMED	Tetramethyl ethylenediamine
TEV	Tobacco etch virus
TLC	Thin layer chromatography
TRIS	Tris(hydroxymethyl)aminomethane
tRNA	Transfer RNA
TROSY	Transverse relaxation optimized spectroscopy
U	(Reaction-) unit
UV	Ultraviolet
µg	Microgram
µL	Microlitre
µM	Micromolar
µm	Micrometre
µs	Microsecond
V	Volt
Vis	Visible light
(v/v)	Volume per volume
W	Watt
WaterGATE	Water suppression by gradient tailored excitation
(w/v)	Weight per volume

Chapter 1 Introduction

1.1 RNA Metabolism

1.1.1 Molecular Structure of RNA

Ribonucleic acids (RNA) represent a functionally diverse class of biomacromolecules fundamental to all forms of life, first and foremost because of their involvement in every step of the protein biosynthesis: While messenger RNAs (mRNA) convey the genetic code transcribed from desoxyribonucleic acids (DNA) to the protein synthesizing ribosome, transfer RNAs (tRNA) serve as adaptor molecules to translate the genetic information into amino acids[23,24]. The ribosome itself is composed of both a variety proteins and ribosomal RNA molecules (rRNA)[25]. Beyond that, non-coding RNAs (ncRNA) have become of increasing scientific interest[26], as they, among other things, are responsible for the maturation of pre-mRNA[27] as well as for gene regulation on transcriptional (siRNA, riboswitch)[28,29] and translational level (riboswitch)[30–32].
Structurally, RNA is a single-stranded polymer chain that consists of ribonucleotides, which are composed of one of the four nitrogenous heterocyclic nucleobases adenine (A), cytosine (C), guanine (G) or uracil (U) appended to a cyclic ribofuranosyl sugar. A negatively charged phosphodiester group connects the 3’ position of the ribose with the 5’ position of the next. In a sequence of nucleotides, the termini are therefore conventionally defined as 5’ end and 3’ end for their unbound ribose OH-groups (Figure 5). Hydrogen bonds (H-bond) are typically formed via imino hydrogens between the opposing nucleobases adenine and uracil and between guanine and cytosine[33].

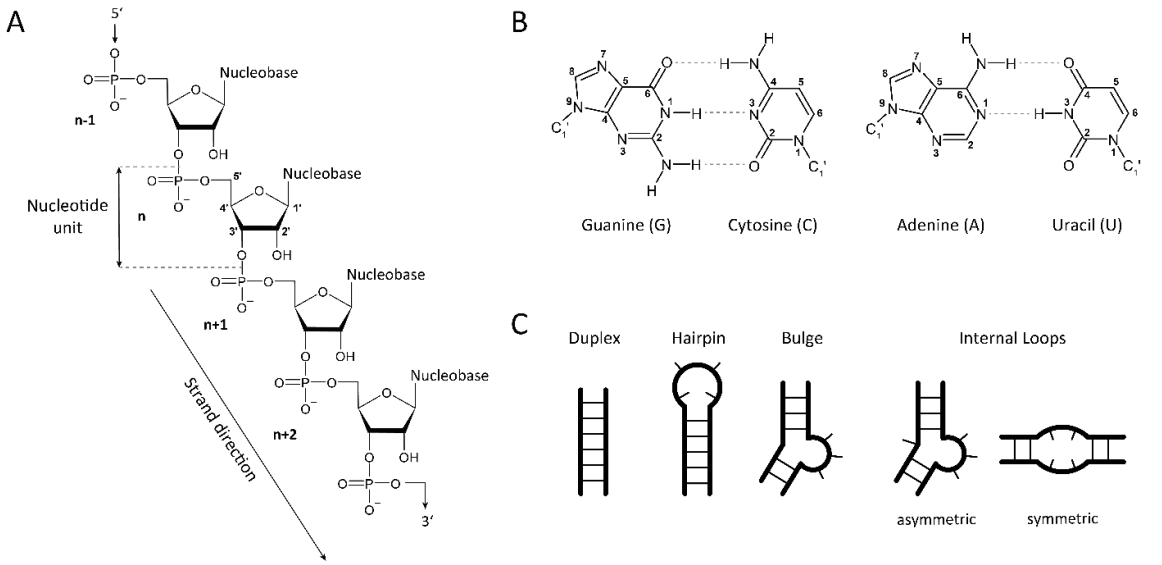

Figure 5. RNA building blocks and secondary structures. A. Chemical structure of RNA polynucleotide strand. Each individual nucleotide is connected via a phosphodiester bond at its 3’-OH group to the 5’-OH group of the following nucleotide. Strand direction is defined from 5’ to 3’ end. B. Chemical structure and Watson-Crick base pairing patterns between the purine (Guanine, Adenine) and pyrimidine (Cytosine, Uracil) nucleobases. Hydrogen bonds are indicated by dashed lines. C. Typical secondary structure elements formed by RNA strands.

In addition to those so called canonical Watson-Crick base pairs between purine and pyrimidine bases, ribonucleotides can also form a variety of non-canonical interactions and base combinations employing the Hoogsteen- and sugar-edge of the nucleobase and utilizing alternative orientations of the glycosidic bond in relation to the hydrogen bond[34]. This variety of H-bond formations together with its single-stranded nature enables the RNA strand to form a multitude of different secondary structures such as helices (A-form helix), hairpins, bulges, internal loops, or junctions and even the structurally very distinct G-quadruplexes [35–39], while desoxyribonucleic acid (DNA) almost exclusively adopts a double-stranded B-form helix conformation[40–43], but also for DNA, exceptions are important including especially DNA G-quadruplexes[44,45]. By combinations of multiple structural elements even more complex tertiary structures can be formed resulting in RNA molecules that can fulfil precise regulatory functions and even catalyse enzymatic reactions[46–48].

But the multitude of possible base-pairings together with the high stability of RNA duplexes can also lead to alternative conformations adopting structures that can be as thermodynamically stable as the native, functional fold[49,50]. Furthermore, the rugged free energy landscape of RNAs often leads to intermediates and long-lived misfolded states being kinetically trapped[51,52]. The energy barriers necessary to overcome these misfolded secondary structures can be so high that the corresponding refolding reaction only occurs on a slow second to minute time scale *in vitro*, as it was observed for ribozymes and other short bistable RNAs[53–55].
To overcome these barriers *in vivo*, RNAs are assisted by a variety of RNA remodelling proteins such as RNA chaperones, helicases, and other binding proteins which un-or refold the RNA and thereby prevent it from being trapped in a non-functional conformation[56–58]. Especially the mode of operation of RNA helicases will be examined in greater detail in chapter 1.3.

1.1.2 RNA Degradation

A large number of cellular mechanisms can only be fully understood with a thorough knowledge of RNA metabolism. Protein production depends strongly on the levels of mRNAs and the translation of mRNA is mediated by tRNAs, rRNA and other functional RNAs. Although it might appear wasteful at first, for a bacterial cell to be able to rapidly adapt to changing environments and respond to metabolic needs, a system of fast and continuous synthesis and degradation of mRNA has evolved[1].The major processes that encompass the life cycle of an mRNA – transcription, translation, and degradation – are each carried out by separate multi enzyme complexes and are closely coupled *in vivo*. As a result of this, the average half-life time of an mRNA lies only between 1 and 10 min[59,60]. rRNA and tRNA on the other hand are usually more stable once they are processed into their mature form. In addition, different factors can modulate the stability of mRNAs: During translation, bound ribosomes provide protection from degrading

enzymes, and after translation higher order structures like stem-loop forming REP sequences within the transcript can further limit the accessibility of sites for nucleases[1,61]. Other mRNA sequences on the other hand directly serve as target regions for ribonucleases or form secondary structures that are bound by RNA unwinding enzymes like helicases, which in turn increase the RNAs accessibility for ribonucleases. Interestingly, polyadenylation, the addition of a poly(A) tail to the 3' end of the mRNA transcript, has a converse effect in eukaryotes compared to bacteria: the posttranscriptional modification strongly increase the RNA stability in eukaryotes by protecting the 3' end from nucleolytic degradation whereas in some degradation pathways in bacteria poly(A) tails provide a "toe-hold" to which exonucleases can bind[1].

The enzymes in the centre of all cellular RNA degradation are ribonucleases or RNases. The different RNases function as a global regulatory network within the cell, not only by controlling the total RNA decay, but also by contributing to recycling of nucleotides, RNA quality control and maturation of rRNA from precursors. They are categorized into endo- and exoribonucleases. Exoribonucleases degrade RNA by removing terminal nucleotides from either the 3' or 5' end of the RNA strand whereas endoribonucleases cleave the phosphodiester bond between two nucleotides within the RNA strand. Beyond that, endo-and exoribonucleases differ mainly in their preferences for specific RNA sequences and whether they can cleave single or double stranded substrates. Although factors like cellular stress conditions, growth phase and whether an mRNA is mono- or polycistronic can have an effect on which degradation pathway the RNA will take, the majority of prokaryotic transcripts starts with an endoribonucleolytic cleavage by RNase E followed by the exonucleolytic degradation by either the polynucleotide phosphorylase (PNPase), RNase R or RNase II[1]. RNases in bacteria can operate alone but they are mostly found as part of the RNA degradosome, a multi protein complex comprised of an endo- and an exonuclease as well as varying subsidiary proteins. RNA degradation in eukaryotes is much more complex and involves more protein factors, primarily due to compartmentalization into nucleus and cytoplasm. In eukaryotes, the exosome is the protein complex that carries out the majority of 3'-to-5' decay. At its core, it consists of a six-membered ring structure in which each subunit is a homologue to the bacterial PNPase[62].

1.2 RNA Degradosome

1.2.1 Functionality of the Degradosome

The essential role of the degradosome in bacterial mRNA decay has been increasingly researched since the 1990s, where the processive degradation of mRNA by the PNPase and the direct interaction of the two ribonucleases RNase E and PNPase could be demonstrated for the first time *in vitro*[3,63,64]. While endonuclease RNase E cleaves the long single-stranded transcript into shorter fragments - preferably at A/U rich sequences within the mRNA substrate - the PNPase is responsible for the further break down of those fragments into single nucleotides from 3' to 5' end in a series of consecutive phosphorolytic cleavage reactions[3,5,9]. Depending on the mRNA transcript, PNPase sometimes additionally requires the RNA substrate to be poly-adenylated at the 3' end in order to bind and start the exonucleolytic attack[10]. It was quickly discovered that an isolated complex formed by the two nucleases failed to degrade small stem-loop structures in the mRNA substrates. Secondary structure elements in the RNA would present a double stranded roadblock to PNPase and impede 3'-to-5' exonucleolytic cleavage[10]. This prompted researchers to expect the involvement of another RNA-unwinding protein component *in vivo* and shortly after the binding of DEAD-Box helicase RhlB to RNase E could be confirmed by copurification experiments[4,64]. In an ATP-dependent unwinding reaction RhlB could melt RNA stem-loops which led to the RNA substrates being fully digested by the nucleases[10]. Around the same time Py *et al.* also identified enolase as the fourth complex component of the degradosome, although it was not clear at the time, what function the glycolytic enzyme played in the RNA degradation machinery[4]. *In vitro* experiments using only RNase E, PNPase and RhlB – a complex referred to as a the minimal degradosome – could demonstrate that those three proteins were sufficient to reconstitute a functional degradation complex, so that enolase was often disregarded as dispensable in context of the degradation machinery[10]. Later it was proposed that enolase can couple the metabolic status of the cell with RNA degradation in *E. coli*, as it is important for the degradation of mRNAs encoding central metabolism proteins[65]. In addition, a more recent study with extensive DNA microarray experiments assessed the effect of deletion of any of the four complex proteins on the global abundance and half-life of mRNA transcripts *in vivo*. By investigating the effects on over 4000 mRNAs they could reveal that all four proteins were essential for normal mRNA turnover and that deletion mutants of any complex component significantly prolonged the half-life time of many mRNA transcripts[66]. At the same time the deletion of PNPase, enolase or RhlB was shown to not preclude bacterial viability[67].

1.2.2 Complex Composition and Localization

In all three domains of life – bacteria, archaea, and eukaryotes – the RNA degradation machinery takes the form of a multicomponent assembly. While the RNA degradosome is best studied in the γ-proteobacterium *Escherichia coli*, degradosome-like complexes have been highly conserved throughout all phylogenetic trees of bacteria[68,69]. The degradosome of *E. coli* comprises the following four principal enzymes: the essential endoribonuclease RNase E[2,64,70], the phosphorolytic 3'-to-5'-exoribonuclease PNPase[3,71], the ATP-dependent RNA helicase RhlB, and the glycolytic enzyme enolase[4].

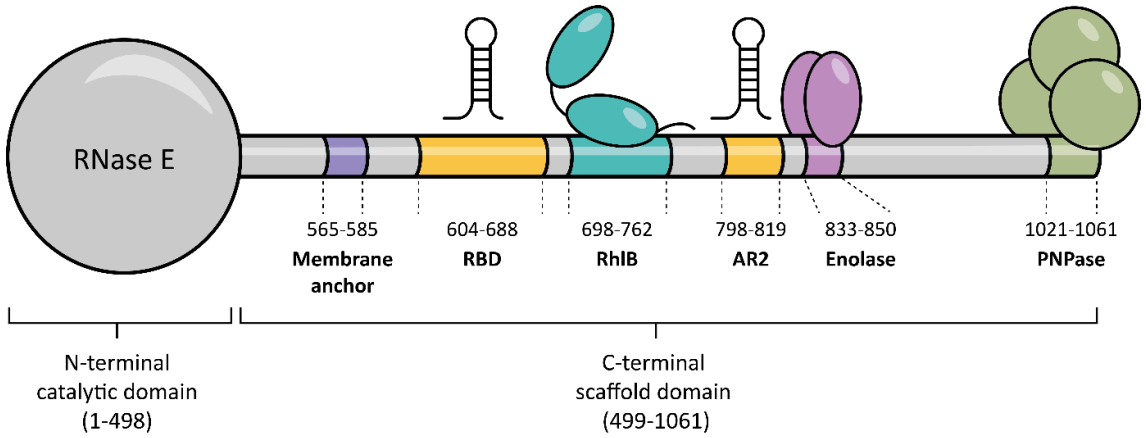

Figure 6. Schematic representation of canonical degradosome complex composition including interaction sites between C-terminal scaffold domain of RNase E and complex partner RhlB, Enolase and PNPase. RNA binding sites RBD and AR2 are indicated in orange and membrane anchor in purple.

As illustrated schematically in Figure 6, the 118 kDa protein RNase E provides the core structure of the complex by serving as a binding platform for the three other degradosome components. While the N-terminal half (residues 1-498) of RNase E contains the catalytic domain that performs endonucleolytic cleavage of RNA substrates, the C-terminal domain (residues 499-1061) is non-catalytic and predominantly unstructured with small regions of higher secondary structures propensity but restricted to specific sites of recognition[5]. Identified first by Carpousis in 1994, the binding site for PNPase is located at the very C-terminus of RNase E (residues 1021-1061)[19]. The interactions with the remaining complex proteins were discovered in the following years and the recognition sites were consecutively narrowed down for enolase (residues 833-850)[72] and RhlB (residues 698-762)[18]. Flanking the binding site for RhlB, RNase E's C-terminal domain also encompasses the two arginine-rich RNA binding domains RBD (residues 604-688) and AR2 (798-819)[6,7]. Under normal growth conditions, biophysical and crystallographic analysis revealed that the complex assembly follows an equimolar stoichiometry, where one PNPase trimer, one enolase dimer and one helicase monomer interact with one RNase E monomer[18,72–74].

It is important to note that the degradosome is a highly dynamic assembly that changes in response to environmental clues through transient recruitment of various additional proteins: During cold-shock conditions in *E. coli*, RNase E has been shown to replace RhlB

by another DEAD-Box helicase called CsdA, that is otherwise involved in the assembly of the ribosome. *In vitro* studies of Khemici *et al.* further revealed that the DEAD-Box helicases RhlE and SrmB can also bind to RNase E and functionally replace RhlB in RNA degradation assays. While SrmB is involved in ribosome biogenesis, RhlE interacts with other DEAD-Box helicases during ribosome maturation to modulate their function under low temperature conditions[75]. However, all three DEAD-box helicases bind at a site that is different from RhlB, showing that RNase E has at least two binding sites for helicases: one that is RhlB-specific, and one additional helicase-unspecific site requiring contacts between residues 791-843[20,76].

Beyond this complex composition required for general mRNA decay, RNase E can also recruit the RNA chaperone Hfq and small regulatory RNAs (sRNA) to assist in a targeted, sRNA-mediated mRNA degradation. For this specific degradation pathway Hfq displaces RhlB, as the binding site for the chaperone (residues 711-750) in large part overlaps with the helicase[77]. Other proteins central to bacterial RNA degradation like the poly(A) polymerase I, RNase II, RNase R or DnaK have also been shown to directly interact with the degradosome, which highlights the interlaced and spatially close relationship of different RNA metabolic pathway components *in vivo*[78–81].

It was furthermore discovered in 2003 via X-ray crystallography, that *E. coli* RNase E forms homo-tetramers through complex interactions in its N-terminal domain. It is suspected that this oligomerization, in which the four subunits are arranged as dimer-of-dimers, promotes the simultaneous cleavage of longer RNA substrates and that it is important for the catalytic activity of the RNase[82,83]. Since the trimeric PNPase bound to RNase E could in principle bind up to three RNase E molecules, even larger complex assemblies of multiple degradosome tetramers have been postulated. This proposal is supported by the considerable heterogeneity of RNA degradosome's size observed in sedimentation experiments and different models of the possible structure of such an assembly have been proposed. However, further research is required to assess the detailed interaction network *in vivo*[83,84].

Even more striking, those tetrameric or potentially even larger assemblies are anchored to the cytoplasmic membrane. This interaction has first been observed in 1991 and was considered controversial until both immunofluorescence and direct fluorescence experiments using RNase E and other degradosome components allowed the detection of the degradation complex as part of the filaments that spiral along the interior surface. The membrane association is enabled by a short amphipathic helix in the C-terminal domain of RNase E (residues 565-585) that functions as a membrane anchor[8,83,85,86]. As this amphipathic helix is predicted to be highly conserved among all γ-proteobacteria, it can be expected that orthologs of RNase E also localize at the bacterial membrane[87]. Considering the close functional and temporal relationship between mRNA translation and degradation, (pointed out in section 1.1.2), it is not surprising that direct interactions between the degradosome and the 70S ribosome and polysomes have also been detected *in vivo*. Native electrophoresis and surface plasmon resonance data indicate that the RNA

binding domains in the C-terminal part of RNase E as well as RhlB are involved in the interaction with the ribosome, but many important aspects of this complex formation remain to be elucidated, including whether the ribosomal proteins or RNAs are bound by the degradosome[88].

In the next chapter we will learn more about DEAD-Box helicases, what distinguishes RhlB from other *E. coli* helicases and why the interaction between RhlB and RNase E is essential to the degradosome's functionality.

1.3 The RNA DEAD-box Helicase RhlB

1.3.1 Classification and Functions of RNA Helicases

When it comes to protein-assisted RNA refolding, two major categories of proteins must be distinguished: RNA chaperones and RNA helicases. While they are both essential in refolding RNAs from a misfolded, kinetically trapped conformation into their native, functional fold in a variety of cellular processes, they differ fundamentally in their structural and mechanistic properties.

First of all, the term "RNA chaperone" has been used for all proteins with a RNA chaperone activity, which means they are able to resolve non-functional RNA structures. From that categorization emerged a very heterogenous group of proteins with no common sequence, motif, or fold. This is also due to the fact that the chaperone activity is often a secondary feature of proteins with very different primary functions in the cell's metabolism. The knowledge about their mode of action is limited, but the interaction with RNA, which is transient and mainly of electrostatic nature, appears to be driven by an entropy transfer and does not require an external energy source such as ATP[56]. The classification of helicases forms a strong contrast to that. A total of 6 helicase superfamilies (SF) have been established so far. Although they perform an extraordinary variety of cellular functions, they share numerous distinct sequence motifs, domain folds and an ATP dependent reaction mechanism (which will be explained in more detail in the chapter 1.3.2). While superfamilies 3-6 contain hexameric enzymes that target predominantly DNA substrates, all RNA helicases - with the exception of a few viral proteins - can be sorted into either superfamily 1 (SF1) or superfamily 2 (SF2) and comprise a single polypeptide chain with two domains[13,89]. As illustrated in a cladogram in Figure 7, those superfamilies can be further subdivided into 13 families, of which the DEAD-Box helicases form the largest family[15].

The essential physiological role of DEAD-Box helicases cannot be underestimated. They can be found in all three domains of life and are involved in virtually all aspects of RNA metabolism, ranging from transcription, translation and RNA decay to ribosome and snRNP biogenesis as well as pre-mRNA splicing[15]. Most of the knowledge about the structure and mechanism of this helicase family has been obtained from research on a

handful of eukaryotic prototype DEAD-Box helicases: the eukaryotic initiation factor 4A (eIF4A)[90–92], the translation regulating helicase Vasa from *Drosophila melanogaster*[21], Ded1p – a yeast helicase involved in translation initiation[93–95], and RNA helicases CYT-19 and Mss116, which are required for mitochondrial intron splicing[22,96–98]. Because DEAD-Box proteins from all phyla show strong evolutionary conservation in their helicase core domains, general principles on structure and function could be inferred for prokaryotic helicases as well[14].

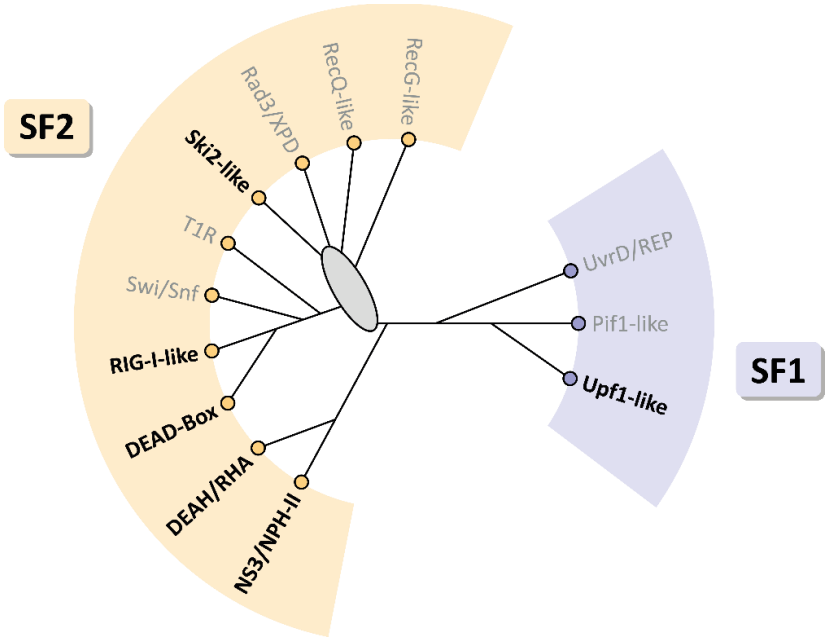

Figure 7. The families of the SF1 and SF2 helicases. Schematic, unrooted cladogram showing the three identified families of the SF1 (right) and the ten families of the SF2 (left). Branch lengths are not to scale. The oval indicates significant uncertainty in the tree topology in this region. Families were named according to names in use, or according to prominent members. Families containing RNA helicases are written in ***bold****, the other families contain DNA helicases. Figure modified after [99,100].*

The well-studied model organism *Escherichia coli* contains the following five DEAD-Box helicases: DbpA, RhlB, RhlE, SrmB and CsdA. SrmB, DbpA and CsdA have all been shown to be involved in the biogenesis and assembly of the large ribosomal subunit, where DbpA is interacting with the ribosomal RNA 23S, while CsdA and SrmB have been shown to be essential for ribosome assembly by modulating L13 (SrmB) as well as S1 and S2 (CsdA)[76,101–103]. The cellular function of RhlE was discovered a few years later: RhlE genetically interacts with SrmB and DbpA to modulate their function during ribosome maturation[75]. Apart from being replaced by CsdA during cold-shock conditions, RhlB is the only DEAD-Box protein in *E. coli* that primarily participates in mRNA processing and decay, making it an unique player in the RNA degradation machinery [76].

1.3.2 Structure and Sequence Motifs of RNA DEAD-Box Helicases

The core region of DEAD-Box helicases displays a high degree of sequence and three-dimensional structure conservation across all phyla. It consists of two recombinase A

(RecA)-like domains that are connected by a flexible linker[12,13,99]. Up to now, 12 conserved sequence motifs have been identified in the helicase core domains. While some of the motifs can also be found in other SF2 families, others are exclusive to DEAD-Box proteins and their presence is required to determine whether a given helicase belongs to the DEAD-Box family[15,99]. Motif II, which contains the amino acid sequence Asp-Glu-Ala-Asp (DEAD), was eponymous for the family name when it was first discovered in 1989[104]. Beyond that, an immense amount of research went into defining and characterizing all sequence motifs. The 12 conserved motifs are located in the cleft between the two domains and have been identified to be either involved in ATP binding and hydrolysis (Q, I, II, VI), RNA binding (I_a, I_b, I_c, IV, IV_a, V) or the communication between ATP and RNA binding (III, V_a). As visualized by colour coding in Figure 8A, motifs required in the formation of the RNA and ATP binding pocket are spaced out between both domains, which indicates that the two domains must close the cleft to functionally bind both substrates[105]. Those findings were first corroborated by Sengoku *et al.* in 2006, who contributed immensely to the understanding of helicases by publishing the first crystal structure of *Drosophila melanogaster* DEAD-Box helicase Vasa in complex with both single-stranded RNA (ssRNA) and an ATP analogue, thereby confirming that a closed interdomain cleft forms the binding pockets for RNA and ATP[15,21]. Figure 8B shows a homology model of RhlB based on the crystal structure of Vasa, visualizing the position of RNA and the nucleotide between the two domains.

Within the ATP binding site, motifs VI and Q are coordinating the nucleobase moiety, while the motifs I, II and VI are responsible for coordinating the three phosphate groups of ATP[21]. This tight interaction and precise positioning within the binding pocket allows the helicase to clearly distinguish between ATP and ADP, which is essential for the coupling of ATP hydrolysis and product release with changes in RNA affinity during the reaction cycle[106]. Crystallization studies on several DEAD-Box proteins including Vasa and Mss116p further revealed that a complex with helicase and RNA substrate is formed even in the presence of non-hydrolysable ground-state or transition-state ATP analogues[21,98].

In the RNA binding site motifs, I_a, I_b, I_c, IV, IV_a and V establish contacts with the RNA strand over a sequence of five adjacent nucleotides[21]. The contacts are formed exclusively with the backbone phosphate groups and the 2'OH group of the sugar moieties of the RNA. Consequently, the RNA binding pocket provides no sequence selectivity for the RNA substrate but can clearly distinguish between RNA and DNA strands on basis of the 2'OH interactions. Nevertheless, the directional orientation of the RNA strand within the binding site is the same for all published crystal structures, with 3' and 5' ends engaged by the N-terminal and C-terminal domains, respectively[14]. Since all published crystal structures of DEAD-Box proteins show the helicase in complex with a RNA single strand in the binding pocket, it remains to be elucidated whether and how DEAD-Box proteins generally distinguish between double-stranded RNA (dsRNA) and ssRNA on a structural level[14]. The bound RNA is characteristically bent into a

conformation that is a key feature exclusive to the RNA unwinding mechanism of DEAD-Box helicases, as will be explained in depth in section 1.3.3.

The communication between RNA and ATP binding pocket, too, is only partially understood, but the two motifs III and V_a have been identified to be important. Motif III does not directly coordinate ATP but instead interacts with motifs II (NTD) and VI (CTD). Because its deletion has been shown to inhibit ATP hydrolysis and reduces the affinity for single-stranded RNA, it is considered to be responsible for interdomain communication[107]. As has been reported for Mss116p, motif V_a appears to contribute to initial binding of the duplex RNA and rearranges to help form the ATPase active site in the closed state[22].

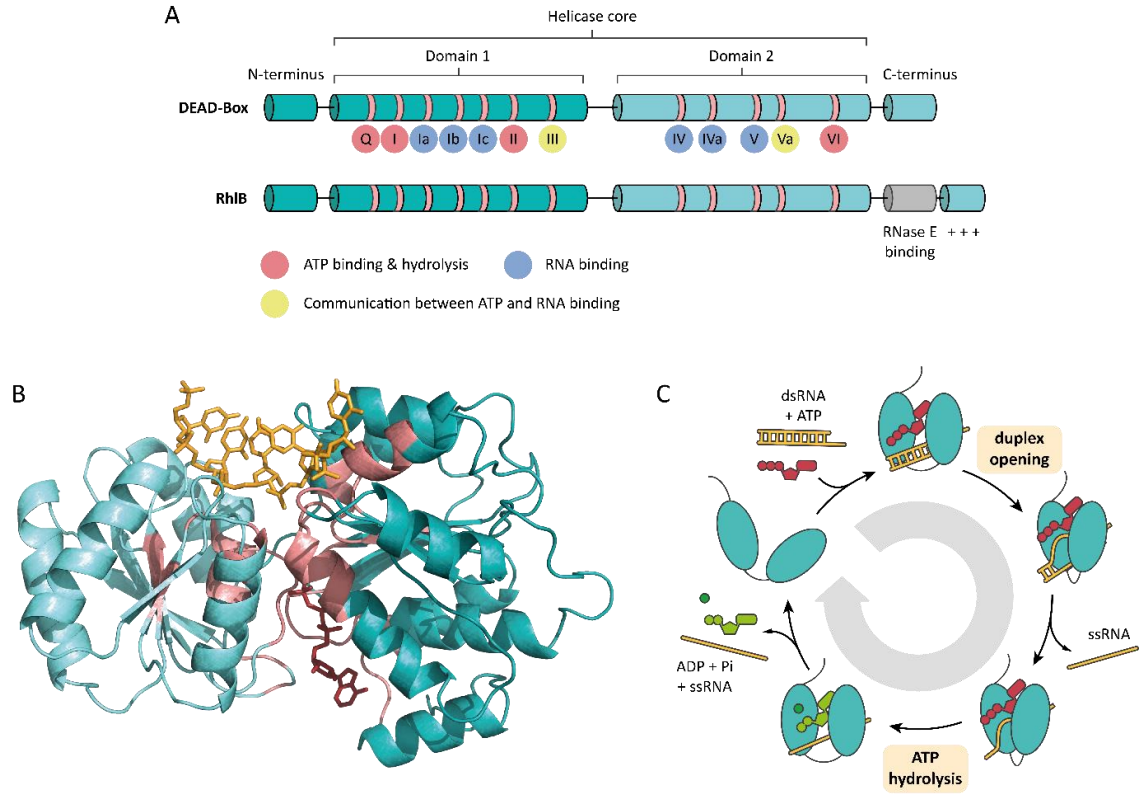

Figure 8. Structure of the DEAD-Box helicase core and the unwinding mechanism. A. Schematic representation of DEAD-Box helicase core domains and of domain architecture of RhlB including ancillary C-terminal RNase E binding site and positively charged C-terminal extension (CTE). Conserved sequence motifs are shown with colour coding corresponding to their primary function (red – ATP binding and hydrolysis, blue – RNA binding, yellow – communication between ATP and RNA binding). B. Homology model of RhlB based on crystal structure of DEAD-Box helicase Vasa[21]. Bound ssRNA (yellow) and AMP-PNP (red) are shown. Model does not include RNase E binding site and CTE of RhlB. C. Schematic pathway of DEAD-Box helicases ATP-dependent RNA unwinding reaction: Substrate free helicase in open conformation. RNA substrate (yellow) and ATP (red) are bound, and helicase adopts closed conformation. Helicase pries open RNA duplex and ssRNA is released. ATP is hydrolysed to ADP and P_i (green). Pi, ADP and remaining ssRNA are released from helicase as helicase is set back to open conformation.

Flanking the core domains, DEAD-Box helicases often contain additional carboxy-terminal or amino-terminal domains that can range from a few to several hundred amino acids in length[99]. In fact, these ancillary domains enable the helicase to target specific RNAs or interact with particular protein partners required for their individual cellular function. For instance, the C-terminal domain of Mss116p extends the RNA binding surface by

coordinating additional nucleotides in the counter strand of the bound dsRNA, whereas *E. coli* helicase DbpA and its *B. subtilis* ortholog YxiN contain a second RNA binding site that lets them bind specifically to hairpin 92 of the 23S ribosomal RNA[98,108,109].

RhlB has a molecular weight of 47.13 kDa and 421 AA, where amino acids 1-357 encompass the DEAD-box helicase core domain. As depicted in Figure 9, the C-terminal domain additionally comprises a binding site for RNase E (residues 368-397) as well as an arginine-rich C-terminal extension (CTE) (residues 397-421)[18]. The C-terminal tail is flexible and unstructured and - similar to CTE of other DEAD-Box helicases - is proposed to assist in RNA binding of RNA, as truncated RhlB derivatives lacking the tail showed a reduced affinity for RNA compared to the full length protein[18]. The further significance of the interaction with RNase E will be explained in greater detail in chapter 1.3.4.

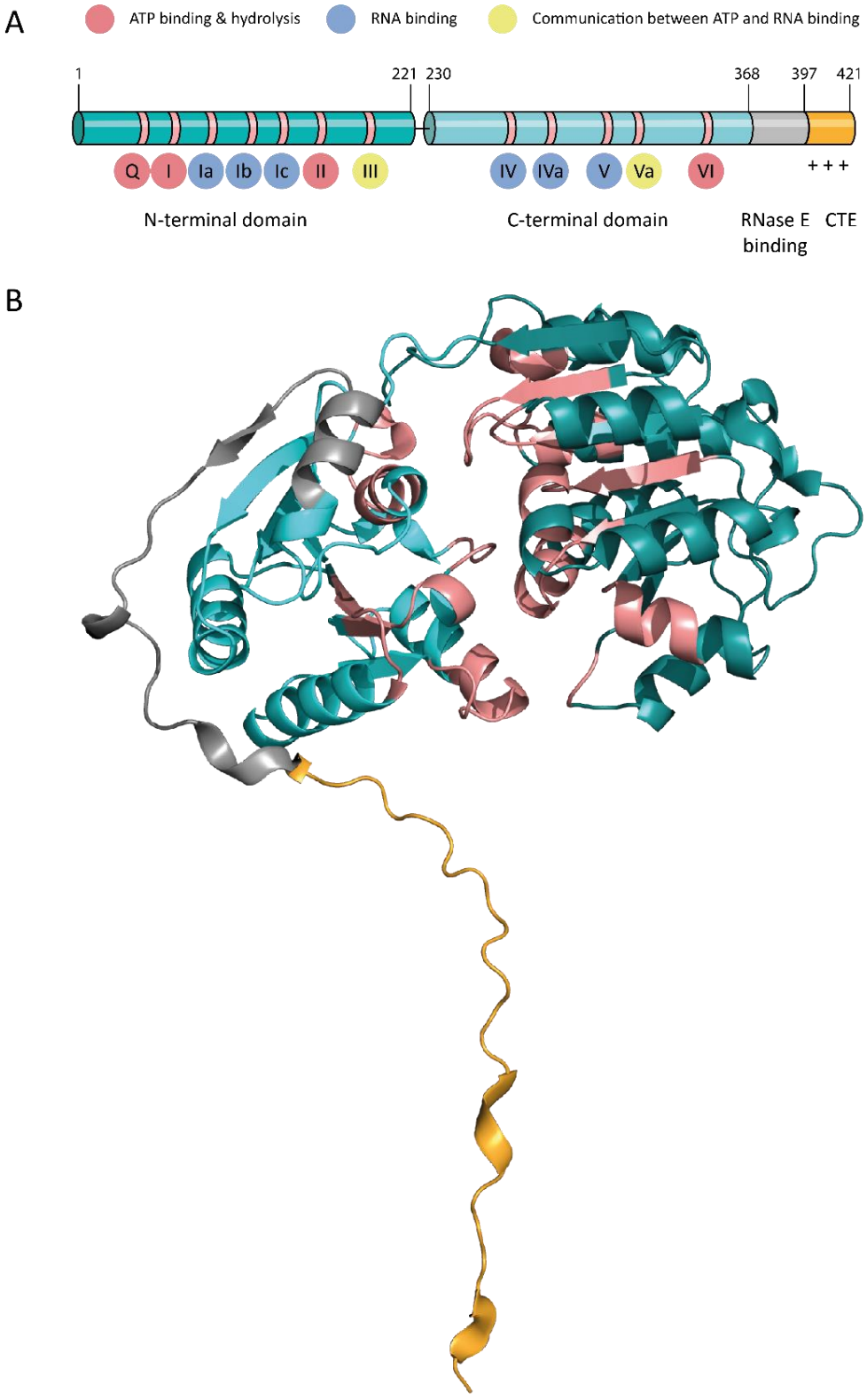

Figure 9. Structure of the DEAD-Box helicase RhlB. A. Schematic representation of RhlB's domain architecture including ancillary C-terminal RNase E binding site (grey) and positively charged C-terminal extension (CTE; orange). Conserved sequence motifs are shown with colour coding corresponding to their primary function (red – ATP binding and hydrolysis, blue – RNA binding, yellow – communication between ATP and RNA binding. B. AlphaFold structure prediction of RhlB based on UniProt entry P0A8J8; AlphaFold DB version 2022-11-01, downloaded on 09. May 2023. Colour coding matches domain architecture in A.[110,111]. For model confidence per residue please see Figure 62.

1.3.3 RNA Unwinding Mechanism

The translocating unwinding mechanism of helicases, especially the ones involved in DNA replication and transcription, has been studied extensively, and so it was a significant revelation when it was first discovered that the family of DEAD-Box helicases exhibits a reaction mechanism completely distinct from other helicase families. A schematic pathway of the reaction mechanism is depicted in Figure 8C.

In absence of any substrate, DEAD-Box helicases populate an ensemble of "open" conformations in which the two core domains are separate and have some independent mobility[112–115]. Upon binding of both RNA and ATP the helicase adopts a "closed" conformation, where both domains are packed together to form the binding pockets for the two substrates in the cleft between them. ATP and RNA are bound cooperatively, as was demonstrated for several DEAD-Box helicases such as mammalian eIF4A and *E. coli* DbpA [90,106,114,116–118]. With both domains of the helicase forming contacts with the ligands, the domain closure is most likely to be the origin of this cooperativity[21,96]. It is this closed conformation that introduces a strong bent to the RNA backbone which is incompatible with A-form helix conformation and therefore pries the RNA duplex apart. This unwinding mode, which has been termed "local strand separation", introduces a duplex opening that for short duplexes not exceeding two helical turns is sufficient to cause the destabilization and consequently dissociation of the counter strand as a whole[14,15,21,98,105,119]. With the remaining RNA single strand still tightly gripped the helicase hydrolyses ATP to ADP and P_i. P_i is immediately released so that the ATP hydrolysis is essentially irreversible. The helicases affinity for ADP is significantly weaker than for ATP - in some cases even anti-cooperativity with ligand RNA was demonstrated – and so both the remaining single strand as well as ADP are rapidly released from the protein and the helicases conformation is set back to its initial open state[90,96,120]. Therefore, despite initial assumptions, the hydrolysis of ATP is not required for RNA unwinding but rather for helicase recycling, as could also be demonstrated by Liu *et al.* by effective duplex unwinding using non-hydrolysable ATP analogues[121,122].

This RNA unwinding mechanism of DEAD-Box helicases is unique and stands in strong contrast to the processive mechanisms adopted by helicases from other families. The processive mechanism involves helicases to travers along a substrate RNA while unzipping it. The DEAD-Box protein's mechanism instead allows them to directly load onto the duplex and locally destabilize it. This of course has substantial implications on the range of substrate RNAs that can be effectively unwound. First and foremost, the duplex length: Most DEAD-Box proteins have been shown to unwind short RNA helices no longer than 10-15 bp, which corresponds to a free energy (ΔG) of approximately 15-25 kcal/mol[16,17,91,92,95,102,123]. However, the overall energy barrier set by the duplex stability is what is setting the limits more than the strand length, therefore effective unwinding of duplexes with up to 21 bp could be demonstrated for RNA substrates with a low GC content of 14.3 % (ΔG 23 kcal/mol) for *E. coli* DEAD-Box helicase CsdA[124].

Similarly, the experiments by Roger *et al.* could confirm that two duplexes of different lengths but of equal stabilities would be unwound by eIF4A with similar efficiencies[92]. The exact values for unwinding limits of course also depend on experimental conditions such as temperature and reaction solution.

Studies that measured ATP hydrolysis and RNA unwinding rates in parallel demonstrated furthermore that unwinding of a short 10-15 bp RNA duplex requires a single cycle of ATP binding and hydrolysis, showing that the stoichiometry between ATP and unwound duplex is equimolar. For longer or more stable RNAs, more ATP was consumed per RNA molecule, indicating that ATP hydrolysis does not lead to complete unwinding in every reaction cycle and futile ATPase cycles might occur[16,93,94,125]. Even in absence of RNA, ATP turnover can take place but only with significantly reduced activity as only cooperative binding of both substrates yields in the optimal reaction cycle[17].

In the past, many studies investigating the mechanism of DEAD-Box helicases utilized a crude RNA mix extracted from yeast ("bulk RNA") as substrate for their ATPase assays, and so it took until the early 2000's before researchers used specific RNA substrates and noticed that individual DEAD-Box proteins had preferences or restraints regarding substrate RNAs with 5'- or 3'-single strand extensions. Because those preferences in length and directionality of single-stranded regions are linked to the specific biological function of the helicase within the cell, they can differ significantly and must be assessed on a case-by-case basis. The first helicase, for which this was analysed in detail, was eIF4A: Unwinding experiments with duplexes containing either 5'- or 3'-single-stranded regions ranging from 0 to 25 Nt revealed that eIF4A was more active in unwinding 5' tailed RNAs at shorter (0-8 nt) single-stranded region length and that the activity lowers when the single-stranded region is shortened below 8 nt. Surprisingly, eIF4A was also capable of unwinding blunt-ended RNA substrates with 70% activity[92]. Similar unwinding capabilities were identified for DEAD-Box helicase RhlE from *E. coli* centred studies, which was able to unwind various 14-mer duplexes independent of length and position of single-stranded extensions present (or even blunt end). In contrast, the unwinding activities of CsdA and SrmB were far more elusive. Both helicases were only able to unwind duplexes with extensions of 12 Nt or more – no matter the direction of the single strand - and no unwinding was observed for shorter single strands or even blunt ended RNA substrates[16]. The ATPase activity, too, was probed for all three *E. coli* helicases. While most of the differential responses to RNA substrates matched the observation for RNA unwinding, some RNA substrates previously resistant to unwinding in SrmB showed a significant stimulation of ATPase activity. This again highlights that successful ATP turnover does not have to concur with effective RNA unwinding. To provide an explanation for the differential RNA substrate preferences and restrictions between DEAD-Box helicases, it was proposed that the attention should be turned beyond the conserved helicase core to the N-or C-terminal extensions. The positively charged C-terminal extensions present in SrmB, CsdA and RhlE, could provide another RNA binding

site, which would strengthen the RNA interaction for substrates with a single-stranded extension[16].

While all other *E. coli* DEAD-Box helicases have been studied quite in-depth in this regard, the same assessment of RNA substrate preferences are scarce for RhlB. When assessing RhlBs ATPase activity like for many other helicases in the past, researches used bulk RNA from *S. cerevisiae*, which unfortunately does not provide any insight into RNA substrate restrictions as it is a mix of different RNAs with undefined conformations[17]. The previously mentioned studies assessing RhlBs function within the degradosome and its ability to facilitate the RNA degradation used the *malEF* REP RNA, a 375 Nt long intercistronic region of a particular transcript known to contain stem-loop structures stable to 3'-exnoculeolytic digests. Because it could not be assessed from the electrophoretic assays where RhlB binds to the RNA, these studies, too, only slightly narrow down the helicases RNA constraints if they exist[4,10]. By far the largest contribution to this topic came from the research group of Luisi in 2007: They examined RhlB's unwinding activity with a 12mer RNA substrate with 12 Nt 5' extension in direct comparison to the same substrate with 3' extension and a blunt end 12mer and could unambiguously show a higher unwinding rate for the 5' substrate (60%) compared to the 3' substrate (15%). The blunt-ended duplex was not unwound at all. However, those results were obtained for experiments in presence of RNase E, which has been shown to stimulate RhlB's ATPase and unwinding activity. Under the experimental conditions with RhlB alone no unwinding could be detected for either RNA. The activating effect of RNase on RhlB and the proteins interaction will be elucidated in further detail in the following section, but with respect to the substrate specificity of RhlB it is so far not clear whether the 5' single strand preference is induced by the RNase interaction or is also present in RhlB alone. Chandran *et al.* further tested the individual domains of RhlB for RNA binding in absence of ATP and revealed that while RhlB-CTD forms a complex with RNA, RhlB-NTD did not show measurable binding to RNA. This suggests that RhlB's C-terminal domain is the primary site for RNA interaction in the absence of ATP[18].

1.3.4 RhlB and its Interaction with RNase E

The enzymatic activity of a DEAD-Box helicase is usually measured either in terms of ATP turnover rate or unwinding activity. The ATP turnover rate can be measured by a variety of enzymatic assays[16,17,106,126], whereas RNA unwinding can for example be quantified by the amount of RNA duplex to monomer conversion of radiolabelled RNA in native PAGE[18]. Based on these metrics RhlB appears to be a remarkably poor helicase. In site-by-site comparison with *E. coli* DEAD-Box helicases RhlE and SrmB the two helicases reached ATPase activities of approximately 32 and 34 mol P_i/min/mol protein, respectively, while RhlB displayed an activity on the borderline for the sensitivity of the assay (≤ 0.5 mol P_i/min/mol protein)[17]. Similarly poor results were obtained by Chandran *et al.* regarding RhlB's unwinding activity tested in an electrophoretic assay with either 3'-

or 5'-tailed 12mer RNA in the presence of ATP: Under the investigated conditions barely 10% of the substrate RNA were unwound after 30 min, which was only a minute increase from the negative control experiment omitting ATP (0-2%)[18].

Those experimental results pose the question, how this barely active enzyme can have such a significant effect on the efficiency of the degradosome, where RhlB is essential to facilitate the full degradation of otherwise stable mRNA transcripts[10]. In 1998, Vanzo *et al.* were the first to detect that the C-terminal domain of RNase E not only binds to RhlB but also stimulates the helicases ATPase activity by a factor of at least 15[6]. This was later corroborated by Callaghan, who even measured a 25-fold ATP turnover increase for the helicase in complex with RNase E[19]. The variation in activation factors can be assumed to result from the differences in experimental conditions and assay sensitivity. As expected, RhlB's RNA unwinding activity was also strongly influenced by the presence of RNase E as demonstrated by the increased unwinding activity from 8% to 60% when bound to RNase E[18].

It was of course of great interest to researchers to gain more insight into this stimulating protein interaction and understand how RNase E causes this activity increase in RhlB. First of all, RhlB has been demonstrated to be a monomer in isolation as well as in complex with RNase E and tightly binds to RNase E in a 1:1 ratio with a K_D of about 50 nM[19,20]. Using limited proteolysis, size exclusion chromatography and crosslinking experiments, the binding site within RhlB could be narrowed down to residues 368-397 in the C-terminal domain, which are not part of the conserved helicase signature core (see Figure 9). Despite the proximity, the flanking flexible C-terminal tail of RhlB (residues 397-421) was shown to be not required for the protein-protein recognition, as a RhlB construct lacking the tail was still able to bind the RNase[18]. Based on protein structure predictions (AlphaFold) and homology models of RhlB with the crystal structure of DEAD-Box helicase Vasa the binding site for RNase E is approximately 20 Å away from the ATP binding pocket and RNA binding pocket (see Figure 9). Since this is a comparatively large distance, it was inferred that RNase E must have an indirect effect on the helicases ATPase activity and that an allosteric activation was taking place[17].

Until now, relatively little is known about the global or local structural effects of RNase E on RhlB, or which conformational changes occur in the helicase during the allosteric activation. Based on the increased ATPase activity, a structural change in the ATP binding pocket was assumed to be most likely. In 2018, in hydrogen-deuterium exchange (HDX-MS) experiments, a reduced surface exposure in presence of RNase E was detected for several regions within RhlB, all of which were located in the C-terminal domain. Besides the postulated RNase E binding site, those regions encompassed motif IV (RNA binding), parts of motif V_a (communication between ATP and RNA binding) and a stretch in close proximity to motif VI (ATP binding and hydrolysis)[127]. Those results revealed that not only the ATP binding site but also the interaction with RNA might be affected by this allosteric activation.

An exciting detail in the relationship between the two proteins that also needs to be touched on is RhlB's unique sequence difference in one of the 12 conserved DEAD-Box protein motifs: the highly conserved aspartate in motif V_a at position 320, which is shared by all its closest homologues and all other *E. coli* DEAD-Box proteins, is replaced by a histidine in RhlB. Notably, this D320H mutation in motif V_a, appears to be co-conserved with some residues in the predicted RNase E binding site, suggesting a strong evolutionary link between RNase E binding and ATPase activation[18]. Most strikingly, a back-mutation of H320 back to aspartate has been shown to completely remove any ATPase activity in the helicase and even abolish the stimulating effect of RNase E that was observed for the wild-type[17]. Unfortunately, analogue tests for RhlB's unwinding activity have not been conducted, yet.

The unwinding assays from Chandran also revealed, that RNase E activation is not the same for different RNA substrates: The unwinding activity of a 5'-tailed 12mer RNA was boosted from 8% to 60% in presence of RNase E, whereas the same 3'-tailed substrate only increased from 12% to 15%[18]. Not only does this indicate that RNase E's stimulating effect is substrate dependent, but it could also hint at RNase E altering RhlB's innate substrate preferences, which could certainly involve structural changes of the RNA binding site.

When assessing the extent of the stimulation of RhlB by the RNase in previously published studies, another important aspect must be taken into consideration: which fragment of RNase E was used. While the full-length endoribonuclease has been be expressed in bacteria in the past, having a fully functional RNA-degrading enzyme in the reaction mix prohibits any experiments on the helicase in presence of RNA substrate. So as soon as the recognition site for RhlB was first identified by Vanzo *et al.* to be located within AA 628-843 in the C-terminal part of RNase E, the complete catalytic N-terminal half as well as most of the flanking C-terminal sections of the RNase could be removed for further experiments where RNA degradation was undesirable[6]. More research groups continued to use the established fragment RNase E (628-843) because of it's convenient cloning sites and the comparability it provided[18–20]. As shown in Figure 10, RNase E (628-843) also encompasses half of the RNA binding site RBD as well as the complete RNA binding site AR2 and part of the binding site for enolase. Experiments with the smaller RNase E fragments narrowed down the binding site for RhlB even further to fragment RNase E (694-790) and demonstrated that the flanking RBDs are not required for direct RhlB recognition[20]. With limited proteolysis Chandran *et al.* could set even smaller limits to the interaction site, stating that the minimal region for RhlB recognition by RNase E must be located between residues 698-762 of RNase E[18].

As discussed above, the C-terminal domain of RNase E has a low structural complexity. The under-representation of hydrophobic amino acids, the enrichment of polar and charged residues clusters and the proline-rich stretches together with the scarcity of predicted secondary structures provided strong evidence and reason for an unstructured protein domain. This was corroborated by small angle X-ray solution scattering (SAXS) and

CD-spectroscopic data[19,83]. The proline-rich clusters are assumed to contribute to a more rigid, extended conformation that is inferred from the SAXS data. However, this unstructured C-terminal domain that has little apparent conservation is interspersed with short highly conserved segments. Those segments have been shown to coincide with the binding sites for the cytoplasmic membrane, RNA and cognate proteins[7,83]. Apart from the segments 633-662 and 685-712 all conserved sites show a great predicted structural propensity. 633-662 and 685-712 on the other hand are predicted to form coiled coils, but the isolated peptide did not adopt a helical structure in isolation. It remains to be further investigated, whether RNA or RhlB binding is required for a disorder-to-formation[7,19,83].

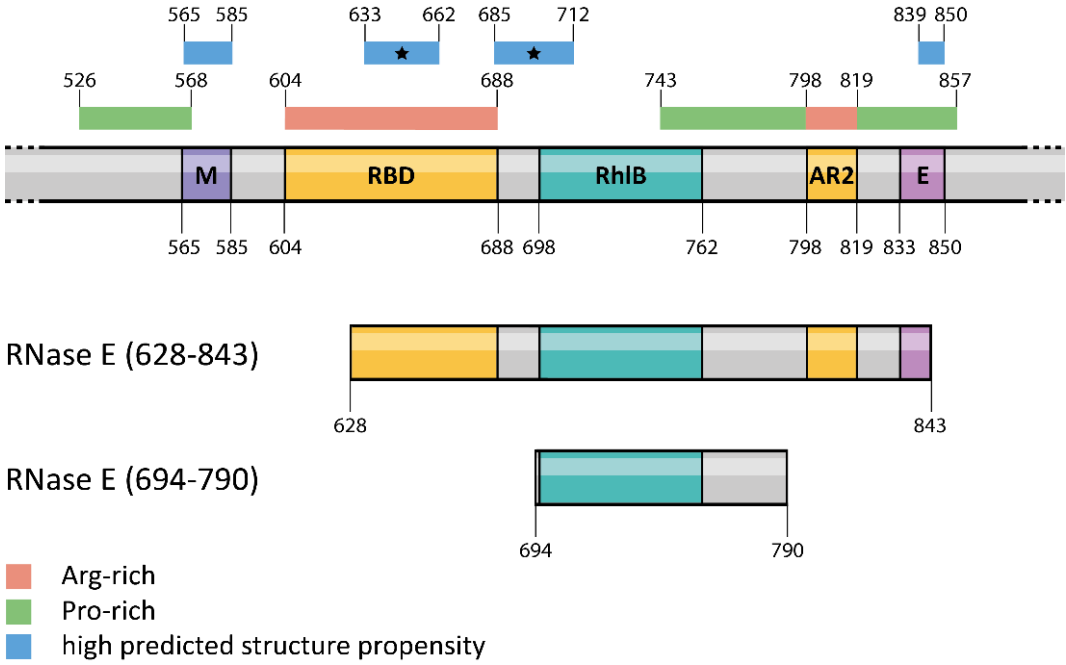

Figure 10. Schematic representation of the primary structure the C-terminal domain of RNase E (approximately residues 500 to 900). Indicated are the identified binding sites for complex partners: RBD (604-688) and AR2 (798-819) bind RNA, RhlB recognition site (698-762) and Enolase (E) (833-850). The membrane anchor encompasses residues 565-585. Arginine-rich (red) and proline-rich (green) regions are colour coded as well as regions with greater predicted propensity for structure formation (blue); the asterisk marks the segments with high propensity for coiled-coil formation[19]. Below are the schematic representations of RNase (628-843) and RNase E (694-790).

By comparison, both RNase E (628-843) and RNase E (694-790) have been shown to be roughly equivalent in stimulating the ATPase activity of RhlB in assays performed with yeast bulk RNA, their effect on RhlB's unwinding activity however differs quite significantly: RhlB's unwinding activity was boosted from 8% to 60% with RNase E (628-843) whereas RNase E (694-790) increased the activity only to 14%[18,20]. This difference was proposed to be due to the RNA binding segments RBD and AR2, suggesting a dynamic interplay between the RNA binding segments of the RNase E and the helicase during the unwinding process. As of now, no further experiments have been performed with the minimal interaction fragment RNase E 698-762 to assess its stimulating effect on RhlB's ATPase or unwinding activity.

So, while the effects of the communication between RNase binding and ATPase activation have been attested, the question as to how the interaction boosts the helicases activity and whether and how the helicases own RNA substrate preferences are intertwined with it, are unfortunately still not clear.

Chapter 2 Motivation

DEAD-Box helicases play a significant role in virtually all aspects of eukaryotic and prokaryotic RNA metabolism, from transcription, translation, and RNA decay to ribosome and snRNP biogenesis. Substantial insight has been gained in the past decades regarding the process of the RNA unwinding mechanism of DEAD-Box helicases. Simultaneously, research has been recognizing and increasingly been focusing on how individual DEAD-Box proteins utilize ancillary domains and interaction partners to adapt and expand the fundamental unwinding mechanisms to fulfil their specific role within the cell. Although *E. coli* frequently serves as prototype organism when it comes to research on prokaryotic proteins, its DEAD-Box helicase RhlB and how the helicase facilitates the degradation of RNA transcripts as part of the multi-protein degradosome complex by interacting with complex partner RNase E is only partially understood.

To investigate the potential interplay of substrate selectivity and allosteric activation for the unique DEAD-Box helicase RhlB, it was the aim pf the PhD thesis to study the interaction of RhlB with RNA substrates with distinct topological features in the presence and absence of complex partner RNase E. Those experiments should reveal whether RhlB does have an innate RNA substrate preference in other aspects of the helicase reaction (RNA binding, ATP turnover) and whether those preferences are selectively affected by the allosteric activation of RNase E fragments (628-843) and (694-790). This thesis will also focus on assessing the protein activation from the perspective of the bound RNA substrate to obtain a deeper understanding of possible structural changes in the RNA binding pocket caused by the allosteric interaction. To gain novel insights into the structural effects on the RNA substrate solution-state NMR-spectroscopic methods will be utilized, as it enables structure assessments with nucleotide and even atomic resolution at measurement conditions that are significantly more physiological than previously published X-ray crystallography data of related DEAD-Box helicases using Poly(U) RNAs[21,98].

This thesis is divided into two parts: The first part describes the design, preparation, and biochemical as well as NMR-spectroscopic characterization of potential substrate RNAs for further investigation of RhlB's helicase reaction. The aim is to identify RNAs constructs that will differ in their featured single-stranded tails and are suitable for NMR-spectroscopic investigations. This also includes NMR resonance assignment of imino and nucleobase resonances for auspicious RNA constructs. In the second part the assessment of RhlB's ATPase activity and RNA binding affinity for the different RNA substrates under the influence of RNase E (628-843) or RNase E (694-790) is described. ATPase assay and standard as well as real-time kinetic NMR experiments will reveal the RNases contribution to other aspects of RhlB's reaction cycle. To further probe the conformation of the RNA substrate in the binding pocket 2D NMR titration experiments with selectively labelled RNA were conducted. In addition, the overall effects of RNase E (694-790) on the dynamic and structure of RhlB and vice versa were investigated by NMR spectroscopy on isotope-

labelled protein-protein complexes and it was examined whether RNase E (694-790) undergoes a structure formation when binding to RhlB.

Chapter 3 Materials and Methods

3.1 Buffer, Media, and Solutions

All required recipes for buffer, media and solutions used in this work are listed in appendix 8.1.

3.2 RNA Construct Design

The RNA constructs were designed to meet both the requirements of model substrate RNAs suitable for the DEAD-Box helicase RhlB as well as fulfil the size constraints given by NMR spectroscopy. As explained in more detail in section 1.3.2 DEAD-Box helicases are capable of efficiently unwinding RNA duplexes with a length of up to two helical turns, which corresponds to 22 Nt in A-type RNA[128,129]. In previous studies performed on both prokaryotic and eukaryotic DEAD-Box helicases, the duplex length of artificial substrate RNAs ranged from about 10-20 Nt with a thermal stability of around 15-25 kcal mol^{-1}[16,18,91]. Stampfl *et al.* could even demonstrate a successful duplex displacement for RNA substrates with a minimum free energy of up to 31 kcal mol^{-1} for the DEAD-Box helicase CsdA[124]. On basis of that, the RNA constructs were adapted from Stampfl's RNA substrate J2h-M2 with a 21 Nt double stranded and a 3' hairpin segment as well as a minimum free energy of 23.0 kcal mol^{-1}, since variations of this RNA duplex have been successfully used as substrates for both CsdA and DbpA studies[102,108].

In an approach to build an accessible RNA toolbox, other RNA substrates were constructed as fragments of the RNA strands J2h and M2. This did not only allow for a straight-forward strand selective isotope labelling but also reduced the number of individual RNA strands that needed to be synthesized and assigned by NMR spectroscopy. Figure 11 illustrates, how the individual RNA strands can be combined to receive RNA heteroduplexes with varying features. In total, 6 different RNA duplexes were chosen for investigation: the prototype construct J2h-M2 with a 2 Nt overhang at the 5' end and a hairpin at the 3' end, two blunt end duplexes J2Δ8-M3 and J2-M2 with 13 Nt and 21 Nt duplex length, respectively, duplex J2Δ8-M2 with an 8 Nt long single strand overhang at the 3' end and 13 Nt duplex, and J2Δ14-M2 with an 8 Nt single strand overhang at the 5' end and 13 Nt duplex. For the 3' and 5' tailed constructs we aimed for matching duplex length and comparable duplex stability (predicted minimum free energy) to limit possible differences in performance to the single strand extension. But because of the non-palindromic nature of the sequence the predicted minimum free energy differed considerably with identical duplex length of 13 Nt (15.5 and 9.8 kcal mol^{-1}). We therefore included an additional 5' tailed construct with an extended duplex of 15 Nt and 6 Nt single strand (J2Δ16-M2) that resembled a closer match to the corresponding 3' tailed construct in terms of duplex stability.

The theoretical melting temperatures, predicted with the Pairfold application of the RNAsoft programs, and the theoretical minimum free energies of the paired duplexes, calculated with the RNAcofold application of the ViennaRNA web suite, are listed in Table 1 together with the duplex lengths[130–132]. Any secondary structure predictions for RNA single strands were performed at 15°C and a fixed ionic strength of 1M NaCl using the mFold web application version 2.3 of the UNAFold web server[133].

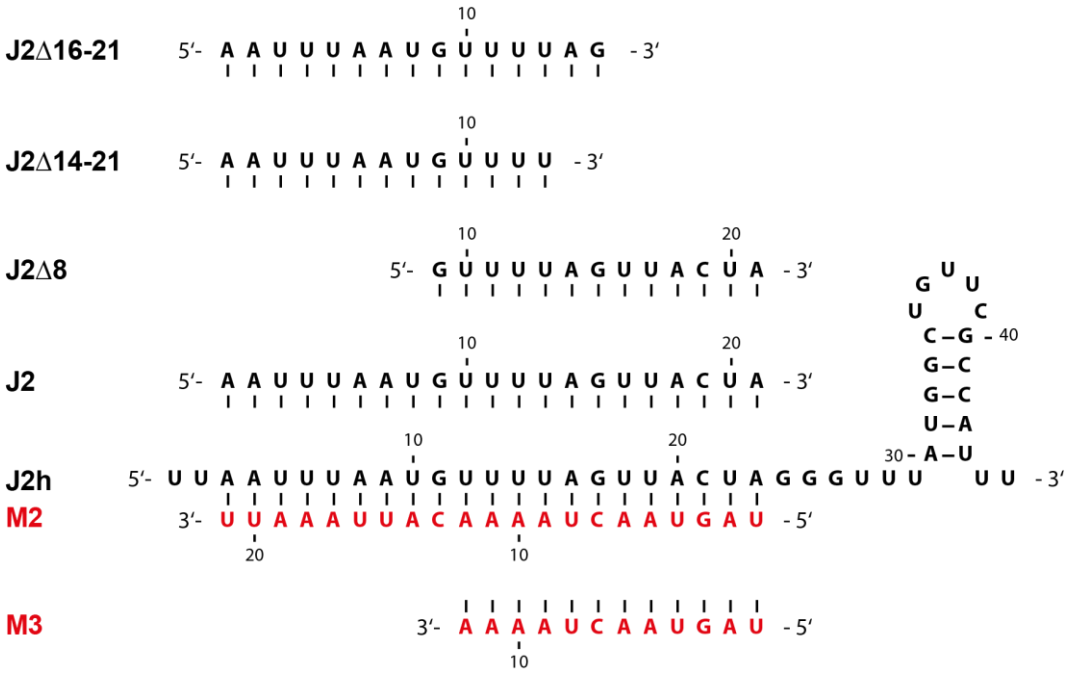

Figure 11. RNA sequences of hetero duplex constructs, aligned to highlight duplex pairing of black with red strands. The constructs J2Δ14-21 and J2Δ16-21 will subsequently be referred to as J2Δ14 and J2Δ16, respectively.

Table 1. Length, predicted minimum free energy, and predicted melting temperature for all RNA heteroduplexes.

RNA duplex	Duplex length [Nt]	Predicted Minimum free energy ΔG [kcal mol^{-1}]	Predicted melting temperature T_m [°C][a]
J2h-M2	21 + 5	-23.02[b]	64.3
J2-M2	21	-22.60	62.0
J2Δ8-M2	13	-15.50	54.4
J2Δ8-M3	13	-13.89	50.8
J2Δ14-M2	13	-9.82	36.1
J2Δ16-M2	15	-13.63	47.8

[a] predicted with PairFold for RNA concentration of 10 µM and 150 mM ionic strength.

[b] Hetero dimer formation

3.3 DNA Plasmid Preparation

3.3.1 DNA Templates for RNA Synthesis

Given in Table 2 are the sequences of all RNA constructs studied in this project. Unlabelled RNAs were both purchased as solid-phase synthesis products by Integrated DNA Technologies Inc. and synthesized in-house by *in vitro* transcription (except for M3 and J2Δ14-21). Full $^{13}C^{15}N$ labelled M2 RNA was also synthesized by *in vitro* transcription, while M2 with selective [5-D, 6-^{13}C] and [8-^{13}C] labelling at position 1 to 8 was purchased by INNotope GmbH.

Table 2. Sequences and synthesis information on all studied RNA constructs.

Construct name	Sequence (5' → 3')	Length [Nt]	*In vitro* transcribed	Purchased
M2	UAGUAACUAAAACAUUAAAUU	21	unlabelled, full ^{13}C,^{15}N lab.	unlabelled, sel. ^{13}C lab. [a]
M3	UAGUAACUAAAAC	13	-	unlabelled
J2h	UUAAUUUAAUGUUUUAGUUACUAGGGUUUAUGGCUGUUCGCCAUUU	46	unlabelled	unlabelled
J2	AAUUUAAUGUUUUAGUUACUA	21	unlabelled	unlabelled
J2Δ16	AAUUUAAUGUUUUAG	15	unlabelled	unlabelled
J2Δ14	AAUUUAAUGUUUU	13	-	unlabelled
J2Δ8	GUUUUAGUUACUA	13	unlabelled	unlabelled

[a] Position of selective [5-D, 6-^{13}C] and [8-^{13}C] labelling is underlined in sequence.

DNA sequences for *in vitro* transcribed RNA constructs were designed with both 5' hammerhead (HH) ribozyme (51 Nt) and 3' Hepatitis delta virus (HDV) ribozyme (67 Nt) to increase transcription efficiency and product homogeneity[134]. In case of J2h the ribozyme sequences were elongated to 75 Nt as otherwise its length would have matched that of the RNA of interest. For J2, J2Δ8 and J2Δ16-21 the P1 stem was elongated by one base pair to increase its stability. Since the P1 stem of the hammerhead ribozyme is formed by both the ribozyme and the downstream RNA sequence, this segment of the HH ribozyme was adjusted with each new RNA construct. A restriction site for *SmaI* or *EcoRI* was included at the 3' of the cassette to perform run-off transcriptions with a linearized DNA template. The DNA sequences of the RNA transcription cassettes shown in Table 3 were purchased by GenScript inserted into pUC57 or pEX-A vector (pEX-A vector in case of J2Δ16-21).

Table 3. 5' to 3' DNA sequences of RNA transcription cassettes including T7 promotor (green), hammerhead ribozyme (red), RNA of interest (black), HDV ribozyme (blue) and SmaI (CCCGGG) or EcoRI (GAATTC) restriction site (orange).

RNA construct	DNA sequence of RNA transcription cassette (5'→3')
M2	TAATACGACTCACTATAGGGAATTTACTACTGATGAGAGCGAAAGCTCGAAACAGCTG TGAAGCTGTCTAGTAACTAAAACATTAAATTGGGTCGGCATGGCATCTCCACCTCCTCG CGGTCCGACCTGGGCTACTTCGGTAGGCTAAGGGAGAAGCCCGGG
J2h	TAATACGACTCACTATAGGGAATATTAAATTAACTGATGAGAGCTCCAGAAGCGGAAA CGCTTCTGGAGCTCGAAACAGCTGTGAAGCTGTCTTAATTTAATGTTTTAGTTACTAGG GTTTATGGCTGTTCGCCATTTGGGTCGGCATGGCATCTCCACCTCCTCGCGGTCCGACC TGGGCTACCAAGTTCGCTTGGTAGGCTAAGGGAGAAGCCCGGG
J2	TAATACGACTCACTATAGGGAATTTAAATTCTGATGAGAGCGAAAGCTCGAAACAGCT GTGAAGCTGTCAATTTAATGTTTTAGTTACTAGGGTCGGCATGGCATCTCCACCTCCTC GCGGTCCGACCTGGGCTACTTCGGTAGGCTAAGGGAGAAGCCCGGG
J2Δ16	TAATACGACTCACTATAGGGAATTTAAATTCTGATGAGAGCGAAAGCTCGAAACAGCT GTGAAGCTGTCAATTTAATGTTTTAGGGGTCGGCATGGCATCTCCACCTCCTCGCGGT CCGACCTGGGCTACTTCGGTAGGCTAAGGGAGAAGGAATTC
J2Δ8	TAATACGACTCACTATAGGGAATCTAAAACCTGATGAGAGCGAAAGCTCGAAACAGCT GTGAAGCTGTCGTTTTAGTTACTAGGGTCGGCATGGCATCTCCACCTCCTCGCGGTCCG ACCTGGGCTACTTCGGTAGGCTAAGGGAGAAGCCCGGG

3.3.2 Transformation

The plasmid amplification was carried out in *E. coli* DH5α cells. For the transformation 50 µL of the competent cells were incubated with 50-100 ng of plasmid DNA on ice for 30 min and subsequently heat shocked for 45 s at 42°C in a water bath. After incubation on ice for 5 min, 450 µL SOC medium were added. The sample was incubated at 37°C for 1 h, before 100 µL were spread onto a lysogeny broth (LB) agar plate containing 100 µg/mL ampicillin. The plates were then incubated overnight at 37°C. After about 16 h bacteria colonies could be picked and after 24 h the agar plates were sealed with Parafilm and stored at 4°C for up to 4 weeks.

3.3.3 DNA Amplification and Purification

Small scale DNA preparation (Miniprep)

For small scale plasmid DNA preparation an individual bacteria colony was picked from the agar plate and used to inoculate 5 mL lysogeny broth (LB) medium containing 100 µg/mL ampicillin. The 5 mL culture was incubated at 37°C and 160 rpm for 10-16 h and cells were harvested by centrifugation at 6800 *g* and 4°C for 10 min. The cell lysis and

DNA isolation were performed according to the QIAprep® Spin Miniprep Kit (QIAGEN) and the concentration of extracted DNA was determined using UV/Vis spectroscopy. The sequence identity of the insert was verified via sequencing.

Large scale DNA preparation (Megaprep)

For large scale plasmid DNA preparation 5 mL LB medium with 100 µg/mL ampicillin were inoculated with either an individually picked bacteria colony from an agar plate or 20-50 µL of prepared glycerol stocks. This starter culture was incubated overnight at 37°C and 160 rpm. The starter culture was then transferred into 1 L super broth (SB) medium with 100 µg/mL ampicillin to inoculate the main culture. 20-40 µL antifoam Y-30 solution (Sigma Aldrich) were added to ensure optimal aeration. The main culture was incubated for 16 h at 37°C and 120 rpm. Before cell harvest, samples for a glycerol stock solution could be prepared by mixing 500 µL of cell suspension with 500 µL 50% glycerol and flash-freezing it in liquid nitrogen. The glycerol stocks were then stored at -80°C and used for later inoculations. For the cell harvest the culture solution was centrifuged for 15 min at 6000 *g* and 4°C. Subsequent cell lysis and DNA isolation were performed using the NucleoBond® Plasmid purification kit (Macherey Nagel) with NucleoBond® AX 2000 columns. The concentration of extracted DNA was determined using UV/Vis spectroscopy.

3.3.4 Plasmid Linearization by Restriction Digest

Plasmids were linearized by restriction digestion using restriction enzyme *SmaI* or *EcoRI*. For an analytical digest 1 µg of plasmid DNA was incubated with 10 U of *SmaI* or *EcoRI-HF®* enzyme (New England Biolabs) in 1x CutSmart® Buffer (New England Biolabs) in a total reaction volume of 20 µL for 4 h at 25 °C or 37°C for *SmaI* or *EcoRI*, respectively. The completion of the linearization was verified by agarose gel electrophoresis. For the preparative scale linearization 1 mg of plasmid DNA was incubated with 250 U of restriction enzyme in 400 µL 1x CutSmart® Buffer for 16 h at corresponding reaction temperature.

The linearized DNA was then purified from the reaction solution by phenol-chloroform-isoamyl alcohol (PCI) extraction. Therefore, the aqueous solution was mixed thoroughly with an equal volume of ROTI® Phenol-Chloroform-Isoamyl alcohol (ROTH) and filled into a 5PRIME Phase Lock Gel tube (Quantabio). After phase separation by centrifugation at 12-16000 *g* for 5 min, the aqueous upper phase was carefully pipetted into a fresh Phase Lock Gel tube. The extraction was repeated, and the aqueous phase transferred into a regular reaction tube. For precipitation the DNA solution was then mixed with one-tenth volume of 3 M sodium acetate, before 2.5 volumes of absolute ethanol were added. The solution was stored at -20°C for 1 h, centrifuged for 30 min at 9000 *g* and 4°C and the supernatant carefully decanted from the pelleted DNA. The pellet was subsequently washed with 5 mL of 70% ethanol before centrifuging again for 15 min at 9000 *g* and 4°C.

The supernatant was removed, and the pellet air-dried for 10-15 min before being resolved in 0.5-2 mL of double distilled and deionized water (ddH_2O).

3.4 RNA Synthesis and Purification

3.4.1 *In vitro* Transcription

In vitro transcriptions of RNA from linearized plasmid were performed with the P266L mutant of T7 polymerase (made in-house). While Table 4 displays the composition of the standard transcription reaction, several reaction conditions such as concentration of DNA plasmid, NTPs, DMSO and $Mg(OAc)_2$ as well as incubation time (4-16 h) had to be optimized on an analytical scale for each individual RNA construct. Therefore, 25 µL transcription reactions were incubated at 37°C for 4 h and analysed by denaturing RNA-PAGE. Conditions providing the highest yield as well as the least amount of side products were chosen for preparative scale transcription. For the synthesis of isotope labelled RNAs fully ^{13}C ^{15}N labelled rNTPs (Silantes) were used.

Table 4. Composition of in vitro transcription reaction. Components that require optimization are highlighted in blue and a typical concentration range is given.

Component	Concentration
Tris/glutamate pH 8.1	100 mM
DTT	20 mM
Spermidine	2 mM
$Mg(OAc)_2$	10-60 mM
rNTPs [a]	5-20 mM
Linearized DNA plasmid	50-200 ng/µL
DMSO	0-25 %
T7 polymerase (P266L)	70 µg/mL

[a] Ratio of individual rNTPs are adjusted to NTP composition of transcript

For a large-scale transcription reaction (10-15 mL) all components besides T7 polymerase and DTT were mixed and incubated at 37°C and 100 rpm for about 30 min before the remaining components were added. The incubation at 37°C was continued for the time previously optimized (typically 8 h) with the addition of 1-2 U yeast inorganic pyrophosphatase (New England Biolab) after 2 h to hydrolyse insoluble magnesium pyrophosphate. Afterwards the reaction mix was centrifuged for 30 min at 8000 *g* and 4°C and the supernatant was carefully decanted from possible salt pellet. The salt pellet was then washed with 5 mL ddH_2O, centrifuged again for 15 min at 8000 *g* and the supernatant pooled with the previous one.

3.4.2 RNA Purification via Anion Exchange Chromatography

In the first purification step anion exchange chromatography with diethylaminoethyl (DEAE) resin was used to separate proteins, DNA plasmid and remaining rNTPs from the reaction mix. Therefore, a self-packed gravity-flow chromatography column (inner diameter 5.0 cm) filled with 10 mL DEAE Sepharose™ Fast flow resin (GE Healthcare) was used. The packed column was washed with 100 mL 0.1% (v/v) diethylpyrocarbonate (DEPC) solution before being incubated in 0.1% DEPC overnight. The column was then washed with 200 mL hot ddH_2O (~60°C) to remove residual DEPC before being equilibrated with 50 mL 0.1 M sodium acetate (pH 5.5). The combined supernatant of the centrifuged transcription solution was applied to the column and the resulting flow through loaded onto the column again. The column was successively washed with 50 mL of 0.6 M, 1 M, 2 M and 3 M sodium acetate and the eluate collected in fraction of 10 mL. For all fraction UV absorption at 260 nm was measured and samples with a high absorption were further analysed via denaturing RNA-PAGE. Fractions that were identified to contain the RNA products of interest were then diluted with ddH_2O to approximately 0.6 M sodium acetate and precipitated with 4 volumes of ice-cold absolute ethanol and stored at -20°C overnight. The samples were then centrifuged for 30 min at 8000 *g* and 4°C, the supernatant was carefully decanted, and the pellet air-dried for 10-15 min before being resolved in enough ddH_2O to yield an absorption at 260 nm of approximately 100 A.U..

The samples, which at this point only contained the desired RNA product as well as RNA ribozymes and/or uncleaved full-length transcript, could subsequently be purified by either reversed-phase (rp) HPLC or preparative polyacrylamide gel electrophoresis. While rp-HPLC was the preferred method, denaturing PAGE was necessary in cases where the difference in length between ribozymes and RNA of interest were too small for a complete separation via HPLC.

3.4.3 RNA Purification via Reversed-phase HPLC

All HPLC purifications were performed by Elke Stirnal.

For the reversed-phase high performance liquid chromatography (rp-HPLC) purification a Hitachi D-7000 HPLC system equipped with a UV/Vis detector was used. The RNA sample was loaded onto a Perfectsil RP18 300A 5µm 10x250 mm column (MZ Analysentechnik) equilibrated with HPLC buffer A at 60°C. At a flow rate of 5 mL/min the following gradients of HPLC buffer B were applied: 0-37% in 5 min, 37-40% in 30 min and 40-100% in 5 min. The HPLC elution fractions were analysed via denaturing RNA-PAGE to identify fractions with the desired RNA product. Those fractions were lyophilized for acetonitrile removal, combined, and subsequently desalted and concentrated using Vivaspin® 20 Ultrafiltration units (Sartorius) with a cut-off of approximately a third of the RNAs

molecular weight. To substitute the RNA's counterions from ammonium to lithium, the RNA samples were precipitated by addition of 5 volumes 2% (w/v) lithium perchlorate in acetone. After incubation at -20°C for at least one hour, the RNA was pelleted by centrifugation at 8000 *g* and 4°C for 30 min. After removal of the supernatant, the pellet was washed with 70% ethanol and centrifuged again for 15 min at 8000 *g* before being air-dried for 15 min and resuspended in 1-2 mL ddH_2O.

3.4.4 RNA Purification via preparative Polyacrylamide Gel Electrophoresis

If rp-HPLC was not applicable, RNA products could be further separated from ribozyme byproducts via preparative scale denaturing polyacrylamide gel electrophoresis. Therefore a 16.0 x 47.5 cm 10% denaturing polyacrylamide gel was prepared with 10% (v/v) acrylamide/ bis-acrylamide (29:1) and 7 M urea in TBE buffer. For polymerization 0.1% (w/v) ammonium persulfate (APS) and 0.1% (v/v) tetramethylethylenediamine (TEMED) were added to the gel solution. The RNA sample was mixed with an equal volume of formamide, and a total of 2.5 mL sample was loaded into the gel pocket. A separate pocket filled with denaturing RNA loading dye was used as a reference to monitor the gel migration front. The gel was run in TBE buffer with a TVS1000 sequencer electrophoresis system (Biostep) at 50 W for 5-7 h with the integrated ventilator being turned on after about one hour. The separated bands were visualized with a UV lamp at 254 nm over fluorescent TLC plates. In the process, UV exposure of the RNA bands was kept to a minimum to avoid RNA damage. Relevant RNA bands were excised and pushed through a syringe to maximize gel surface. 15-25 mL 0.3 M sodium acetate (pH 5.5) were added to the gel pieces to extract the RNA over an incubation time of 24-48 h at room temperature on a rocking shaker. The RNA containing sodium acetate solution was subsequently sterile filtered and the RNA precipitated by addition of 2.5 volume of absolute ethanol. After incubation at -20°C overnight, the RNA was pelleted by centrifugation at 8000 *g* and 4°C for 30 min. The pellet was let to air-dry for 15 min before being resolved in 1-2 mL ddH_2O.

3.4.5 RNA Folding and Buffer Exchange

RNA samples were exchanged into NMR buffer using Vivaspin® 2 Ultrafiltration units (Sartorius). For homogenous duplex formation and correct folding equimolar amounts of the corresponding RNA strands were mixed, heated up to 95°C for 5 min and then rapidly cooled on ice for 5 min. The folding was inspected by native polyacrylamide gel electrophoresis.

3.5 Protein Expression and Purification

3.5.1 DNA Templates for Protein Expression

The protein sequences of full-length wild-type RhlB and RNase E were obtained from the Universal Protein Resource (UniProt) database (Entry no. #P0A8J8 and #P21513) and RNase E fragments RNase E (694-790) and RNase E (628-843) were chosen for biosynthesis. Both RNase E constructs, in the further course of this thesis in part also abbreviated as RnE 694 and RnE 628, were designed with a N-terminal TEV-cleavable His_6-tag, while previous expression tests with RhlB showed a higher expression rate for a C-terminally His_6-tagged protein (Diplomarbeit Heidi Zetzsche). Figure 12 shows the designated expression platforms for each protein construct.

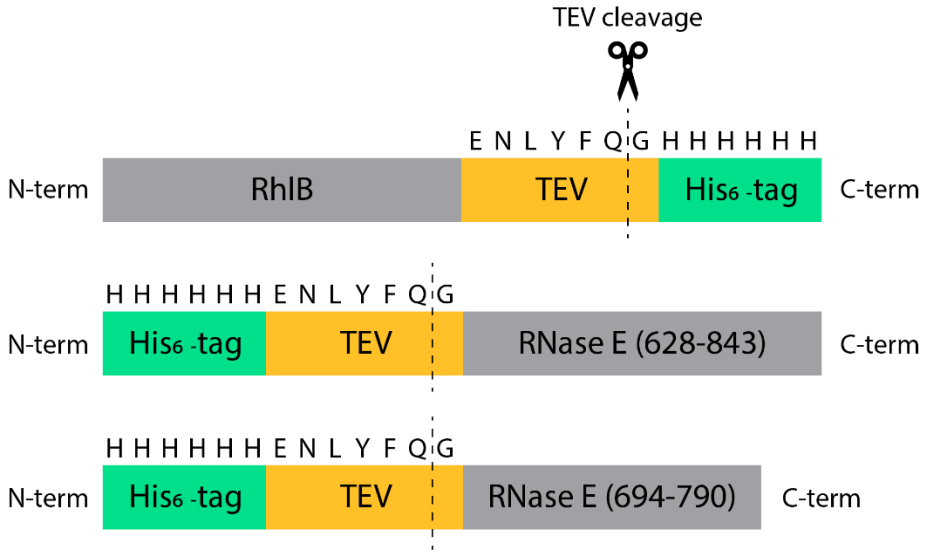

Figure 12. Design of expression platforms for RhlB, RNase E (628-843) and RNase E (694-790). All proteins were expressed in fusion with N- or C-terminal His_6-tag and TEV cleavage site.

Table 5 summarizes the molecular weights and isoelectric points of both the His_6-tagged proteins as well as the proteins after TEV cleavage. The pH of each protein buffer was adjusted to be at least one pH unit off from the isoelectric point to avoid precipitation.

Table 5. Molecular weights (MW), isoelectric points (pI) and number of amino acids (AA) of expressed unlabelled proteins with His_6-Tag and after TEV cleavage. Both molecular weights and isoelectric points were calculated by the protParam tool of the Expasy web suite.

Protein construct	**+ His_6-tag-TEV**			**tag cleaved off [a]**		
	MW [Da]	**pI**	**AA**	**MW [Da]**	**pI**	**AA**
RhlB	48800.7	7.10	434	47920.8	6.98	427
RNase E (694-790)	12608.9	5.20	110	10991.2	4.65	98
RNase E (628-843)	26775.3	9.40	229	25157.6	9.81	217

[a] after cleavage of N-terminal or C-terminal tag via TEV protease; the target protein remains with an additional G or ENLYFQ from the TEV cleavage site, respectively.

The corresponding gene sequences were purchased from Dharmacon™ cloned into pET11A (RhlB, RNase E (628-843)) and pET21A (RNase E (694-790)) expression vectors via *NdeI* and *BamHI* restriction sites. Full amino acid and DNA sequences are given in.

3.5.2 Transformation

DNA plasmids for protein expression were transformed into *E. coli* BL21 DE3 cells as described in section 3.3.2.

3.5.3 Protein Expression

Analytical scale expression

To optimize expression conditions for individual plasmids test expressions were performed on an analytical scale. Therefore, 5 mL lysogeny broth (LB) or terrific broth (TB) medium supplemented with 100 µg/mL ampicillin were inoculated from either agar plate colonies or glycerol stock and incubated at 37°C and 120 rpm. OD_{600} was measured in regular intervals until 0.6 or 1.5 for LB and TB medium was reached, respectively. The cultures were split into volumes of 1 mL and expression was induced by addition of 0.3-1 mM isopropyl β-d-1-thiogalactopyranoside (IPTG). As a negative control a 1 mL culture was kept uninduced. Incubation was continued at either 37°C or 21°C for 3-16 h before 50 µL of cells were harvested by centrifugation at 8000 *g* and 4°C for 10 min. Cell pellets were resuspended in 30 µL SDS loading buffer, denatured at 95°C for 20 min and 10 µL analysed by SDS-PAGE.

Preparative scale expression

For a large-scale expression of unlabelled protein, a 50 mL starter culture (LB medium and 100 µg/mL ampicillin) was inoculated from either agar plate colonies or glycerol stock. The culture was incubated overnight at 37°C and 120 rpm and used to inoculate 1 L main culture (LB or TB medium, 100 µg/mL ampicillin and 30 µL/L antifoam Y-30 solution) with a starting OD_{600} of 0.1. The main culture was incubated at 37°C and 120 rpm and the OD_{600} monitored in regular intervals. When OD_{600} reached 1.5 (optimal induction point for TB medium; OD_{600} of 0.6-0.8 for LB medium) the expression was induced by addition of IPTG with a final concentration of 0.3-1 mM depending on the expression plasmid. Incubation was then continued at 21°C and 120 rpm for 3-4 h (TB medium) or overnight (LB medium) before cells were harvested by centrifugation at 6000 *g* and 4°C for 15 min.

Expression of isotope labelled proteins was achieved with M9 minimal medium supplemented with $^{15}NH_4Cl$ and ^{13}C glucose in the main culture. For that, the full-grown LB medium starter cultures were centrifuged for 10 min at 4000 *g* and 4°C and the pelleted cells were resuspended in 50 mL M9 minimal medium. After additional incubation for 4-6 h at 120 rpm and 37°C this culture was then used to inoculate 1 L main

culture (M9 minimal medium and 100 µg/mL ampicillin) to a starting OD_{600} of 0.1. Incubation was continued at 37°C and 120 rpm and expression was induced by addition of 0.3-1 mM IPTG at OD_{600} of 0.6-0.8. Expression was continued at 37°C and 120 rpm overnight and cells were harvested by centrifugation at 6000 *g* and 4°C for 15 min.

3.5.4 Protein Purification

All media, buffer, and solutions used for protein expression and purification can be found in appendix 8.1.

The cells were resuspended on ice in buffer A (5-10 g cells per 50 mL buffer) supplemented with cOmplete™ EDTA-free protease inhibitor tablets (Roche) and lysed via high-pressure homogenization using the Microfluidics M-110P at 15000 PSI. The lysate was then clarified by centrifugation for 45 min at 16000 *g* and 4°C. To remove nucleic acids from the cell suspension, 0.03% (w/v) polyethylenimine (PEI) were added to the lysate. After incubating for 15 min at 4°C on a rocking shaker, the formed precipitate was pelleted by centrifugation for 15 min at 8000 *g* and 4°C.

For all the following chromatographic purification steps the ÄKTApurifier™ FPLC system (GE Healthcare) was used at 4°C. The clarified supernatant was loaded onto a 5 mL HisTrap™ HP column (GE Healthcare) equilibrated with buffer A and washed with 5-10 column volumes (CV) buffer A followed by 5 CV of LiCl buffer to remove protein-bound nucleic acids (this step was only performed for purification of RhlB and RNase E (628-843)). The proteins were eluted with a gradient of 100% buffer B over 100 mL and collected in 5 mL fractions. Protein containing fraction were identified by SDS-PAGE and combined. The His-Tag was cleaved off through addition of TEV protease (tobacco etch virus) during overnight dialysis against 5 L buffer A and subsequently removed via reverse Ni-NTA chromatography. Promising protein fractions were again identified via SDS-PAGE, pooled, and concentrated to approximately 3 mg/mL using Vivaspin® 20 Ultrafiltration units (Sartorius) with a molecular weight cut-off of at least half the proteins molecular weight. 10 mL of the protein solutions were then applied via a 10 mL loop onto a HiLoad® 26/600 Superdex® 75 pg or HiLoad® 26/600 Superdex® 200 pg size exclusion column (GE Healthcare) equilibrated in buffer C and eluted with 1.2 CV of buffer C in fractions of 5 mL. Elution fractions were analysed by SDS-PAGE and fractions containing pure protein were combined, flash-frozen in liquid nitrogen and stored at -80°C.

For both NMR experiments and assays, suitable amounts of proteins were freshly thawed and rebuffered into pH 8 NMR buffer using Vivaspin™ centrifugal concentrators.

3.5.5 RNase Contamination Test

As RNAs, particularly unstructured single strands, are prone to degradation by RNases, meticulous attention to RNase-free work conditions is crucial. Those RNase impurities can arise either from contaminated lab equipment or insufficient removal of *E. coli* internal RNases during protein purification. Besides using as much sterile lab equipment as possible and incubating materials in DEPC solution prior to use every newly purified protein batch was therefore tested for RNase contamination before being used for any experiments involving RNA substrates. For the test 100 pmol of single stranded RNA (typically single stranded substrate RNA "M2") were incubated with 0.5-1 nmol protein in 10 µL pH 8 NMR buffer for 1-4 h at 37°C and analysed on a denaturing RNA-PAGE against a control sample without protein. In case of a RNase contamination the RNA substrate would either appear fragmented or completely degraded into single nucleotides.

Protein samples that exhibited RNase activity were applied onto a 5 mL HiTrap™ Heparin HP column (GE Healthcare) equilibrated in pH 8 NMR buffer, washed with 10 CV of pH 8 NMR buffer, and eluted in 5 mL fractions with a gradient of 100% buffer H in 65 min. Protein containing fractions were identified via SDS-PAGE and combined. A second RNase test was performed to verify whether RNase impurities were in fact removed. After a successful removal of RNase contaminants, the buffer of the protein fractions was exchanged back to pH 8 NMR buffer either via dialysis or centrifugal concentration.

3.6 Analytics

3.6.1 Analytical Gel Electrophoresis

Agarose gel electrophoresis:

Agarose gel electrophoresis was used to analyse both supercoiled and linearized plasmid DNA. For the agarose gel 20-30 mL 1% (w/v) agarose powder were dissolved in TBE or TAE buffer, heated and cast into a gel. 200 ng of DNA were prepared in gel loading dye (New England Biolabs) or house-made DNA loading with a total volume of 6 µL and loaded into gel pockets. 1 kb Plus DNA ladder (New England Biolabs) was used as a size reference. DNA samples were then separated for 45 min at 120 V in TBE buffer using a Sub-Cell® horizontal electrophoresis system (Bio-Rad). After staining the gel with GelRed® solution (Biotium) for 10-20 min, DNA bands were visualized via UV-transillumination in a Gel iX20 Imager (Intas Science Imaging).

Polyacrylamide gel electrophoresis (PAGE):

For the analysis of RNAs polyacrylamide gel electrophoresis (RNA-PAGE) was utilized. The separation of RNA strands exclusively by molecular weight was achieved under denaturing conditions as follows:

10 mL of 10-15% denaturing RNA-PAGE gel solution prepared and polymerized by addition of 0.1% (w/v) ammonium persulfate and 0.1% (v/v) TEMED. Typically, 10-100 pmol of RNA sample were prepared in denaturing RNA loading dye (New England Biolabs) or house-made denaturing RNA loading buffer in a total volume of 6-10 µL and heated to 95°C for 5 min, before being loaded into the gel pockets. The gel was run for 30 min at 220 V in TBE buffer in a Multigel system (Biometra), stained with GelRed® solution (Biotium) for 10-20 min and the RNA bands visualized via UV-transillumination. In the case of optical interference of RNA bands either dye, visualization by UV shadowing could be performed over a fluorescent TLC plate. If so, increasing the RNA concentration in the gel samples by a factor of at least 5 was necessary.

To distinguish RNA molecules also by secondary structure conformation and to inspect RNA folding, polyacrylamide gel electrophoresis was performed under non-denaturing conditions (here called "native PAGE"). Therefore, the following changes were made from the protocol above: Urea was omitted from gel solutions and native gel loading buffer was used for sample preparation. RNA samples were not heated to 95°C before loading into gel pockets and electrophoresis was performed either with water cooling or at 4°C at about 70 V (0.5 -1 W) over a time of 2-4 h.

Electrophoretic mobility shift assay:

To assess the binding of RNA to proteins, an electrophoretic mobility shift assay (EMSA) was used. This method utilized the difference in electrophoretic mobility of unbound RNA vs protein-bound RNA, where the band of the less mobile complex probe will be "shifted" up on the gel compared to the free RNA.

The binding reactions were carried out in EMSA buffer (150 mM NaCl, 25 mM Tris, 1.5 mM $MgCl_2$, 5 mM DTT, pH 8.3) and 10% (v/v) glycerol was added instead of RNA loading dye. 20 µL reactions were prepared with 1.5 µM RNA, increasing amounts of protein and excess of 3 mM ATP or AMP-PNP and incubated for 15 min at 4°C before being loaded onto a discontinued native gel (6% + 10% acrylamide). The gels were run in TA buffer (EDTA was omitted due to ATP) at 4°C with low power of about 30-70 V (~0.5 W) over a time of 5 h. Gels were again stained with GelRed® solution (Biotium) for 10-20 min and the RNA bands visualized via UV-transillumination. Band intensities were quantified using Gels could additionally be stained for protein using Coomassie staining solution and digitalized by Gel iX20 Imager (Intas Science Imaging). The fraction of bound RNA was calculated by quantifying the intensity of the free RNA band (GelRed staining) relative to the intensity of the lane containing only the RNA using ImageJ (after background subtraction). The fraction of bound RNA was plotted against protein concentration and the data were fitted for a simple single binding site function:

$$y = B \cdot \frac{x}{(K_D + x)} \qquad \textit{Equation 1}$$

where B is the upper plateau of the binding curve and K_D is the dissociation constant.

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE):

Proteins were separated and analysed by mass using sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). The gel was prepared by layered polymerization of 7 mL resolving gel and 3 mL stacking gel. Protein samples were prepared in SDS loading buffer, heated up to 95°C for 15 min and loaded into gel pockets. The Low Molecular Weight SDS Calibration Kit (GE Healthcare) was used for size reference. The gel was run for 50 min at 200 V in SDS running buffer using a Mini-PROTEAN® Tetra vertical electrophoresis cell (Bio-Rad). Gels were stained with Coomassie staining solution and digitalized by Gel iX20 Imager (Intas Science Imaging).

3.6.2 Concentration Measurement via UV-Vis Spectrophotometry

The concentrations of the studied biomolecules were determined by measuring their specific absorbance at their corresponding absorption maximum λ_{max} (260 nm for nucleic acids; 280 nm for proteins) using a NanoDrop™ One or NanoDrop™ 1000 spectrophotometer (Thermo Fisher Scientific) and calculating the concentration using Lambert-Beer law:

$$E = \varepsilon \cdot c \cdot d \qquad \textit{Equation 2}$$

with E being the absorbance/extinction, ε being the molar extinction coefficient in [$L \cdot mol^{-1} \cdot cm^{-1}$], c being the concentration in [mol·L], and d being the path length in [cm].

For single stranded and double stranded DNA plasmids, the constant default nucleic acids extinction coefficients of 0.02 µL ng^{-1} cm^{-1} (dsDNA) and 0.03 µL ng^{-1} cm^{-1} (ssDNA) set by Nanodrop were used and concentrations were given in ng μL^{-1}. In case of RNA constructs, the extinction coefficients were calculated using the OligoAnalyser™ tool (Integrated DNA Technologies). Extinction coefficients of proteins were calculated using the ProtParam tool provided by the ExPASy Bioinformatics Resource Portal (Swiss Institute of Bioinformatics).

The following table lists the extinction coefficients and molecular weights of all studied RNA and protein constructs.

Table 6. Extinction coefficients and molecular weights of studied RNA and protein constructs.

Molecule	Extinction coefficient ε [L·mol^{-1}·cm^{-1}]	Molecular weight [g·mol^{-1}]
RNA constructs [a]		
M2	230600	6658.1
M3	141500	4116.6
J2h	465600	14575.6
J2	227300	6605.9
J2Δ8	138000	4064.4
J2Δ14-21	139600	4049.4
J2Δ16-21	163900	4723.8
Protein constructs [b]		
RhlB	37360	47920.8
RNase E (694-790)	1490	10991.2
RNase E (628-843)	9970	25157.6

[a] Molecular weights given are for solid-phase synthesized RNAs with dephosphorylated 5' and 3' end; RNAs transcribed by *in vitro* transcription feature a 2',3'-cyclic phosphate at the 3' end and therefore have a molecular weight of + 62 g mol^{-1}.
[b] Extinction coefficients and molecular weights given are for TEV protease cleaved proteins.

3.7 Circular Dichroism (CD) Spectroscopy

3.7.1 CD Spectroscopy of RNA

The melting point and overall thermal stability is an important parameter in the characterization of a new RNA construct, as it aids in identifying optimal experimental conditions such as buffer composition and a suitable temperature range for measurements.

To determine the melting point, the optical phenomenon of circular dichroism (CD) is utilized. Circular dichroism describes the ability of chiral molecules to absorb right and left circularly polarized light to different extents. Experimentally, this difference in absorption is measured as the so called ellipticity Θ_λ, which is defined by the following equation:

$$tan\theta_\lambda = \frac{I_R - I_L}{I_R + I_L} \qquad \textit{Equation 3}$$

with $I_{R,\,L}$ being the intensity of right and left circularly polarized light after the absorption and $\tan\Theta_\lambda$ approximating Θ_λ measured in [mdeg]. Since this absorption behaviour changes significantly with the secondary structure of the nucleic acid, this method is widely used to study folding and conformations of both DNA and RNA. In both types of nucleic acids, a helical double strand exhibits a significantly higher ellipticity than a single

strand, which allows us to monitor transitions from folded to unfolded state in a variable temperature measurement.

All measurements were performed on a Jasco J-810 spectropolarimeter (JASCO) equipped with a PTC-4235 L Peltier thermostatic cell holder (JASCO) using 1 mm (QS) quarts glass high performance macro-cuvettes (Hellma) with 200 µL of a 10 µM RNA sample in buffer. A constant purging of the measurement chamber with 2.0-2.2 L/min N_2 gas prevented the formation of ozone during operation. CD spectrum measurements were blanked for baseline correction beforehand with buffer and the wavelength of absorption maximum determined in CD spectra was set as the monitoring wavelength during variable temperature measurements. Unless stated otherwise, all CD spectra and melting curves of RNA constructs were recorded with the measurement parameters listed in Table 7.

Table 7. Measurement parameter set for CD spectra and CD melting curves of RNA constructs.

Parameter	Setting
CD spectrum	
Wavelength range	200-320 nm
Scanning speed	50 nm/min
Scanning mode	continuous
Data pitch (acquisition interval)	0.1 nm
Sensitivity	standard
D.I.T. (digital integration time)	1 s
Band width	1 nm
Accumulations	5
CD melting curve	
Start temperature	5 °C
Target temperature	95 °C
Data acquisition interval	0.2 °C
Ramp rate	1°C/ min
Wavelength	variable [a]
Sensitivity	standard
D.I.T. (digital integration time)	1-4 s
Band width	1 nm

[a] set to absorption maximum determined in CD spectrum.

Since the CD data measured are strongly concentration dependent, ellipticity Θ was converted to molar ellipticity [Θ] with the following equation:

$$[\theta] = \frac{\theta \cdot M}{10 \cdot c \cdot d} \qquad \textit{Equation 4}$$

In which Θ is the ellipticity in [mdeg], M is the molecular weight in [g·mol^{-1}], c is the concentration in [g·L^{-1}], d is the path length in [cm], and [Θ] is the molar ellipticity in [deg·cm^2·dmol^{-1}]. The melting curves were corrected according to Marky *et al.* (1987)[135] and fit with the following sigmoidal function in SigmaPlot v.12.5:

$$f = y_o + \frac{a}{1+(e^{\frac{-(x-x_0)}{b}})} \qquad \textit{Equation 5}$$

Here, a and b are the amplitude and steepness of the curve and the inflection point x_0 represents the melting temperature of the corresponding RNA. For J2h-M2, the only RNA construct with two separate double stranded segments, this double sigmoidal fit function was utilized:

$$f = y_o + \frac{a}{1+(e^{\frac{-(x-x_0)}{b}})} + \frac{c}{1+(e^{\frac{-(x-x_1)}{d}})} \qquad \textit{Equation 6}$$

The two melting points are described by inflection points x_0 and x_1. The corrected and fit CD data could also be converted into fractions of folded and unfolded RNA as the decrease in ellipticity is caused by duplex unfolding.

3.7.2 CD Spectroscopy of Proteins

Analogue to the investigation of RNA conformations can CD spectroscopy be utilized to assess the secondary structure composition and folding properties of a protein. This is possible because the ellipticity of the polypeptide backbone amides changes with their array alignment resulting in protein secondary structure elements like α-helices, β-sheets, or random coils to exhibited very distinct signature spectra (Figure 13). The overall protein thus gives rise to a CD spectrum that is a linear combination of its individual secondary structure elements [136–138].

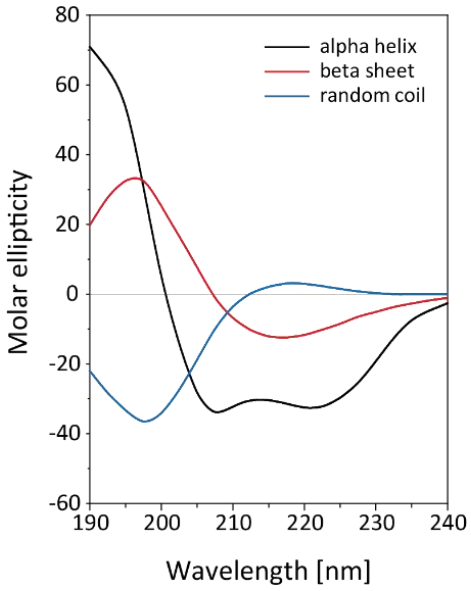

Figure 13. Model CD spectra of protein alpha helix, beta sheet and random coil. Modified after Reed et al. 1997 [136].

All measurements were performed in an analogous manner to the RNA measurements on a Jasco J-810 spectropolarimeter (JASCO) equipped with a PTC-4235 L Peltier thermostatic cell holder (JASCO) using 1 or 2 mm (QS) quarts glass high performance macro-cuvettes (Hellma) with 200 or 400 µL of a 10 µM protein sample in ddH_2O. A constant purging of

the measurement chamber with 2.0-2.2 L/min N_2 gas prevented the formation of ozone during operation. CD spectrum measurements were blanked for baseline correction beforehand with ddH_2O. Unless stated otherwise, all CD spectra of proteins were recorded with the measurement parameters listed in Table 8.

Table 8. Measurement parameter set for CD spectra of protein constructs.

Parameter	Setting
Wavelength range	190-300 nm
Scanning speed	50 nm/min
Scanning mode	continuous
Data pitch (acquisition interval)	0.2 nm
Sensitivity	standard
D.I.T. (digital integration time)	1 s
Band width	1 nm
Accumulations	3

3.8 Analytical Size Exclusion Chromatography

Analytical size exclusion chromatography was performed at 4°C on a ÄKTApurifier™ FPLC system (GE Healthcare) equipped with a Superdex 200 Increase 10/300 GL column (GE Healthcare) equilibrated with NMR buffer or high-salt NMR buffer. Calibration was performed using the Gel Filtration Markers Kit for Protein Molecular Weights 12-200 kDa (Sigma). 100 µL of 10, 20 or 100 µM protein in the correspondent buffer were applied to the column with a glass syringe (Hamilton) and eluted over 1.25 CV (30 mL) at 0.65 mL/min.

3.9 ATP Hydrolysis Assay

The ATPase activity of RNA helicase RhlB was determined spectrophotometrically using the Molecular Probes™ EnzChek™ Phosphate Assay Kit (Thermo Fisher Scientific). The underlying method, which was originally described by Webb[139], is based on the conversion of 2-amino-6-mercapto-7-methylpurine riboside (MESG) to 2-amino-6-mercapto-7-methylpurine and ribose 1-phosphate in the presence of orthophosphate (see Figure 14). This reaction, catalysed by the enzyme purine nucleoside phosphorylase (PNP), is accompanied by a change in absorption maximum from 330 nm to 360 nm that allows the quantification of inorganic phosphate (P_i) released during ATP hydrolysis.

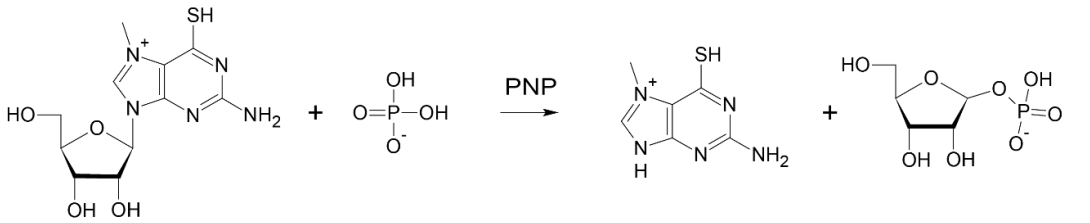

Figure 14. Enzymatic conversion of 2-amino-6-mercapto-7-methylpurine riboside (MESG) with orthophosphate to ribose 1-phosphate and 2-amino-6-mercapto-7-methylpurine by purine nucleoside phosphorylase (PNP).

The original experimental setup was scaled down from 1 mL to 200 µL reaction volume and adapted from 1 mL cuvettes to 96-well microplates to increase output and reproducibility. Reagents were prepared according to the manufacturer's protocols and a phosphate standard curve was provided by measuring samples with variable known concentrations of free phosphate (0-30 nmol).

Assays were performed in triplicates with 6 µM RNA substrate, 2.4 µM protein (RhlB ± RNase E), 0.2 mM MESG, 1 U/mL PNP, 0.2 mM ATP in 111 mM KCl, 68.5 mM Tris, 1 mM $MgCl_2$, 0.1 mM sodium azide at pH 8.1. All reaction components except for ATP were mixed and preincubated at 22°C for 10 min, before the reaction was started by addition of 40 µL ATP to the final concentration of 0.2 mM or addition of 40 µL buffer in case of the negative control. The absorption at 360 nm was monitored for 300 s with a 7 s data point interval on an Infinite® 200 PRO microplate reader (Tecan). The data points were corrected for baseline and negative control, converted into P_i concentrations using the phosphate standard curve and fit linearly from 120-220 s. Finally, the resulting initial rate constants were converted into [mol P_i /min/ mol helicase].

3.10 NMR Spectroscopy

3.10.1 Sample Preparation and Experimental Setup

Unless stated otherwise, all NMR samples containing RNA either prepared in "pH 8 NMR buffer" (150 mM KCl, 25 mM Tris, pH 8.3) or in "pH 6 NMR buffer" (150 mM KCl, 25 mM BisTris, pH 6.5) supplemented with 4.5 mM $MgCl_2$, 100 µM sodium trimethylsilylpropanesulfonate (DSS) as internal ^{1}H chemical shift standard, 6-10% (v/v) D_2O and, in case of protein experiments, 5 mM dithiothreitol (DTT). For ^{31}P NMR experiments additional 0.3-1 mM creatine phosphate (PCr) was added as internal ^{31}P chemical shift standard. ^{15}N and ^{13}C chemical shifts were referenced indirectly using the following chemical shift reference ratios: ^{13}C = 0.25144953 and ^{15}N = 0.101329118[140]. While all RNA samples were prepared in 5 mm Shigemi® NMR micro tubes (SHIGEMI Co.) with a sample volume of 280 µL, 3 mm NMR tubes with a total volume of 180 µL were

utilized for measurements involving proteins to avoid foam formation caused by the plunger. For experiments in D_2O, the RNA NMR sample were lyophilized and resuspended in 280 µL 99.95% (v/v) D_2O twice for complete solvent exchange.

Unless stated otherwise, all NMR measurements with RNA were performed at 288 K. The NMR spectrometers (Bruker) and probe heads used are listed in the table below. Data were processed with TopSpin (Bruker BioSpin) versions 2.1, 3.2 and 4.0.6 and further analysis of 2D experiments was performed with SPARKY version 3.114 (T.D. Goddard, D.G. Kneller, UCSF) and NMRFAM-SPARKY[141].

Table 9. Utilized Bruker Avance (AV) spectrometer and equipped probe heads.

Spectrometer	Probe head
AV III 600 HD	Prodigy 5 mm TCI $^{1}H/^{19}F$ [^{13}C, ^{15}N] Z-GRD
AV II 600	5 mm Cryo TCI ^{1}H [^{13}C, ^{15}N] Z-GRD
AV III 600 (599er)	5 mm Cryo TCI ^{1}H [^{13}C, ^{15}N] Z-GRD
AV 600	5 mm Cryo TCI ^{1}H [^{13}C, ^{31}P] Z-GRD
AV III 700 HD	5 mm Cryo QCI ^{1}H [^{13}C, ^{15}N, ^{31}P] Z-GRD
AV III 800 HD	5 mm Cryo TCI ^{1}H [^{13}C, ^{15}N] Z-GRD
AV III 800	5 mm Cryo TXO ^{13}C [^{1}H, ^{15}N] Z-GRD
AV NEO 900	5 mm Cryo TXI ^{1}H [^{13}C, ^{15}N] Z-GRD
AV III 950	5 mm Cryo TCI ^{1}H [^{13}C, ^{15}N] Z-GRD

3.10.2 NMR Characterization and Assignment of RNAs

In a one-dimensional ^{1}H NMR spectrum of RNA the focus mostly lies on the imino protons of the nucleobases guanine and uracil. Not only are those resonances easily identified because of their characteristic low field chemical shift, the imino protons only give rise to a detectable signal when they are protected from solvent exchange due to stable hydrogen-bond formation. This way, the number of imino proton resonances gives a direct read-out over the number of stable base pairs and conformations of the RNA. Moreover, their distinct chemical shift does not overlap with protein resonances, which makes them the ideal reporter on RNA-protein interaction studies.

The assignment of the imino protons is usually aided by ^{15}N heteronuclear single quantum coherence (HSQC) and ^{1}H nuclear Overhauser effect and exchange spectroscopy (NOESY) experiments. In the ^{15}N HSQC, imino (and amino) protons from the nucleobases can be easily distinguished based on their nitrogen chemical shift. In a 2D NOESY experiment the nuclear Overhauser effect is then used to gain further insight into the spatial relationship of individual protons. Since this cross-relaxation effect is distance dependent, only protons that are closer than 5 Å will have a NOE-contact and therefore show a cross-peak in the spectrum. Those correlations allow the sequential assignment of imino protons along and across strands and give useful information on inter-nucleotide connections and sequentially distant protons.

An assigned set of imino protons is usually sufficient for a secondary structure characterization, but depending on the aim of the project, it might also want to identify aromatic proton resonances of the nucleobases and protons in the sugar moiety. For both proton groups a ^{13}C HSQC is the go-to experiment, since all those covalently bound protons can be detected there, independent of solvent exchange. Figure 15 shows the defined regions in which the individual protons resonate. For the aromatic protons, the adenine C2H2 can easily be distinguished from C6H6 and C8H8, which have a partial overlap. For the sugar moiety especially C1'H1' are easily recognizable because of their exceptionally low field shifted resonances.

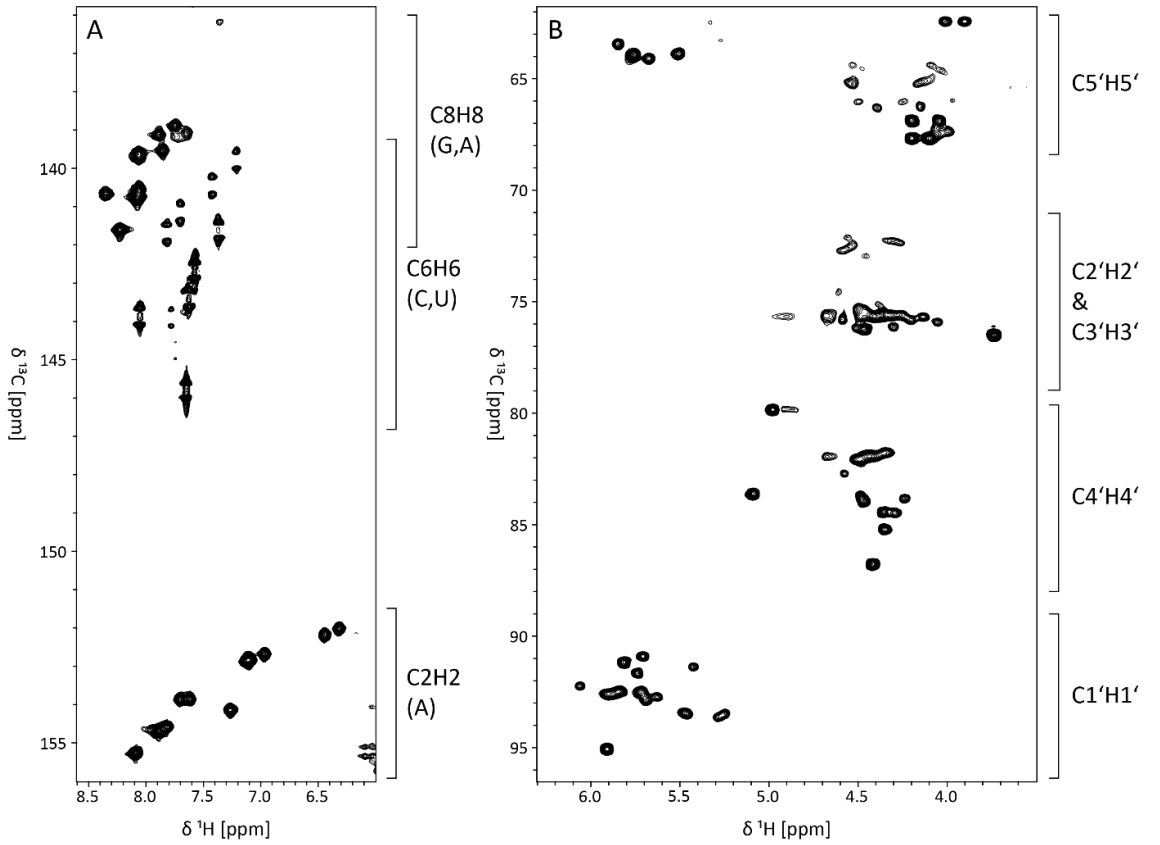

Figure 15. Exemplary ^{13}C HSQC spectra of RNA duplex J2Δ8-M2, illustrating ^{1}H and ^{13}C chemical shift ranges of aromatic nucleobase and sugar resonances.

The assignment of aromatic and sugar resonances is primarily performed within a 2D NOESY (preferably on a deuterated sample). Here, a sequential "walk" from 5' to 3' end using NOE contacts between H6/H8 and H1'/H2' can be used to link and identify the corresponding nucleotides. In case of incomplete assignments, the following additional NMR experiments can be performed to correlate missing signals: HCN triple-resonance experiment for assignment of intramolecular H6/H8-H1' connectivities, TROSY relayed HCCH-COSY experiment for correlating adenine H2 with H8 and ^{13}C NOESY-HSQC. For a full assignment of all sugar resonances, a HCCH-TOCSY would be recommended.

One-dimensional (1D) ^{1}H experiments were performed with a jump-return-echo water suppression scheme ("hs11echo", "hs11echo.dec") [142]. For the measurements 2048 to 4096 points were recorded with 64-256 scans, a spectral width of 25 ppm and a transmitter frequency offset of 4.7 ppm for ^{1}H (optimized to resonance frequency of water) and 153.5 ppm for ^{15}N (in case of ^{15}N labelled RNA samples). The binomial delay (D19) was set to 20-45 µs to maximize excitation on imino proton resonances. In case of ^{15}N labelled RNA a GARP sequence was utilized for decoupling during acquisition time[143]. The spectra were usually fourier transformed with 4096 points and a quadratic sine bell function with a sine bell shift (SSB) of 2 was used as a window function. Spectra were then phased, baseline corrected, and calibrated to the chemical shift of DSS.

For ***$^{1}H,^{1}H$-NOESY*** experiments focusing on imino proton resonance assignment a pulse sequence with jump-return-echo water suppression ("noesy11echo2013.ric") was used and spectra were usually recorded with 266 scans, 256 points in the indirect and 4096 points in the direct dimension[142]. The spectral width was set to 25 ppm in the direct dimension and 16 ppm in the indirect dimension. The experiment also included a frequency jump from 7 to 4.7 ppm in der indirect dimension. The ^{1}H transmitter frequency offset (O1p) was optimized to the resonance frequency of water (4.7 ppm), the mixing time was set to 300 ms, relaxation delay (D1) to 2 s and binomial delay D19 to 30 µs. In case of ^{1}H-^{1}H-NOESY experiments with focus on aromatic and sugar resonance assignment a phase sensitive NOESY sequence with gradient pulses during mixing time ("noesygpphpr") was utilized on a RNA sample in D_2O[144]. Die experiment was run with 192 scans, and 2048 points with a spectral width of 10 ppm were recorded in the direct dimension and 768 points with a spectral width of 10 ppm in the indirect dimension. O1p was set to 4.7 ppm, relaxation delay D1 to 1.3 s and the delay for homospoil/gradient recovery (D16) to 0.2 ms. All NOESY spectra were usually processed with 4096 and 1024 points in the direct and indirect dimension, respectively, using a quadratic sine bell function with SSB of 2.2 to 3. Phase correction and automated baseline correction were performed as well as chemical shift calibration to DSS resonance.

$^{1}H,^{15}N$-HSQC experiments were performed using a phase sensitive pulse sequence with a WATERGATE water suppression scheme ("fhsqcf3gpph")[145]. Depending on the sample concentration, 4-800 scans were measured with 2048-4096 points in the direct dimension and 256 points in the indirect dimension. Spectral widths were set to 25 ppm in the direct dimension and 34-35 ppm in the indirect dimension. The transmitter frequency offsets were set to the water resonance frequency (4.7 ppm) for ^{1}H, 101 ppm for ^{13}C and 153.5 ppm for ^{15}N. The experiments were run with a J-coupling constant for the INEPT transfer (CNST 4) of 110 Hz, a relaxation delay (D1) of 1 s and binomial water suppression delay (D 19) of 50-66 µs. All spectra were processed with 2048-4096 points in the direct dimension and 256-1048 points in the indirect dimension using a quadratic sine bell function with SSB of 2. Phase correction and automated baseline correction were performed, and the chemical shifts were calibrated to the DSS resonance frequency.

For ***^{1}H,^{13}C-HSQC*** measurements both *real-time* and *constant-time* pulse sequences were exploited. While *constant-time* experiments ("hsqcctetgpsp.2") provided an increased spectral resolution through impeded homonuclear ^{13}C coupling[146], the *real-time* versions ("hsqcetgpsisp", "hsqcetgpsisp2") exhibited an enhanced signal-to-noise ratio[147]. Spectra were recorded separately for sugar and aromatic resonances to reduce the spectral width. The experiments were performed with 52 scans, 2048x128 points and a spectral width of 15 and 20 ppm in direct and indirect dimension for aromatics and 15 and 40 ppm for sugars. The ^{1}H/^{13}C/^{15}N transmitter frequency offsets were to 4.7/143.5/190 ppm for aromatic resonances and 4.7/76/150 ppm for sugar resonances. J-coupling constant for INEPT transfer was set to 180 Hz for aromatics and 145 Hz for sugars. The experiments were performed with a relaxation delay of 1.5 s and a GARP4 sequence for decoupling during acquisition. Spectra were usually processed with 2048 points in the direct and 1024 points in the indirect dimension using a quadratic sine bell function with SSB of 2. Phase correction and automated baseline correction (in ^{1}H dimension) were performed, and the chemical shifts were calibrated to the DSS resonance frequency.

2D versions of the ***HCN*** triple resonance experiment ("na_hcnetgpsisp3d") were performed with 384 scans and 1024 points in the ^{1}H dimension, 256 points in the ^{15}N dimension and 1 point in the ^{13}C dimension[148]. The corresponding spectral widths for ^{1}H/^{13}C/^{15}N dimensions were 9 ppm, 16 ppm and 37 ppm, respectively. The experiments were performed with transmitter frequency offsets of 4.7 ppm, 115 ppm and 159.5 ppm for ^{1}H, ^{13}C and ^{15}N, respectively. The relaxation delay D1 was set to 1.1 s and ^{13}C and ^{15}N broadband decoupling during acquisition was accomplished with GARP sequences. The ^{1}H ^{15}N plane was fourier transformed with 2028 points in the ^{1}H and 512 points in the ^{15}N dimension using a quadratic sine bell function (SSB 2). Phase correction and automated baseline correction were performed.

The phase-sensitive 3D double-resonance experiment **NOESY-HSQC** ("noesyhsqcetgpsisp3d") was performed to obtain ^{13}C-edited NOESY spectra of ^{13}C-labeled RNA samples from which homonuclear ^{1}H-^{1}H NOEs can be assigned even in overcrowded regions and nucleobase resonances from a ^{13}C HSQC plane can be correlated with the respective ^{1}H^{1}H NOESY plane[147,149]. Experiments were recorded with 2024, 64 and 256 points for the ^{1}H/^{13}C/^{1}H dimensions, respectively, with corresponding spectral widths of 15 ppm, 20 ppm and 10 ppm. The ^{1}H/^{13}C/^{1}H transmitter frequency offsets were to 4.7/143.5/4.7 ppm. The relaxation delay D1 was set to 1.3 s. Spectra were processed with 2048 points in the direct ^{1}H dimension and 256 points in the indirect ^{1}H dimension, for ^{13}C dimension 128 points were chosen. Quadratic sine bell function (SBB 2) was used. Phase correction and automated baseline correction were performed.

2D versions of the **TROSY relayed HCCH-COSY** experiments ("na_trhcchco3d") were performed to specifically correlate H2/H8 resonances[150]. Spectra were recorded with 2048, 1 and 512 points as well as spectral widths of 26 ppm, 60 ppm and 60 ppm in the ^{1}H/^{13}C/^{13}C dimensions, respectively. The transmitter frequency offsets were 4.7 ppm for the ^{1}H and 140 ppm for the ^{13}C dimensions. The relaxation delay D1 was set to 1.5 s.

Spectra were processed with 1024, 1024 and 256 points for the $^{1}H/^{13}C/^{13}C$ dimensions, respectively, using a Quadratic sine bell function (SBB 2). Phase correction and automated baseline correction were performed.

One-dimensional (1D) ^{31}P experiments were acquired with standard pulse sequences including ^{1}H (and ^{13}C) broad band decoupling during acquisition ("zgig", "zgigf2f3", "zgigf2igf3"). Experiments were usually set up with 1024-8192 points, a spectral width of 10-60 ppm, a relaxation delay of 2 s and 128-4094 scans. Transmitter frequencies were set to 0 ppm for ^{31}P, and to 4.7 ppm (^{1}H) and 72 ppm (^{13}C) for decoupling. Spectra were fourier transformed with 16384 points and an exponential function using line broadening of 5 Hz. After phase correction and automatic and manual baseline correction, the chemical shift was calibrated to the PCr reference frequency.

3.10.3 NMR Characterization of Proteins

For ***one-dimensional (1D) ^{1}H*** experiments on proteins a pulse sequence with water suppression by excitation sculpting with gradients was used ("zgesgp")[151]. Measurements were mainly recorded with 4094 points, 256 scans, a spectral width of 16 ppm, transmitter frequency offsets of 4.7 ppm (^{1}H) and 117 ppm (^{15}N) and a relaxation delay of 1 s. The spectra were processed with at least 4096 points and either a quadratic sine bell function (SSB 2) or an exponential window function with a line broadening (LB) of 0.3-5 Hz. Phase correction and automatic baseline correction were performed and if necessary, the spectra were calibrated to the chemical shift of DSS.

The ***^{15}N-BEST-TROSY*** ("b_trosyetf3gpsi.3") was a frequently used 2D experiment recorded for protein characterization[152,153]. For the band-selective excitation of amide protons, PC9, REBURP and E-BURP2 shaped pulses were utilized with an offset frequency of 8.3 ppm and a bandwidth of 5 ppm. The experiments were typically acquired with 1024-2048 points in the direct and 170-256 points in the indirect dimension with 92 scans. The spectral width was set to 12 and 35 ppm in direct and indirect dimension, respectively, the transmitter frequency offset was set to 4.7ppm, 101 ppm and 117 ppm for ^{1}H, ^{13}C and ^{15}N, and the relaxation delay was set to 0.3 s. Spectra were usually fourier transformed with 2048 and 512 points in direct and indirect dimension, using a quadratic sine bell function with a sine bell shift of 2.5. The spectra were then phase corrected and automatically baseline corrected in the proton dimension before being calibrated to the chemical shift of DSS. To correct for the TROSY shift, the spectrum reference frequency was changed by -47 Hz in the proton dimension and +47 Hz in the nitrogen dimension.

Furthermore, ***$^{1}H,^{15}N$-HSQC*** ("fhsqcf3gpph") experiments were measured using the pulse sequence previously described for RNA measurements in section 3.10.2. For proteins the experiments were usually recorded with 2048 and 256 points in direct and indirect dimension, respectively, with 24 scans and transmitter frequency offsets of 4.7 ppm, 101 ppm and 117 ppm for ^{1}H, ^{13}C and ^{15}N. The relaxation delay was set to 1 s, J-coupling

constant for the INEPT transfer was 90 Hz and the delay for binomial water suppression was set to 62.4 µs. Recorded spectra were processed with 2048 x 256 points as described in section 0.

CON** and **(HACA)CON experiments were performed for ^{13}C detected correlation between C_o and N, using two similar pulse sequences ("c_con_iasq", "c_hcacon_ia")[154,155]. While the CON pulse sequence excites all C_O carbon atoms including side chain carboxamide groups of glutamine and asparagine and transfers the magnetization to N, (HACA)CON begins excitation on H_α atoms before transferring the magnetization to $C_\alpha \rightarrow C_O \rightarrow N$, thereby omitting the sidechain resonances and yielding a spectrum with a higher sensitivity. Both experiments were recorded with 16-32 scans, 1024 points in ^{13}C dimension and 128-512 points in ^{15}N dimension and a spectral width of 33 ppm for ^{13}C and 47 ppm for ^{15}N. The transmitter frequency offsets were set to 173.5 ppm, 123 ppm and 4.7 ppm for ^{13}C, ^{15}N and ^{1}H, respectively. The spectra were virtually decoupled using an in-phase/anti-phase (IPAP) scheme, phase corrected, automatically baseline corrected and referenced to the chemical shift of DSS based on a recorded ^{1}H 1D spectrum.

3.10.4 Real-time Mixing Experiments

Real-time NMR mixing experiments are a powerful tool to investigate chemical processes on a second to hour time scale, including secondary structure transitions and certain chemical reactions. In this project it was utilized to investigate the ATP hydrolysis of RhlB.

All experiments were performed on a AV III 700 HD spectrometer (Bruker) at 288 K utilizing the rapid mixing setup adapted from Mok *et al.*[156]. In this setup, a 300 µL volume of 100 µM unlabelled RNA, 400 µM RhlB (+/- 400 µM RNase E), 1 mM PCr, 100 µM DSS and 6% D_2O were prepared in NMR buffer in a 5 mm Shigemi NMR tube and inserted with a glass capillary containing 40 µL of injection solution (25.5 mM ATP in NMR buffer with 6% D_2O) that would yield a final ATP concentration of 3 mM ATP after mixing. The glass capillary was connected to a pneumatic injection syringe. For the mixing experiment a consecutive series of single scan ^{31}P 1D spectra with 8192 points and a spectral width of 30 ppm were recorded as a pseudo 2D ("rt_zgig_2D.hz"; Appendix Chapter 9), with the injection being triggered after 128 scans. The mixing experiments were recorded with 25600 scans over a period of 14 h and initially processed with 32k points. Without sacrificing significant time resolution, individual 1D spectra were combined to increase the signal-to-noise ratio (minimum 2024 points). Pseudo 2D spectra were then automatically baseline corrected in the direct dimension and phased by transferring phase correction of 1D slice onto 2D. Peak integrals over time were calculated within TopSpin and the kinetic traces were fitted in SigmaPlot v.12.5 with the following double exponential function:

$$I = I_0 + a \cdot e^{(-k_1 \cdot t)} + b \cdot e^{(-k_2 \cdot t)} \qquad \textit{Equation 7}$$

With I being the peak integral, I_0 the starting intensity, k_1 and k_2 the rate constants and a and b being amplitude constants.

3.10.5 Titration Experiments

To monitor the structural effect of RhlB binding to RNA substrates and to determine the K_D, titration experiments have been performed in the following way: 50 or 100 µM unlabelled RNA were prepared in a 3 mm NMR tube as described in section 3.10.1 and the protein of interest was added in titration steps of 0.25, 0.5, 1,2 and 4 equivalents of protein over RNA. For each titration step (including the initial NMR sample without protein) a standard ^{1}H 1D spectrum with jump-return-echo water suppression ("hs11echo") was recorded with 4094 points, 512 scans, a spectral width of 25 ppm and a binomial delay of 45 µs. For determination of apparent K_D, peak intensities of non-overlapping imino proton resonances were extracted from TopSpin, dilution corrected against DSS reference signal and then the normalized reciprocal peak intensities were plotted against the protein concentration and fitted globally based on a single side ligand binding equilibrium function[157].

Chapter 4 Design, Synthesis, and Spectroscopic Analysis of RNA Substrates for RhlB

4.1 RNA Preparation and Purification

Although all RNA required constructs could be purchased in unlabelled form, the high concentration demands of NMR spectroscopy as well as the production of ^{13}C, ^{15}N labelled samples required additional in-house *in vitro* synthesis for most of the RNA substrates. Therefore, the ordered plasmids of RNA constructs M2, J2, J2h, J2Δ8 and J2Δ16 were transformed into *E. coli* DH5α cells, amplified and isolated. The circular plasmids were linearized using *SmaI* or *EcoRI* restriction enzyme, respectively, and the successful digestion confirmed via analytical agarose gel electrophoresis, as shown in Figure 16. The undigested plasmids of the pUC57 show the major band of the compact supercoiled structure between 2 and 3 kBp and one to two faint bands in the range between 6 and 10 kBp representing the plasmid in the relaxed circular or nicked conformation. For each pUC57 construct, restriction digestion produced a single linearized DNA band running slightly below 3 kBp, matching the corresponding vector lengths (2.21 kBp + insert). In case of the smaller pEX-A vector (2.45 kBp + insert) both the supercoiled and the linear plasmid run proportionately lower in the gel.

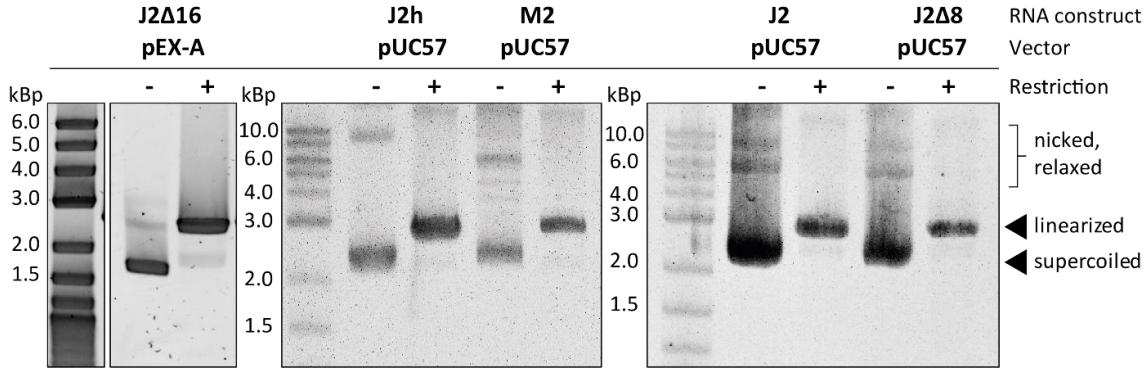

Figure 16. Agarose gel electrophoresis of plasmid linearization via restriction digestion for constructs M2, J2h, J2, J2Δ8 and J2Δ16. Linearized and supercoiled plasmids are marked with arrows.

The linearized plasmids were then purified and used as DNA templates for *in vitro* transcription. As described in 3.4.1, extensive test transcriptions were performed for each construct to optimize the reaction conditions aiming at high yields of cleaved RNA products while minimizing the formation of undesired side products. Optimal conditions were determined via denaturing RNA-PAGE based on band intensity. Figure 17 shows the denaturing RNA-PAGE gels of the optimized conditions for each construct. Since each transcript comprises the RNA of interest and both 5' hammerhead and 3' HDV ribozymes, a transcription under optimal reaction conditions (especially $Mg(OAc)_2$ concentration) aims for a complete self-cleavage of the ribozymes and results in three separate RNA bands. The significantly larger DNA template, which does not migrate into the gel at the chosen electrophoretic conditions, is stained in the gel pocket. Apart from J2, which even

after optimization still exhibited a small fraction of partially cleaved product (band highlighted with an asterisk in Figure 17), all transcribed constructs could be optimized effectively.

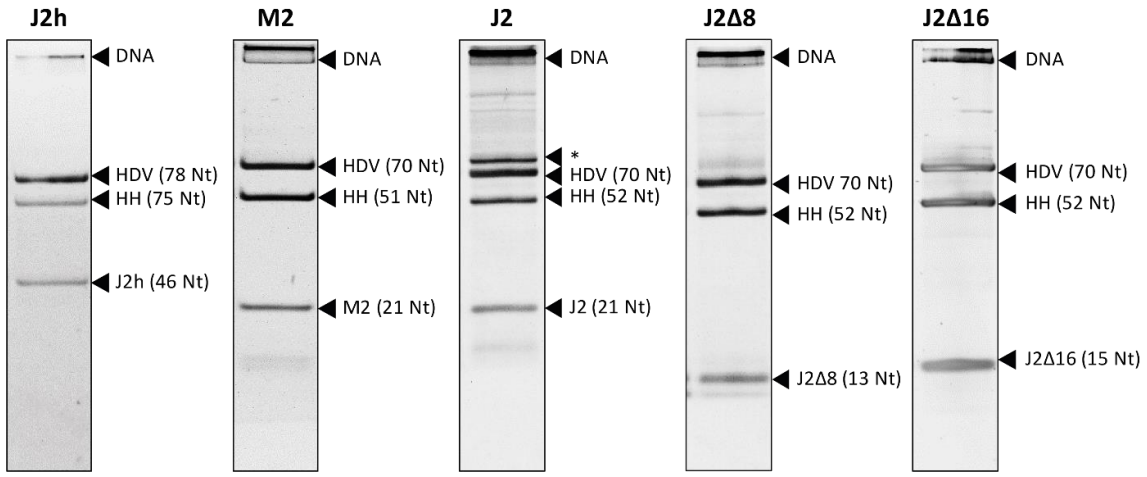

Figure 17. Denaturing RNA-PAGE of analytical transcriptions under optimized reaction conditions for RNA constructs J2h, M2, J2, J2Δ8 and J2Δ16. Gel lanes show the fully cleaved transcripts of 3' HDV ribozyme, 5' HH ribozyme, the RNA of interest as well as the DNA template (remains in gel pocket). Notice that the total length of the ribozymes varies depending on the RNA construct. The thin band in the transcription of J2 marked with an asterisk is partially cleaved product of either J2+HDV (91 Nt) or J2+HH (73 Nt).*

The optimized transcription conditions are summarized in Table 10. With the exception of J2h, all RNA constructs required 30-40 mM $Mg(OAc)_2$ and 75-100 ng/µL DNA template for optimal cleavage. In case of J2h 70 mM $Mg(OAc)_2$ and 150 ng/µL DNA yielded the best results. 10% DMSO was only required to for J2 and M2 to reduce formation of unwanted side products. For the shortest two constructs J2Δ8 and J2Δ16 a slightly higher rNTPs concentration of 10 mM was necessary to achieve sufficient product formation, while 4mM rNTPs were effectual for the other constructs.

Table 10. Optimized transcription reactions for individual RNA constructs. Only reaction components highlighted in blue were varied.

Reaction component	J2h	M2	J2	J2Δ8	J2Δ16
Tris/glutamate pH 8.1	100 mM	100 mM	100 mM	100 mM	100 mM
DTT	20 mM	20 mM	20 mM	20 mM	20 mM
Spermidine	2 mM	2 mM	2 mM	2 mM	2 mM
$Mg(OAc)_2$	70 mM	35 mM	40 mM	30 mM	40 mM
rNTPs [a]	4 mM	4 mM	4 mM	10 mM	10 mM
DNA template	150 ng/µL	100 ng/µL	100 ng/µL	75 ng/µL	100 ng/µL
DMSO	0	10 %	10 %	0	0
T7 polymerase (P266L)	70 µg/mL	70 µg/mL	70 µg/mL	70 µg/mL	70 µg/mL

[a] Ratio of individual rNTPs are adjusted to NTP composition of transcript.

Preparative scale transcriptions (10-20 mL) were then performed with the optimized reaction conditions. Each RNA construct was successively purified using DEAE anion-exchange chromatography and *rp*-HPLC or preparative PAGE as described in sections

3.4.2 to 3.4.4. The corresponding gel analysis of DEAE elution fractions and HPLC-chromatograms are appended in appendix 8.3.

Despite promising test transcriptions did the large-scale synthesis of J2Δ16 fail repeatedly, as the newly transcribed RNA started to degrade during transcription and continued to degrade during the anion-exchange chromatography until J2Δ16 could not be detected anymore. A comprehensive trouble shooting was unsuccessful, resulting in the exclusive use of commercially available unlabelled RNA for further experiments. The RNA constructs J2Δ14 and M3 were only purchased and resuspended in H_2O according to the manufacturer's instructions. For both *in vitro* transcribed and purchased RNA constructs denaturing and native RNA-PAGEs confirmed the purity and folding homogeneity of the samples. All RNA constructs showed one distinct band on both gels and eventual dimer formation could be dismissed (Figure 18).

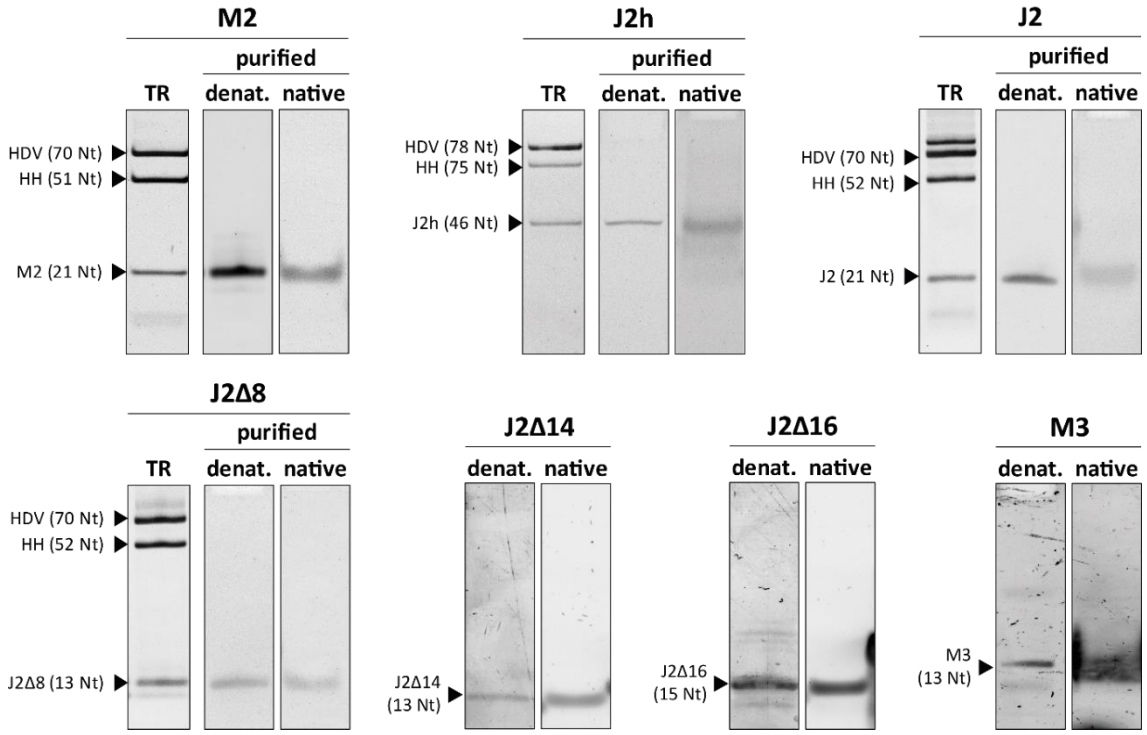

Figure 18. Denaturing and native PAGE for all utilized RNA constructs. For constructs that were also prepared via in vitro transcription, an exemplary transcription reaction (TR) is included. RNA synthesis products and their corresponding sizes are indicated by black arrows.

4.2 Duplex Formation and Stability

4.2.1 Duplex Formation Analysis by native RNA-PAGE

To be suitable RNA substrates for investigation of the helicase mechanism via NMR spectroscopy, the designed RNA single strands had to form homogenous heteroduplexes and adopt a single conformation. For an initial inspection of the annealing homogeneity, native RNA-PAGEs were performed with 30-50 pmol of the relevant RNA single strands both individually and loaded at a 1:1 stoichiometry. Figure 19 shows the results of these tests. Both a significant shift in the band height and an increased GelRed® staining intensity are indicative of a duplex formation, since the dye intercalates better into double stranded nucleic acids. This can be observed most notably for the duplexes J2Δ8-M2 and J2h-M2, but also the duplexes J2-M2, J2Δ16-M2 and J2Δ8-M3 exhibited a strong increase in staining intensity, which strongly suggested duplex formation. While the formed duplex usually runs higher on a gel due to a larger molecular weight, J2Δ14-M2 showed an intense band at the same height as the J2Δ14 single strand and a fainter band matching the height of M2, implying an incomplete duplex formation or unexpected secondary structure. A fainter band could also be observed for J2h-M2 above the main band. With the folding procedure described in section 3.4.5 this minor folding population could be further reduced.

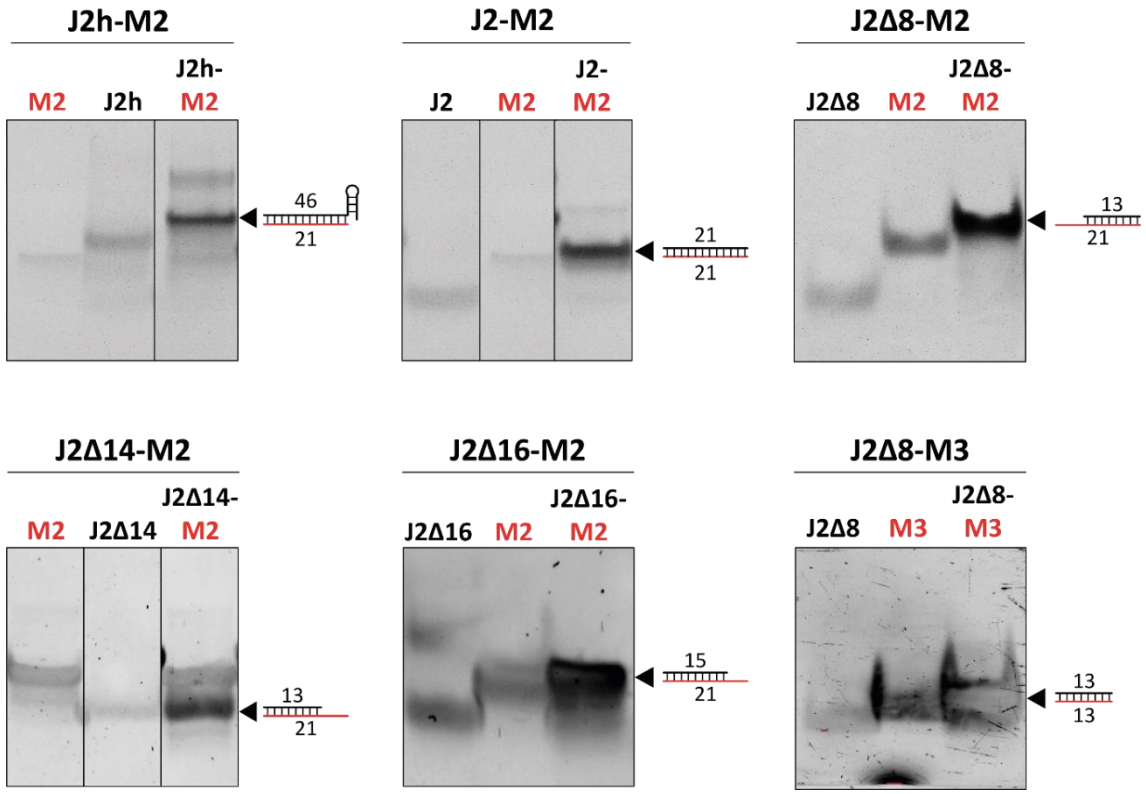

Figure 19. Native PAGE of individual RNA single strands as well as the corresponding 1:1 mix to verify duplex formation. Expected secondary structures of the formed duplexes are indicated schematically and the corresponding strands colour coded in red and black.

4.2.2 Duplex Stability and Melting Temperatures by CD Spectroscopy

To further analyse the thermal stability of the RNA duplexes, CD melting curves were recorded in a temperature interval from 5 to 95°C for all heteroduplex constructs. The temperature dependent molar ellipticity was converted into fractions of folded RNA and the curves are presented in Figure 21. All RNA duplexes apart from J2h-M2 were fitted with a sigmoidal function, from which the melting points were extracted. Since J2h-M2 has an additional five base pair hairpin that unfolds separately, a double sigmoidal regression function was used to reflect the unfolding behaviour and extract two melting points. Because the RNA buffer composition was altered and optimized over the course of the project, CD experiments were in part conducted at pH 6 and pH 8. To evaluate the stabilizing effect of Mg^{2+} ions on the duplexes and the influence on the melting temperatures experiments were performed in the presences and absence of 1.5 or 4.5 mM $MgCl_2$. The predicted and measured melting points are summarized in a bar diagram in Figure 20 and listed in detail in appendix chapter 8.3.

The melting temperatures of the RNA duplexes increased as expected with duplex length and stability from 37.5°C for duplex JΔ14-M2 up to 58.7°C for J2h-M2 in the absence of magnesium. The overall trend also follows the predicted minimum free energy for the individual RNA duplexes (see Table 1). While the predicted melting points were generally in agreement with the relative stability of the duplexes, the exact values did in most cases overestimate the measured melting temperature by about 4-6°C. This illustrates how, despite the advancement of nucleic acid prediction software, one should not rely exclusively on calculated thermodynamic properties but whenever possible include experimental data for verification. Most noticeable was the melting curve of J2h-M2, which did not resemble a two-state unfolding transition like the other constructs but showed two inflection points at 55.7°C and 58.7°C. The T_m prediction by Pairfold is limited to two strand annealing processes and therefore only provided the estimated melting temperature of the heteroduplex formation (64.3°C). Since both measured melting points are in that range, we cannot say with certainty, which one resembles the 21 bp heteroduplex and which one the 5 bp hairpin. Very stable short hairpins with a melting point beyond 50-60°C are not unheard of – the particularly stable tetraloops come to mind[36] - but it is nonetheless unexpected. It is on the other hand also possible that the unfolding transition of the hairpin was not captured in the curve fit and the secondary inflection point is indicative of residual inhomogeneity in the construct conformation. Another interesting finding was the notable increase in T_m from the blunt end 13 Nt duplex J2Δ8-M3 to the same duplex with 3' extension (J2Δ8-M2): Although all base pairs were identical, the stabilizing effect of the 3'-single strand extension increased the melting point about 6.5°C without $MgCl_2$ and 4.9°C with $MgCl_2$. According to Freier *et al.* this effect is caused by the geometry of A-form RNA. Interactions of the 3'-dangling nucleotides with the bases of the opposing strand allow the dangling ends to continue the RNA duplex geometry and provide a favourable ΔG° through additional base stacking[158,159]. The addition of $MgCl_2$ to the buffer visibly increased the melting

temperature for all measured RNA duplexes by 5-7 °C as the Mg^{2+} ions stabilize the duplex formation by non-specific interaction with the negatively charged phosphate backbone[160].

As later NMR ^1H 1D experiments revealed (see section 4.3.8) did the M2 single strand form a two base pair hairpin in absence of a counter strand. Attempts to determine the melting point of this secondary structure via CD spectroscopy weren't successful since the change in ellipticity that accompanied the unfolding was too small for a reliable fit. The melting temperature was therefore determined by measuring ^1H 1D NMR spectra at increasing temperatures and integrating the imino proton resonance signal of G3. A sigmoidal fit of the integrals (see appendix chapter 8.3) yielded a melting point of 26.7°C.

Based on the acquired information on the RNAs thermal stability, all further experiments were performed at 20°C or lower where the fraction of folded duplexes was larger than 95% for all constructs.

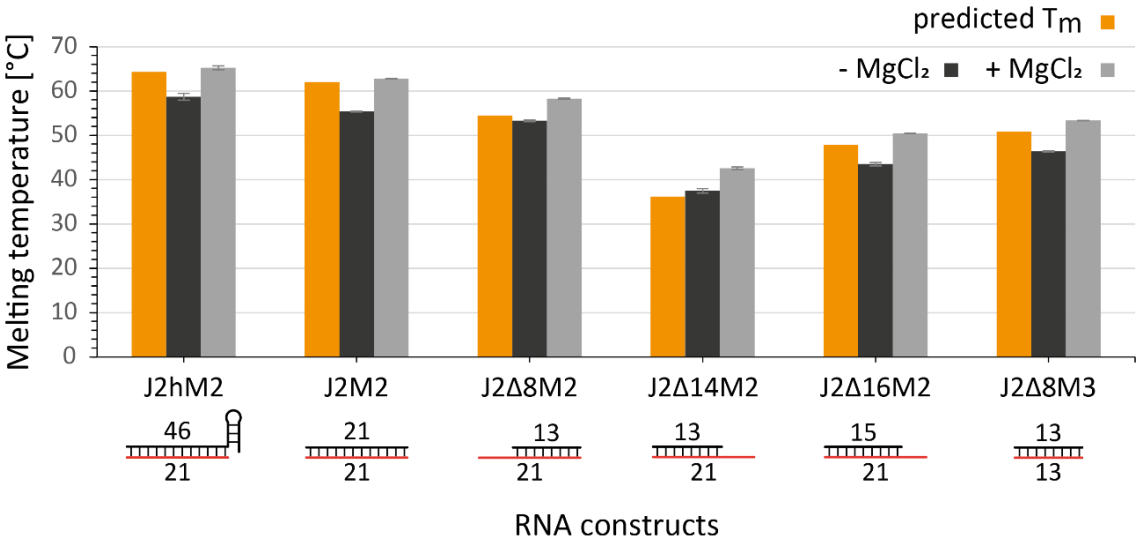

Figure 20. Predicted and measured melting points of all investigated RNA duplex constructs. J2h-M2 was measured in 150 mM KCl and 25 mM K_2HPO_4/KH_2PO_4 at pH 6.2 and all remaining constructs in 150 mM KCl and 25 mM Tris at pH 8.3. Notice that for the "+ $MgCl_2$" measurements J2h-M2 was supplemented with 1.5 mM $MgCl_2$, while for all other constructs 4.5 mM $MgCl_2$ was added. Errors are given as the mean standard deviation of the fit. Of the two melting points derived from double sigmoidal fit of the J2h-M2 melting curve, only the higher T_m was used in this bar diagram. Predicted T_m were calculated with Pairfold for an RNA concentration of 10 µM and an ionic strength of 150 mM.

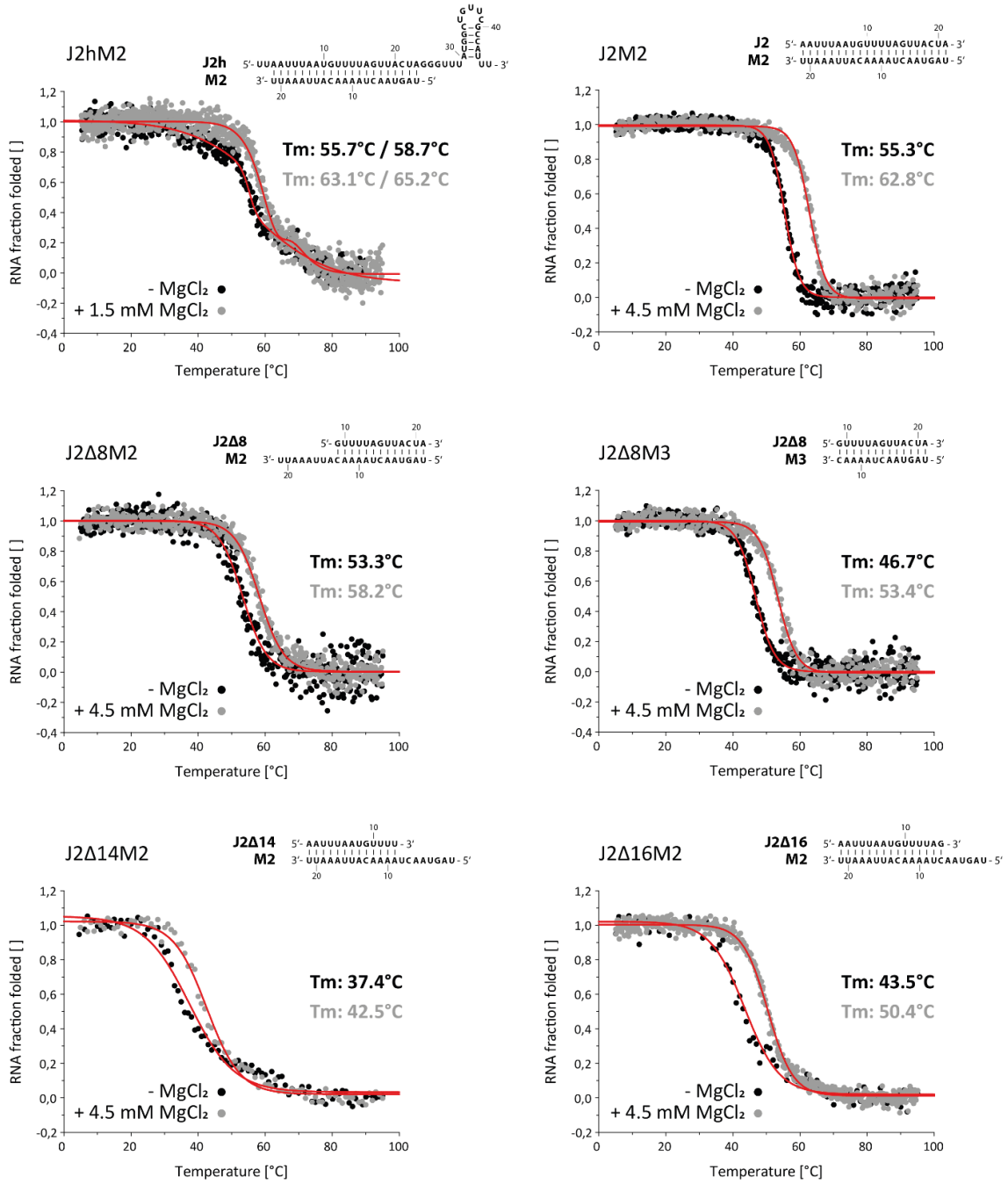

Figure 21. CD melting curves plotted as folded RNA fractions against temperature for RNA duplexes J2h-M2, J2-M2, J2Δ8-M2, J2Δ8-M3, J2Δ14-M2 and J2Δ16-M2. All measurements were acquired with 10 µM RNA in either 150 KCl and 25 mM Tris at pH 8.3 or in 150 mM KCl and 25 mM K_2HPO_4/KH_2PO_4 at pH 6.2 for duplex J2h-M2. Temperature curves were performed at λ_{max} of 258 to 268 nm (determined individually via CD spectrum for each construct). Grey and black data points represent measurements with and without $MgCl_2$, respectively, and sigmoidal and double sigmoidal regression curves are highlighted in red. Both measurements of J2Δ14-M2 and J2Δ16-M2 without $MgCl_2$ were performed with a D.I.T. of 4 s instead of 1 s, which reduced the signal-to-noise ratio compared to the other plots.

4.3 NMR Spectroscopic Characterization of RNA Constructs

The NMR spectroscopic analyses were performed to probe the secondary structure of the designed RNA duplexes as well as to evaluate their conformational homogeneity. Besides a general assessment of the suitability of the constructs for further NMR studies we aimed at a complete imino proton assignment of each construct. Those assignments were in part obtained at pH 6 and different buffer compositions and therefore had to be transferred to spectra measured at the final buffer conditions at pH 8.3 (see "NMR buffer" in section 3.10.1). For RNA constructs that showed promising results in interaction studies with RhlB we proceeded with the nucleobase assignment and partial sugar assignment in ^{13}C HSQC spectra using ^{13}C,^{15}N labelled RNA samples. The homo- and heteronuclear experiments used depended on the assignment strategy and are specified in the section of the corresponding RNA.

4.3.1 Optimization of NMR Conditions

In this project most NMR experiments conducted to study the interaction of RhlB with its RNA substrates were performed with an emphasis on analysing the conformation of the RNA. This provided the unique challenge of finding measurement conditions that are compatible with a variety of different constraints. Firstly, the protein required a buffer composition and experimental setup, at which it was folded, catalytically active and stable for a at least the duration of the measurement, ideally even longer. At the same time, the RNA substrates needed to not only be stable but their experimentally relevant NMR signals had to be detectable with the highest possible resolution and sensitivity while showing the best possible chemical shift dispersion. Thirdly, neither buffer components nor concentration limitations could affect or impede the helicase reaction mechanism itself. Since some of these requirements were mutually exclusive, several compromises had to be made, often while sacrificing quality of NMR spectra over experimental feasibility. The following section shows the most important considerations and results of this optimization process.

Temperature:

The most relevant effect of temperature in NMR spectra of RNA can be observed on exchangeable imino protons and to a lesser extend also on amino protons of the nucleobases. While an increasing temperature leads to a faster molecular tumbling and therefore to narrower peaks, it also increases the exchange rates of the protons with solvent water, causing peak broadening in less stable base pairs. This effect can be nicely examined in Figure 22 with duplex J2Δ8-M2, where the overall peak sharpness increased from 283 to 303 K while the resonance of close-to-terminal U20 broadened beyond detectability. To maximize the number of imino reporter signals available, the upper temperature limit was set to 293 K. In our case, CD melting curves already dictated a

measurement temperature of 293 K or lower for every RNA substrate to adopt a stable heteroduplex conformation. This unfortunately meant that RhlB's could not perform at the ideal reaction temperature of 37°C (310 K) for an *E. coli* protein. Some of the previous studies on RhlB's ATPase activity have also been conducted at 25°C, suggesting that a reduced reaction temperature does not impede with the overall mechanism, but only decreases the absolute reaction rates[17]. Relative rates acquired under the same temperature conditions should therefore still give viable insights into substrate preferences. All those results informed the decision to measure NMR experiments at 288 K.

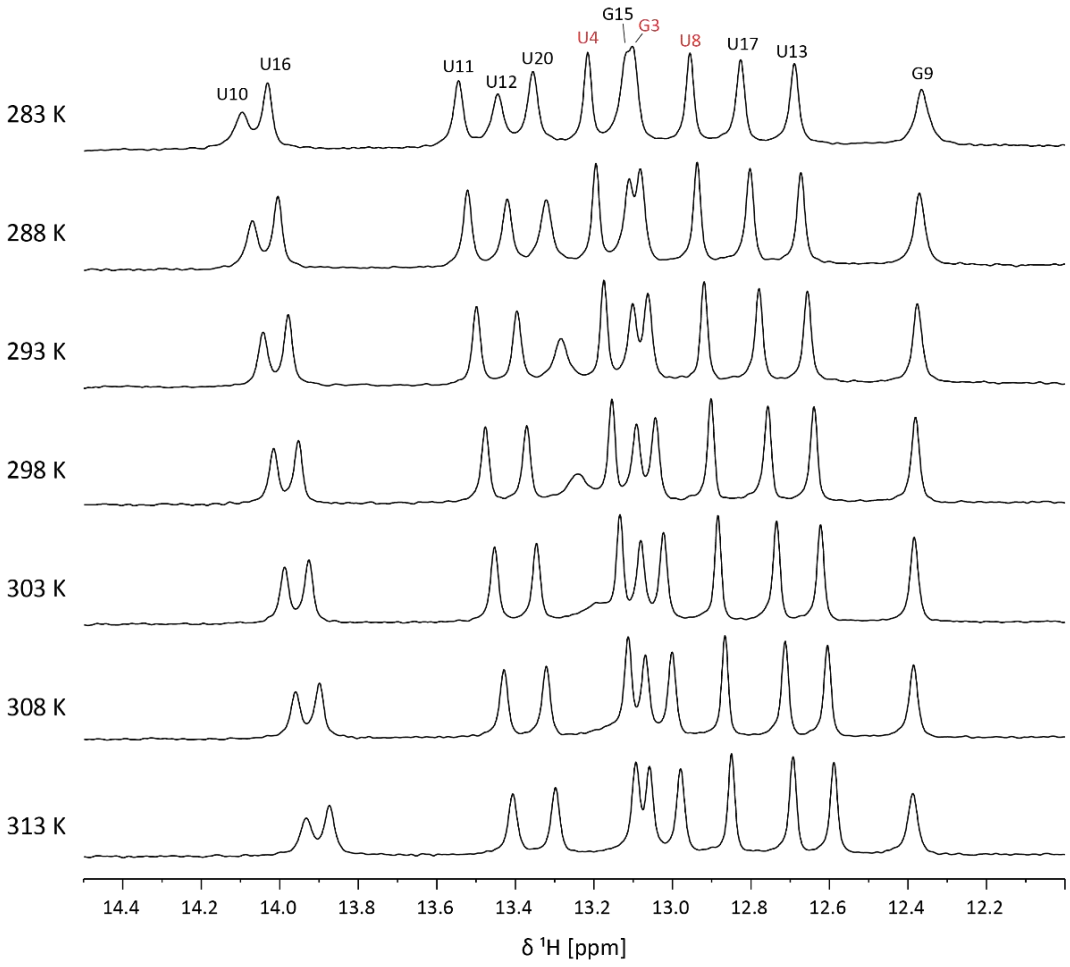

Figure 22. Imino regions of ^{1}H NMR spectra of RNA duplex J2Δ8-M2 (100 µM) in 150 mM KCl, 25 mM KH_2PO_4/K_2HPO_4, 1.5 mM $MgCl_2$ and 7% D_2O. Spectra were recorded at temperatures ranging from 283 to 313 K in increments of 5 degrees with 512 scans at 800 MHz. The relative peak shifts in the individual spectra were calibrated to the resonance of the buffer signal. Peaks annotated in black and red are assigned to J2Δ8 and M2 strand, respectively.

pH:

Because the hydrogen exchange with solvent water can be catalysed by both H^+ and OH^--ions the exchange rates of labile base protons like imino and amino protons of the nucleobases are also affected by the pH. While solvent exposed amino protons will be observable at pH around 7, imino protons signals are always broadened beyond detectability due to their high exchange rate unless the protons are protected in a stable hydrogen bond formation like a base pair[161,162]. The overall long-term stability of the

RNA substrates should also be considered here, as multiple publications highlight the increased rate of RNA degradation through auto-hydrolysis at alkaline pH[163,164]. Particularly single stranded RNA is susceptible for the phosphodiester-cleavage reaction[165]. To not compromise the RNA stability and for optimal imino proton resolution, typically a buffer with a pH around 6 is used for RNA samples.

The DEAD-Box helicase RhlB on the other hand was stably purified in this project at a pH 8.3 and attempts to change the pH to 7 or below as well as a purification at pH 6 resulted in substantial protein degradation and overall low stability. Consequently, it was tested whether NMR experiments with RNA could also be performed at a pH around 8. Figure 23 exemplifies this based on the pH titration of J2Δ16-M2 from a pH of 6.2 to 8.3. No chemical shift perturbations were detected. Although all peaks exhibited a visible line broadening towards the higher pH, all imino signals except for U20 were still clearly detectable even at pH 8.3. Similar to high temperatures are the increased exchange rates at higher pH only problematic for terminal and therefore less stable terminal base pairs like U20, which broadened beyond detectability between pH 7.2 and 7.7. It was concluded that further experiments of the protein RNA complex should be performed at pH 8.3, while RNA assignment experiments could be conducted at pH 6.2 and the assignment transferred to spectra at pH 8.3.

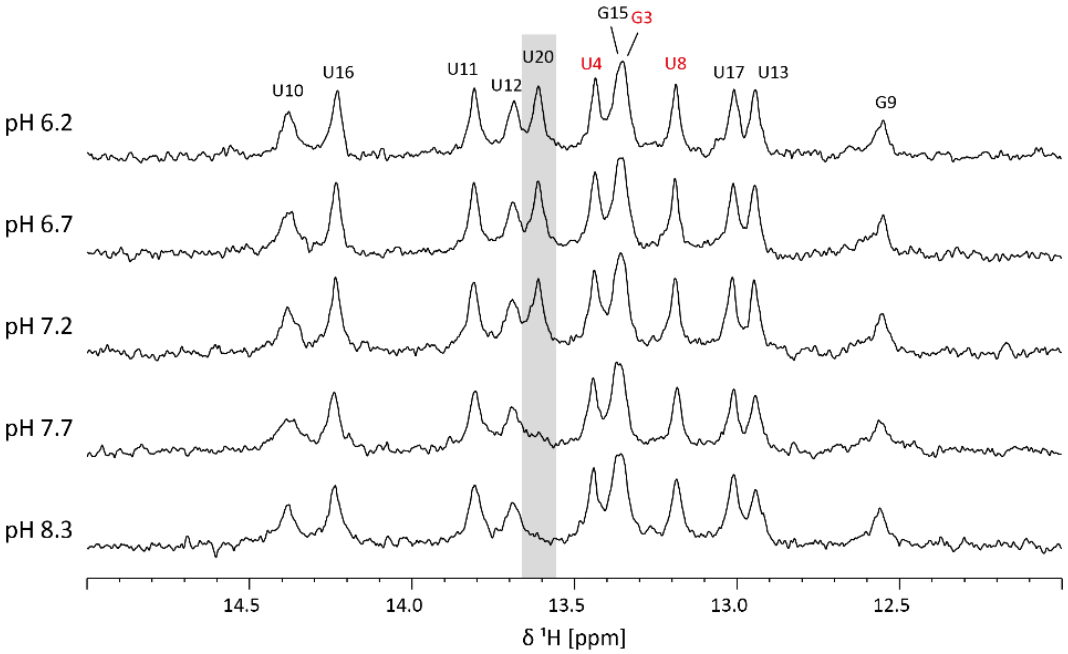

Figure 23. Imino regions of ^{1}H NMR spectra of RNA duplex J2Δ8-M2 (50 µM) in 150 mM KCl, 25 mM BisTris (pH 6.2-7.2) or Tris (pH 7.7-8.3), 50 µM DSS and 7% D_2O with pH ranging from 6.2 to 8.3 in increments of 0.5. Spectra were recorded at 278 K with 192 scans at 600 MHz. Peaks annotated in black and red are assigned to J2Δ8 and M2 strand, respectively. Resonances disappearing with pH increase are highlighted in grey.

Buffer system and ionic strength:

The inorganic buffering agent K_2HPO_4/KH_2PO_4 is conventionally used for NMR buffer of RNA as it does not give rise to proton resonances in the ^{1}H spectrum and thereby obscure any resonances of interest. For RhlB studies however, the organic buffer substances

tris(hydroxymethyl)aminomethane (Tris; at pH 8.3) and bis(2-hydroxyethyl)amino-tris(hydroxymethyl)methane (BisTris; at pH 6.2) were used instead of potassium phosphate to avoid influencing the reaction equilibrium of RhlB's ATP hydrolysis through the addition of phosphate ions. For both buffering agents the methylene and amine groups give rise to proton resonances in the ^{1}H spectrum at around 3.2 and 1.2 ppm, respectively, which do not overlap with RNA resonances of interest. Whenever necessary the corresponding signals have been annotated in the spectrum. Whenever potassium phosphate buffer was used for initial assessment of RNA duplexes in absence of protein via NMR spectroscopy it was noted accordingly.

The typical salt concentration of 50-100 mM KCl for RNA samples was increased to 150 mM KCl to stabilize the helicase. Higher salt concentrations of 300 mM KCl proved to prevent the interaction of RhlB with the substrate RNA and were therefore avoided. Different $MgCl_2$ concentrations were tested to analyse how its duplex stabilizing effect changed linewidths, chemical shift, and spectral overlap. As exemplified in Figure 24 with J2Δ8-M2, the addition of 1.5 mM $MgCl_2$ did have a significant effect on the chemical shift dispersion of the peaks while the linewidths were only lightly affected. An increase of the magnesium concentration to 4.5 mM did not affect the chemical shifts any further. To serve as counterions to the 3 mM ATP that were added in some reaction samples, a final concentration of 4.5 mM $MgCl_2$ was used for experiments.

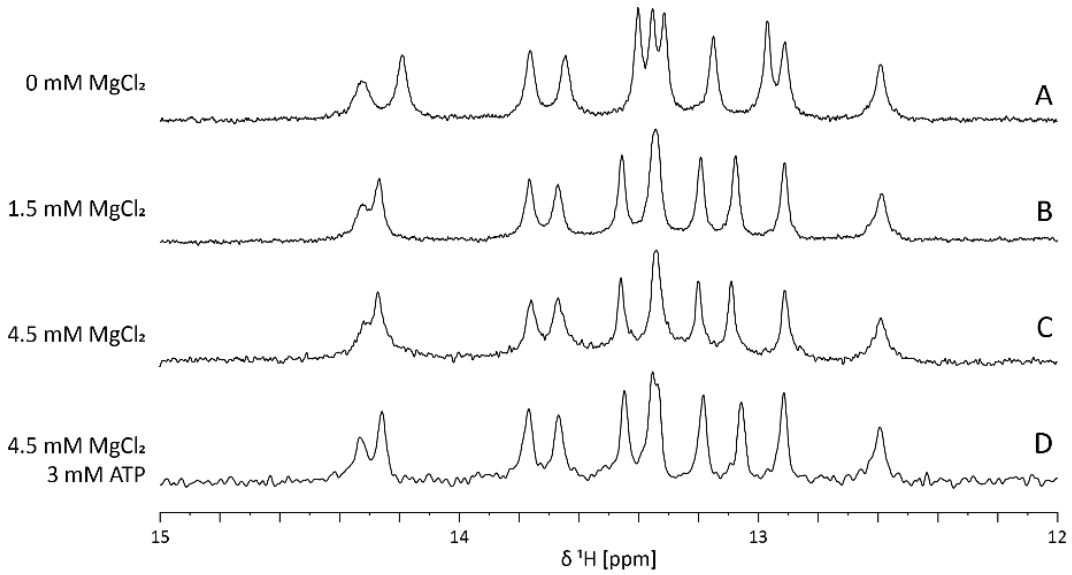

Figure 24. Imino regions of ^{1}H NMR spectra of RNA duplex J2Δ8-M2 at different $MgCl_2$ concentrations. A. 100 µM sample in 150 mM KCl, 25 mM Tris, 100 µM DSS, 5 mM DTT and 7% D_2O at pH 8.3 measured at 288 K with 256 scans at 600 MHz. B. 100 µM sample in 150 mM KCl, 25 mM Tris, 100 µM DSS, 5 mM DTT, 1.5 mM $MgCl_2$ and 7% D_2O at pH 8.3 measured at 288 K with 256 scans at 600 MHz. C. 50 µM sample in 150 mM KCl, 25 mM Tris, 50 µM DSS, 4.5 mM $MgCl_2$ and 7% D_2O at pH 8.3 measured at 288 K with 264 scans at 600 MHz. D. 50 µM sample in 150 mM KCl, 25 mM Tris, 50 µM DSS, 4.5 mM $MgCl_2$, 3 mM ATP and 7% D_2O at pH 8.3 measured at 288 K with 264 scans at 600 MHz.

4.3.2 J2h-M2

The first RNA duplex to be analysed via NMR spectroscopy was the prototype-construct J2h-M2, which consisted of a 21 bp heteroduplex with the 3' end of J2h extending into a 5 bp stem-loop structure, connect through a 6-nucleotide linker. Both 3' and 5' end of J2h were additionally extended by two uridine residues (for secondary structure see Figure 25C).

For the investigation of the secondary structure of the J2h-M2 construct via imino proton assignment, ^{1}H 1D spectrum with jump-return-echo water suppression, ^{1}H,^{15}N-HSQC and ^{1}H,^{1}H-NOESY experiments were recorded at 283 K. The spectra were measured with a 130 µM selectively ^{15}N labelled sample (G/U ^{15}N M2) in 50 mM KCl, 25 mM KH_2PO_4/K_2HPO_4 and 10% D_2O at pH 6.2 (Figure 25). In the NOESY spectrum, 23 distinct imino proton resonances could be identified between 11.8 and 14.0 ppm, which partially overlapped in the ^{1}H 1D spectrum. Since imino protons of canonical Watson-Crick base pairs resonate at a chemical shift between 12-15 ppm, the broadened resonances between 10 and 11 ppm were attributed to non-canonical and less stable base pair interactions, possibly in the loop region of the J2h hairpin. In the ^{1}H,^{15}N-HSQC spectrum, 5 out of the 7 uridines and one guanine resonance (G3) of the ^{15}N labelled M2 strand were detected and could be identified via their characteristic ^{15}N chemical shift. Out of the 5 uridines, 4 could be unambiguously assigned to nucleotides $U4_{M2}$, $U8_{M2}$, $U15_{M2}$ and $U16_{M2}$, while the terminal imino protons $U21_{M2}$ and $U1_{M2}$ as well as $U20_{M2}$ could not be assigned or detected. With those identified resonances as starting points, the remaining imino protons could be assigned in sequential NOE walks within the imino-imino region of the NOESY spectrum. Out of 23 imino peaks, 21 were successfully assigned. Only two resonances at 13.79 and 13.27 ppm did not exhibit reliable NOE cross peaks and could therefore not be assigned with certainty. The assignment showed that the heteroduplex between RNA strands J2h and M2 was formed, but only 18 out of 21 stable base pairs could be detected. The terminal and near-terminal AU base pairs of $U1_{M2}$, $U20_{M2}$ and $U21_{M2}$ appeared to be in fast exchange with solvent water and are therefore broadened beyond detectability. The formation of J2h's hairpin could also be confirmed with the assignment of nucleotides G32, G33 and G40, while neither U31 nor U44 could be identified. It could further be observed for instance on resonances G11 and G32/40 in Figure 25A, that the imino signals of the J2h hairpin show a significantly narrower linewidth and higher signal intensity than the imino protons of the hetero duplex, indicating a more stable secondary structure formation.

As explained in more detail in section 4.3.1 did the NMR buffer undergo several alterations during this project to accommodate for both RNA and protein requirements. Figure 26 illustrates the effects of different salt conditions on J2h-M2's imino proton region of the ^{1}H spectrum. While spectrum A with 50 mM KCl and 25 mM KH_2PO_4/K_2HPO_4 led to a set of distinct resonance peaks, was the spectral quality of spectrum B, a sample prepared with 150 mM KCl and otherwise identical buffer composition, significantly

worse. Most notably, the only sharp resonances that could be identified from the previous assignment, were G32, G33 and G40 located in the hairpin of J2h. Signals from the heteroduplex were barely detectable, which implied an incomplete or incorrect duplex formation. Spectrum C of Figure 25 shows another J2h-M2 RNA sample prepared with the same buffer, supplemented with 1.5 mM $MgCl_2$. Even though this spectrum was recorded at a higher temperature (298 K) than the previous two spectra (283 K), one can assume stable duplex formation based on the melting point of 63°C previously determined via CD melting curve. However, this sample also showed notably narrow and distinct peaks for the J2h hairpin protons while resonances of the hetero duplex were broadened considerably. Repeated measurements with different samples showed ^{1}H 1D spectra of comparable quality. So far it could not be explained why the sheer increase in ionic strength to 150 mM KCl had such an adverse effect on duplex formation. Considering that the linewidth of imino protons would be further increased by measurements at higher pH and possible complex formation with the helicase, it was concluded that the RNA duplex did not fulfil the requirements necessary to perform protein titration experiments at pH 8.3. Not only was it questionable, to what percentage the RNA adopted the duplex fold, but subsequent measurements required a certain number of narrow non-overlapping imino reporter signals. Hence it was decided to dismiss J2h-M2 as a substrate RNA.

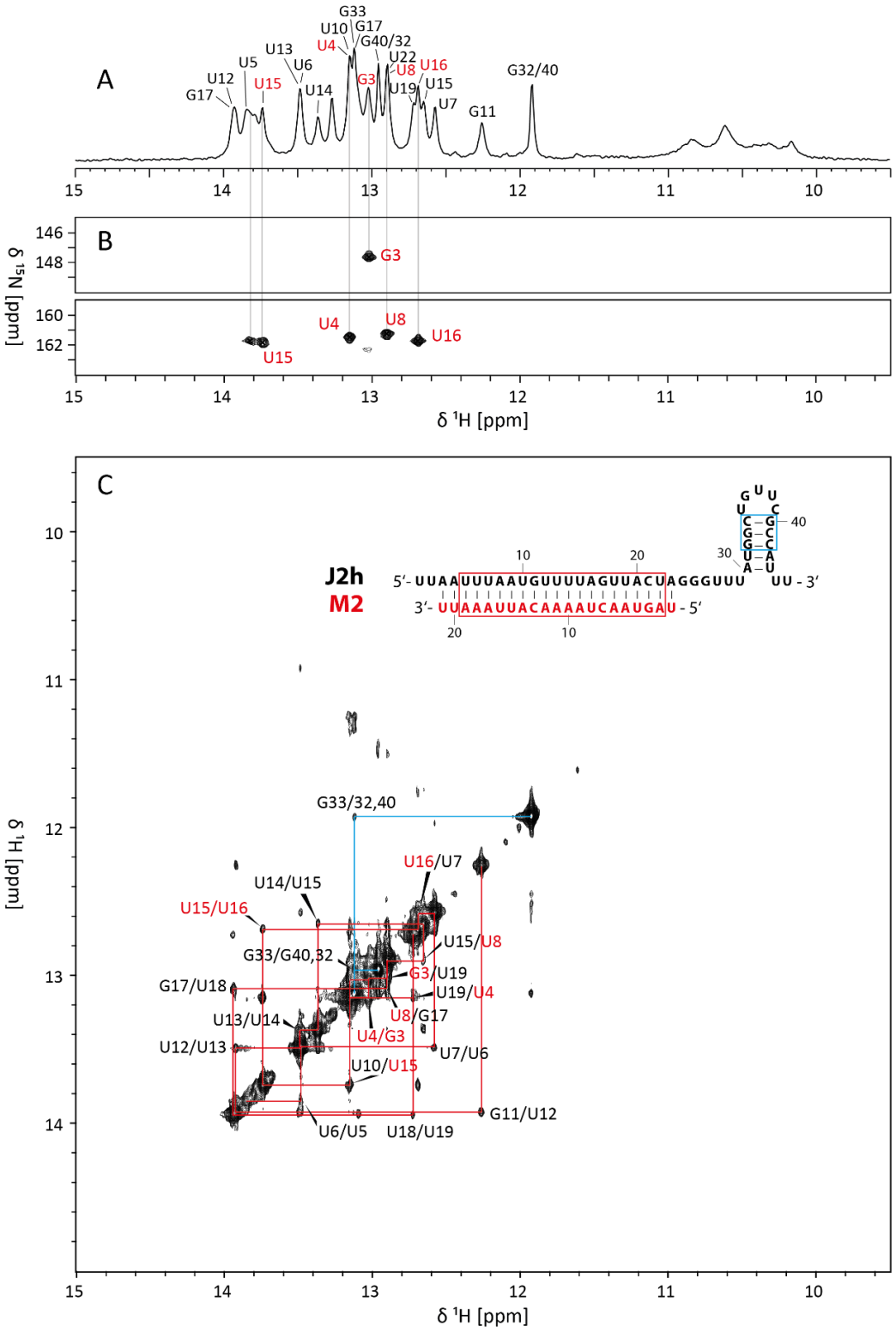

Figure 25. Imino resonance assignment of selectively ^{15}N labelled RNA duplex J2h-M2 (130 µM, M2 ^{15}N G/U labelled) in 50 mM KCl, 25 mM KH_2PO_4/K_2HPO_4 and 10% D_2O at pH 6.2. Imino resonance assignments are colour coded according to J2h (black) and M2 (red) single strand. A. Imino region of ^1H 1D spectrum recorded at 283 K with 128 scans at 800 MHz. B. $^1H,^{15}N$-HSQC spectrum recorded at 283 K with 16 scans at 800 MHz. C. Imino region of $^1H,^1H$-NOESY spectrum recorded at 283 K with 128 scans at 900 MHz. Sequential walks of J2h hairpin (blue) and the heteroduplex (red) are indicated by lines.

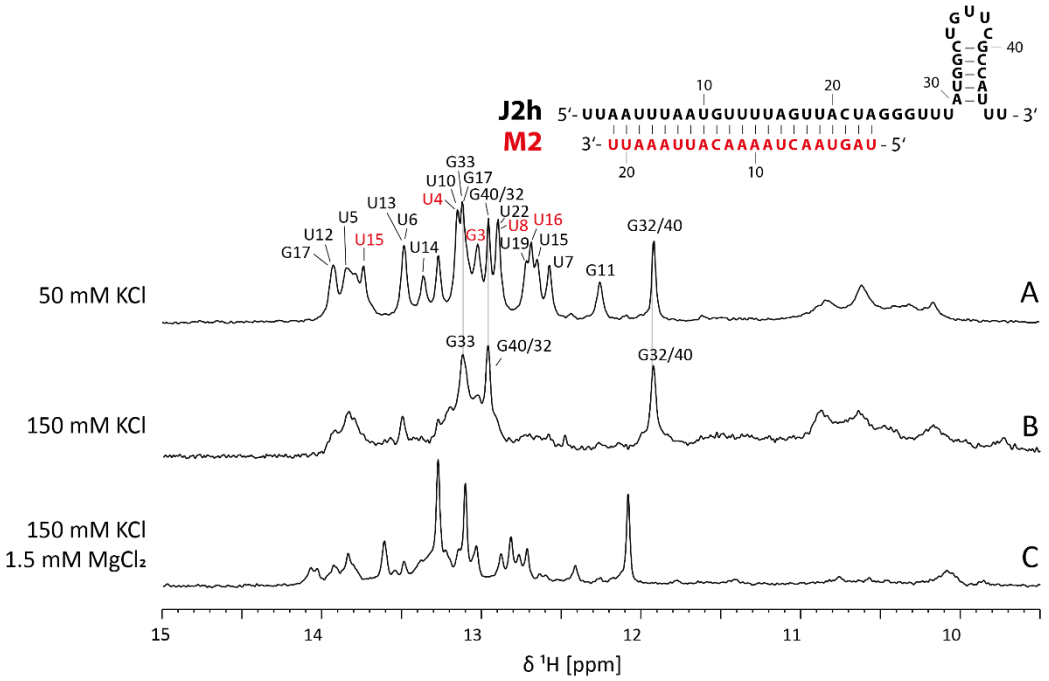

Figure 26. Imino regions of ^{1}H NMR spectra of RNA duplex J2h-M2 at different buffer conditions. A. 130 µM sample in 50 mM KCl, 25 mM KH_2PO_4/K_2HPO_4 and 10% D_2O at pH 6.2 measured at 283 K with 128 scans at 800 MHz. B. 160 µM sample in 150 mM KCl, 25 mM KH_2PO_4/K_2HPO_4, 20 µM DSS and 10% D_2O at pH 6.2 measured at 283 K with 128 scans at 600 MHz. C. 145 µM sample in 150 mM KCl, 25 mM KH_2PO_4/K_2HPO_4, 1.5 mM $MgCl_2$ and 7% D_2O at pH 6.2 measured at 298 K with 512 scans at 800 MHz.

4.3.3 J2Δ8-M2

In RNA duplex J2Δ8-M2 the M2 strand was paired with J2Δ8, a truncated 13 Nt long version of the J2h strand, which should form a 13 Nt heteroduplex with an 8 Nt single stranded overhang at the 3' end of M2. For the imino proton assignment the following experiments were used: ^{1}H 1D, ^{1}H,^{15}N-HSQC and ^{1}H,^{1}H-NOESY. Figure 27 shows the corresponding spectra of an 880 µM selectively ^{15}N labelled (^{15}N M2) sample in 25 mM NaCl, 50 mM BisTris and 7% D_2O at pH 6.5, which was synthesized and measured at 298 K by Katharina Hohmann during her master thesis.

Of the 13 expected imino protons 11 could be detected in the NOESY spectrum between 12.5 and 14.2 ppm. The two uridine and one guanine signal observed in the ^{15}N-HSQC were assigned to $G3_{M2}$, $U4_{M2}$ and $U8_{M2}$, while the expected imino proton resonance for $U1_{M2}$ could not be detected. Using a sequential imino proton walk, all the remaining imino protons were successfully assigned. At 298 K, no signals were observable for the terminal AU base pairs involving $U1_{M2}$ and $U20_{J2\Delta8}$, again hinting at imino protons with high solvent exchange rates. In measurements at a lower temperature (288 K) an additional imino proton with broader line width appeared at 13.2 ppm, as can be seen in spectrum A of Figure 28. This resonance was suspected to be $U20_{J2\Delta8}$.

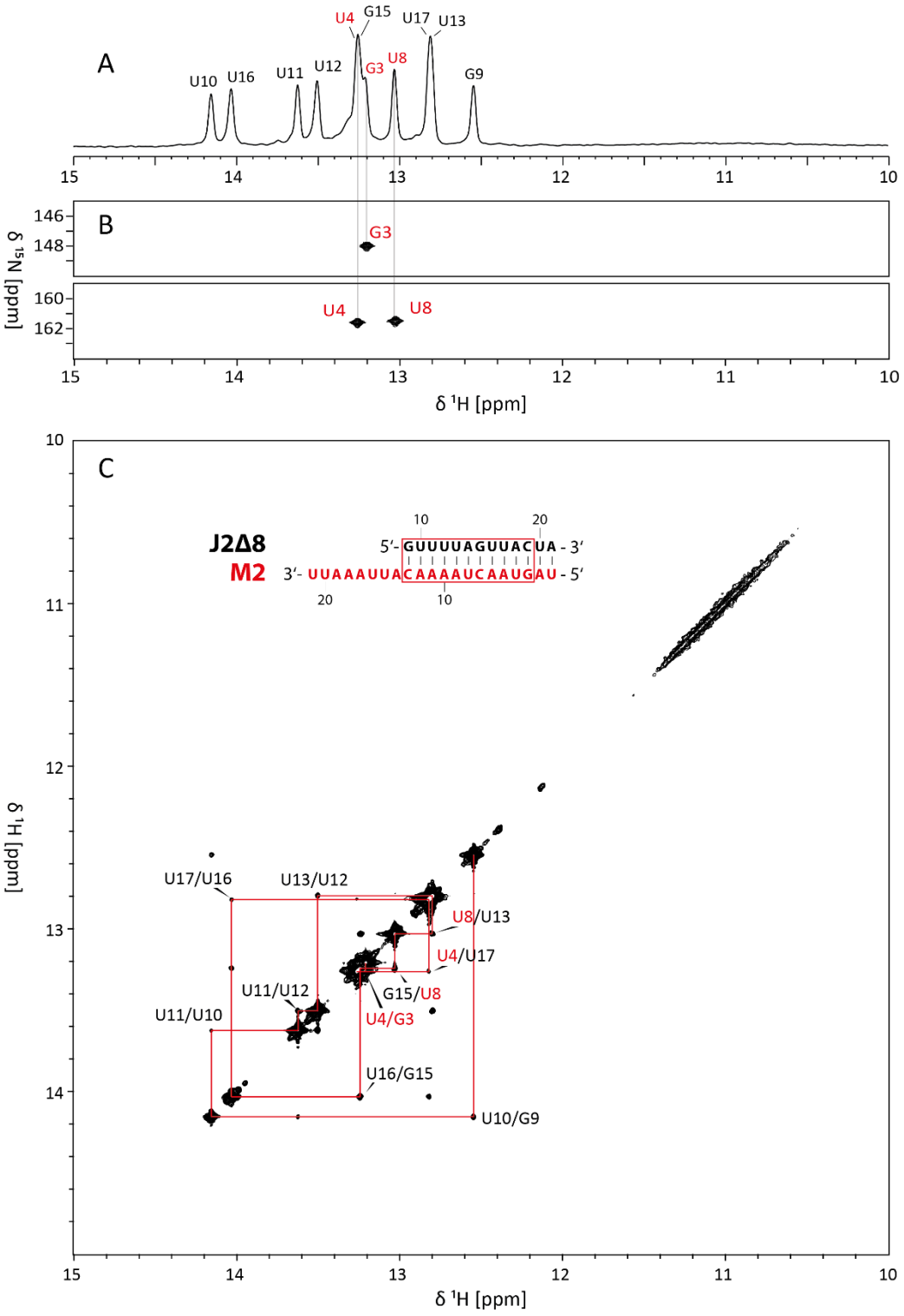

Figure 27. Imino resonance assignment of selectively ^{15}N labelled RNA duplex J2Δ8-M2 (880 µM, M2 ^{15}N G/U labelled) in 25 mM NaCl, 50 mM BisTris and 7% D_2O at pH 6.5. Imino resonance assignments are colour coded according to J2Δ8 (black) and M2 (red) single strand. A. Imino region of ^{1}H 1D spectrum recorded at 298 K with 32 scans at 600 MHz. B. ^{1}H,^{15}N-HSQC spectrum recorded at 298 K with 8 scans at 800 MHz. C. Imino region of ^{1}H,^{1}H-NOESY spectrum recorded at 298 K with 32 scans at 800 MHz. The sequential walk is indicated by red lines. The depicted spectra were recorded in the context of Katharina Hohmann's master thesis.

The overall spectral quality of the imino proton region of J2Δ8-M2 was very promising, so consecutive buffer changes were performed, ultimately resulting in measurements at 288 K in 150 mM KCl, 25 mM Tris and 1.5 mM $MgCl_2$ at pH 8.3 (Figure 28). Under those higher pH conditions minor peak broadening could be observed for all imino resonances which was more pronounced the further the imino protons were located towards the ends of the duplex. $U20_{J2\Delta8}$ even broadened beyond detectability. The addition of Mg^{2+} ions induced several proton chemical shift changes, which slightly improved the overall chemical shift dispersion and therefore the number of reporter signals available for protein interaction experiments.

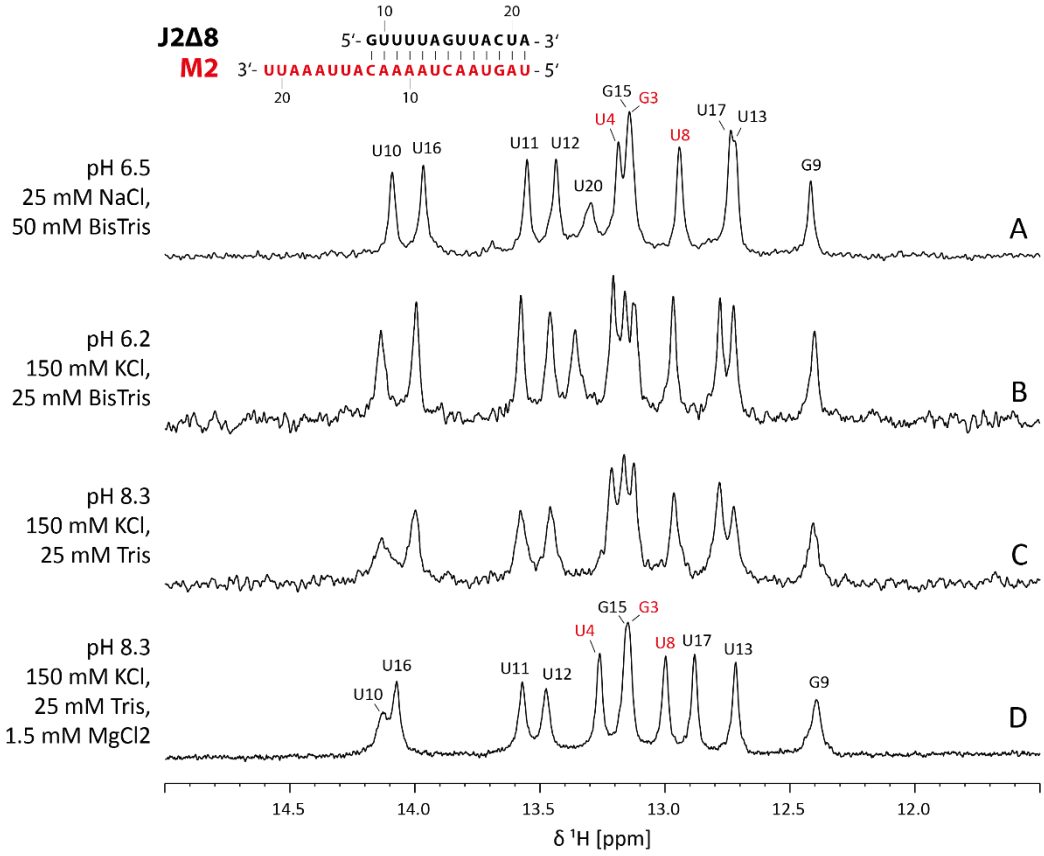

Figure 28. Imino regions of ^1H NMR spectra of RNA duplex J2Δ8-M2 at different buffer conditions. Imino resonance assignments are colour coded according to J2Δ8 (black) and M2 (red) single strand. A. 880 µM sample in 25 mM NaCl, 50 mM BisTris and 7% D_2O at pH 6.5 measured at 288 K with 32 scans at 800 MHz. B. 50 µM sample in 150 mM KCl, 25 mM BisTris, 50 µM DSS and 7% D_2O at pH 6.2 measured at 288 K with 192 scans at 600 MHz. C. 50 µM sample in 150 mM KCl, 25 mM Tris, 50 µM DSS and 7% D_2O at pH 8.3 measured at 288 K with 264 scans at 600 MHz. D. 100 µM sample in 150 mM KCl, 25 mM Tris, 1.5 mM $MgCl_2$, 5 mM DTT, 100 µM DSS and 7% D_2O at pH 8.3 measured at 288 K with 264 scans at 600 MHz.

Figure 29 depicts the aromatic nucleobase resonance regions of the $^{1}H^{13}C$ HSQC spectrum of the selectively ^{13}C labelled J2Δ8-M2 (^{13}C M2) RNA, which was recorded in pH 8.3 buffer at 600 MHz and 288 K. The final RNA assignment of the M2 strand was achieved with a combination of the following NMR experiments: $^{1}H^{1}H$ NOESY, ^{13}C HSQC, ^{15}N HSQC, HCN, ^{1}H NOESY-HSQC and TROSY relayed HCCH-COSY.

As shown in the ^{13}C HSQC spectrum, it was possible to identify all 9 expected C6H6 peaks, resonating between 7.1-8.1 ppm in ^{1}H and 139.5-146 ppm in ^{13}C dimension, as their resonances are split into duplets due to evolution of homonuclear ^{13}C coupling. Of the 12 C8H8 purine resonances, which are located between 136-142 ppm and 7.3-8.4 ppm in ^{13}C and ^{1}H dimension, respectively, all but the following nucleotides could be assigned: For resonances A_9 and A_{14} an unambiguous assignment was not possible due to peak overlap in aromatic-to-sugar region of the complementary NOESY spectrum. In case of the resonances of the adenines from the single stranded overhang A_{17}, A_{18} and A_{19} the single strand flexibility prevented the formation of most amino-to-sugar NOESY cross peaks critical for a distinction of the three nucleotides. The same reason impeded the clear assignment of their C2H2 nucleobase resonances. The remainder of the 11 C2H2 nucleobase signals, that resonate at approximately 152-155 ppm (^{13}C), could also be completely identified and assigned to the respective nucleotide.

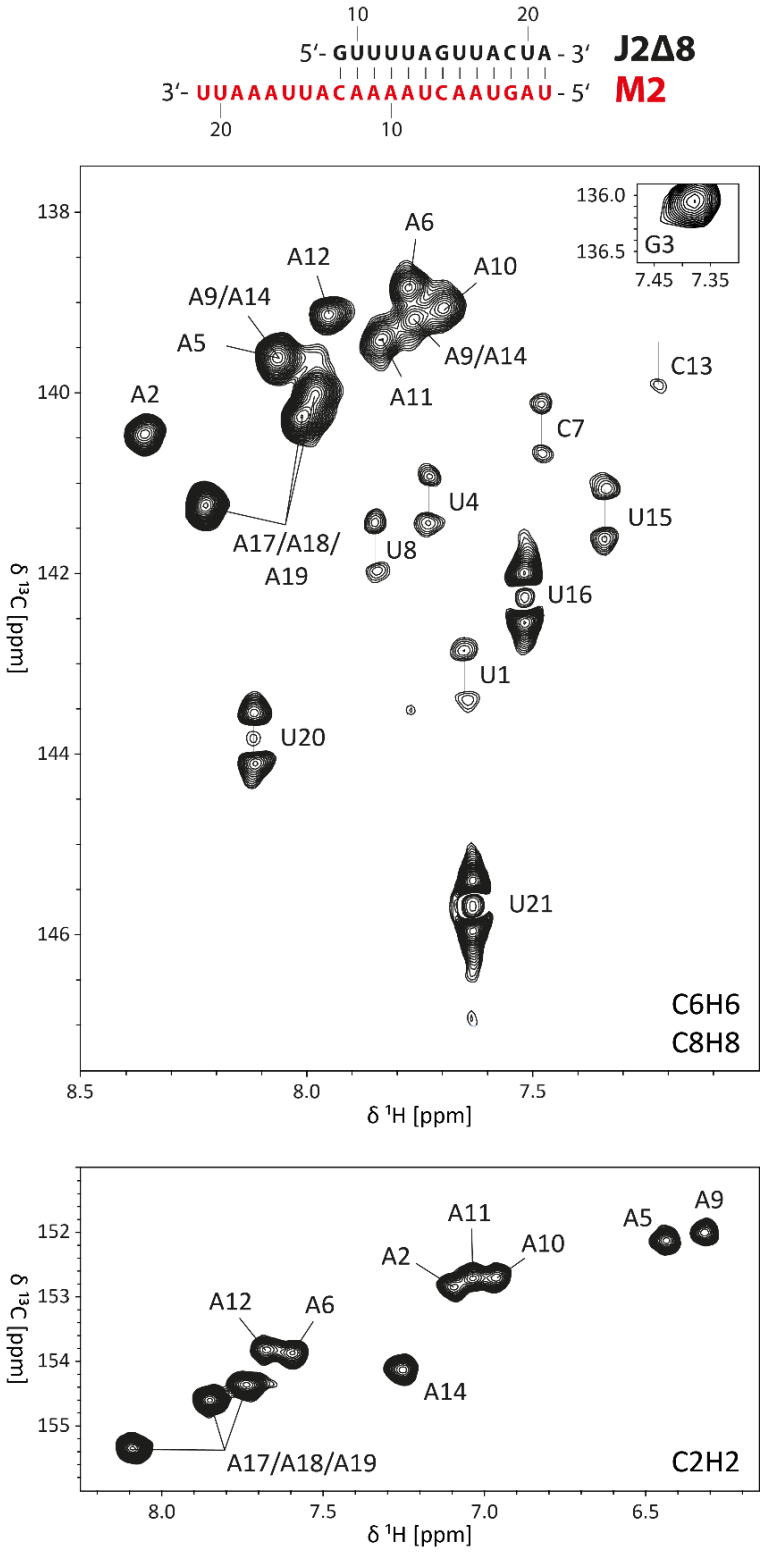

Figure 29. C6H6, C8H8 and C2H2 resonance regions of the $^{1}H^{13}C$ HSQC spectrum of selectively ^{13}C labelled J2Δ8-M2 (M2 fully ^{13}C labelled). Spectrum was recorded in pH 8.3 NMR buffer with 100 µM RNA at 288 K and 600 MHz with 32 scans. The resonance assignment of M2 is annotated in the spectrum. Where an unambiguous assignment was not possible, the labels indicate the possible assignments separated by slashes.

4.3.4 J2Δ16-M2

J2Δ16-M2 was designed as one of two 5' overhang constructs and as the elongated version of J2Δ14-M2. Here the J2 strand was truncated from the 3' end to result in the 15 Nt long J2Δ16 RNA, which, paired with M2, led to a 15 Nt long duplex with a 5' single strand overhang of 6 Nt. The imino proton assignment was performed analogue to previous RNA duplexes with the measurement of an ^{1}H, ^{1}H,^{15}N-HSQC and ^{1}H,^{1}H-NOESY experiments. Figure 30 depicts the results of those measurements on a 100 µM selectively ^{15}N labelled (^{15}N M2) and a 300 µM unlabelled RNA sample at 288 K. In the ^{1}H spectrum, 13 out of 15 expected imino protons could be observed between 12.5 and 14.3 ppm as well as two very broad resonances between 10.5 and 11.5 ppm. The ^{1}H,^{15}N-HSQC showed three uridine residues, which could be identified as $U8_{M2}$, $U15_{M2}$ and $U16_{M2}$ by the sequential NOESY walk. Resonances for both $U20_{M2}$ and $U21_{M2}$ could not be detected. With the assignment of the remaining proton resonances, it was evident that, except for the two 3' terminal base pairs involving $U20_{M2}$ and $U21_{M2}$, all base pairs of the intended hetero duplex could be verified. Identifying the two resonances between 10.5 and 11.5 ppm was unsuccessful. They could neither be detected in the imino-imino region of the NOESY spectrum nor in the ^{15}N HSQC. Their proton chemical shift and broad linewidth were indicative of non-canonical and instable base pair interactions, but this could not be confirmed. Non-canonical base pairs are not present in the hetero duplex and an equivalent second conformation was not detected in native RNA-PAGE. A more transient interaction of two separate single stranded extensions of M2 or interactions of a folded back single strand tail with its own duplex are possible but speculative.

With the conversion of the buffer to 150 mM KCl, 25 mM Tris and 4.5 mM $MgCl_2$ at pH 8.3 (Figure 31), especially the near-terminal imino protons experienced a significant line broadening due to the increased exchange rates, ultimately leading to the disappearance of $G15_{J2\Delta16}$ and $U3_{J2\Delta16}$ and the extensive intensity decrease for $U8_{M2}$. The remaining imino resonances however displayed intensities and a resolution favourable for further investigations.

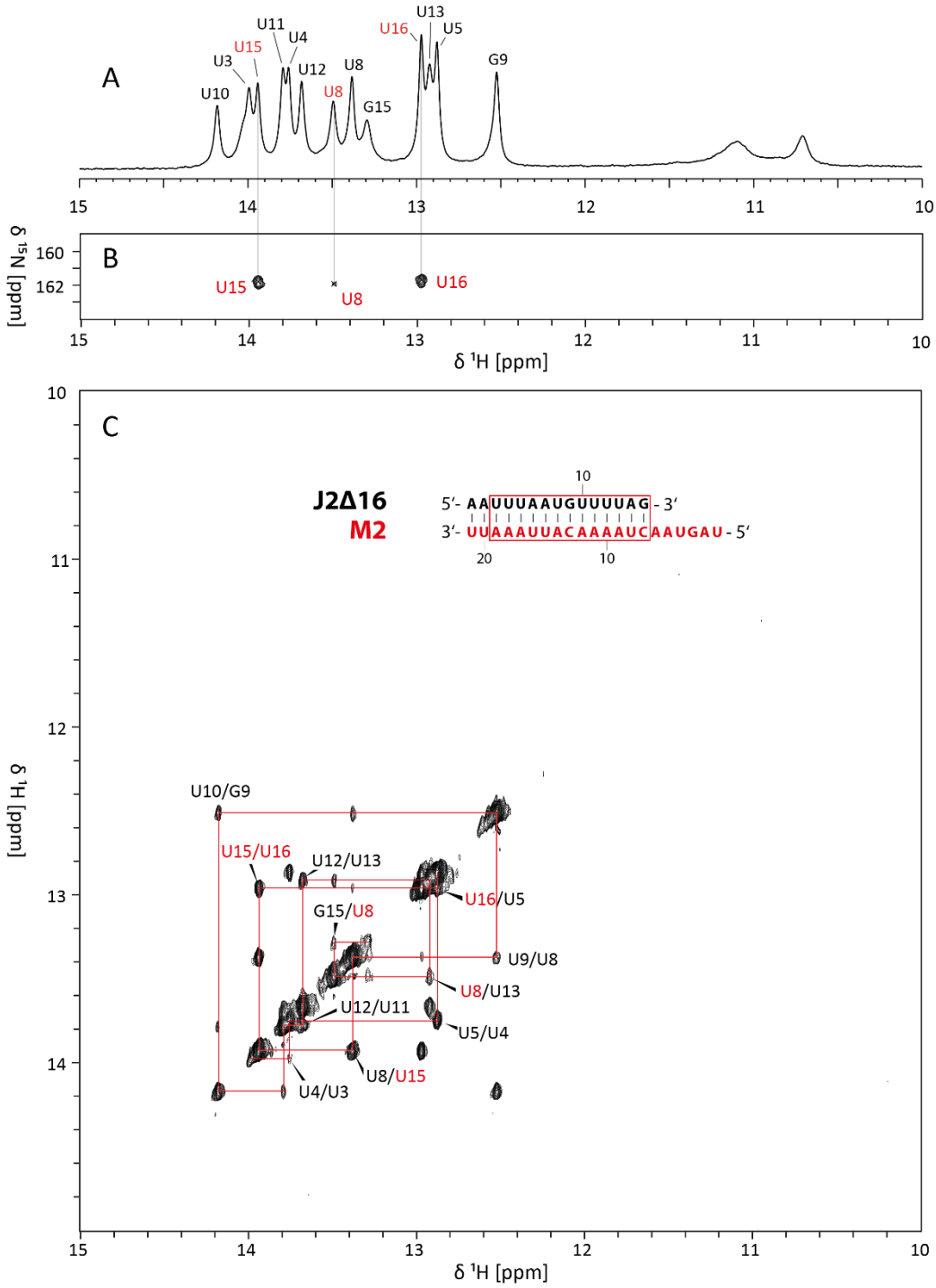

Figure 30. Imino resonance assignment of RNA duplex J2Δ16-M2. Imino resonance assignments are colour coded according to J2Δ16 (black) and M2 (red) single strand. A. Imino region of ^{1}H 1D spectrum of 300 µM unlabelled sample in 150 mM KCl, 25 mM BisTris, 4.5 mM $MgCl_2$, 100 µM DSS and 8% D_2O at pH 6.2 recorded at 288 K with 256 scans at 900 MHz. B. ^{1}H,^{15}N-HSQC spectrum of 100 µM selectively ^{15}N labelled sample (M2 full ^{15}N labelled) in 150 mM KCl, 25 mM Tris, 4.5 mM $MgCl_2$, 100 µM DSS and 8% D_2O at pH 8.3 recorded at 288 K with 128 scans at 600 MHz. C. Imino region of ^{1}H,^{1}H-NOESY spectrum of 300 µM unlabelled sample in 150 mM KCl, 25 mM BisTris, 4.5 mM $MgCl_2$, 100 µM DSS and 8% D_2O at pH 6.2 recorded at 288 K with 256 scans at 900 MHz. The sequential walk is indicated by red lines.

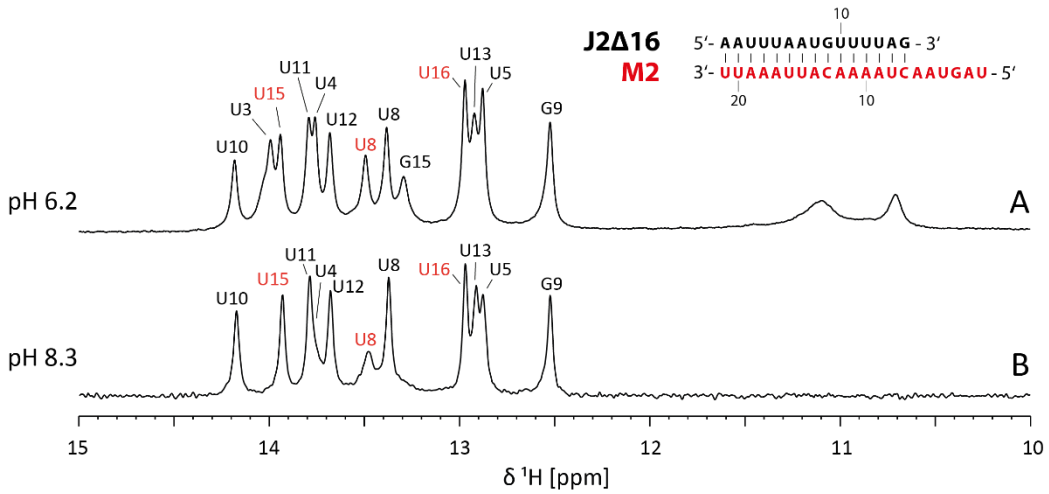

Figure 31. Imino regions of ^{1}H NMR spectra of RNA duplex J2Δ16-M2 at different buffer conditions. Imino resonance assignments are colour coded according to J2Δ16 (black) and M2 (red) single strand. A. 300 µM sample in 150 mM KCl, 25 mM BisTris, 4.5 mM $MgCl_2$, 100 µM DSS and 8% D_2O at pH 6.2 recorded at 288 K with 256 scans at 900 MHz. B. 100 µM sample in 150 mM KCl, 25 mM Tris, 4.5 mM $MgCl_2$, 100 µM DSS and 8% D_2O at pH 8.3 recorded at 288 K with 128 scans at 600 MHz.

Figure 32 depicts the C6H6, C8H8 and C2H2 resonance regions of the $^{1}H^{13}C$ HSQC spectrum of the selectively ^{13}C labelled J2Δ16-M2 (^{13}C M2) which was recorded in pH 8.3 buffer at 700 MHz and 288 K. The final RNA assignment of the M2 strand was achieved with a combination of the following NMR experiments: $^{1}H^{1}H$ NOESY, ^{13}C HSQC, ^{15}N HSQC, HCN, ^{1}H NOESY-HSQC and TROSY relayed HCCH-COSY.

In this ^{13}C experiment again the C6H6 resonances are split into duplets due to homonuclear ^{13}C coupling. With exception of C_7 and C_8, all of the 9 pyrimidine resonances could be assigned unambiguously. An additional 10^{th} resonance was identified and could also be ascribed to the U_{21} nucleotide. Based on their relative peak position, the U_{21} peak at 144 ppm (^{13}C) was assigned to the base paired conformation of the nucleotide ($U21_{ds}$), while the more high-field shifted resonance at 145.7 ppm (^{13}C) was determined to be the conformation of the unpaired fraying terminal base ($U21_{ss}$). This assignment was further corroborated by the matching chemical shift with the unpaired U_{21} resonance of J2Δ8-M2, and by the results in section 5.6, where the selective duplex opening increased the relative population of the unpaired U_{21} signal. Of the 12 C8H8 purine resonances all but A_9/A_{10} could be assigned. For the high-field shifted C2H2 signals an unambiguous assignment of A_2, A_5 and A_6 was not possible due to the absence of some distinctive amino-sugar cross peaks in the NOESY spectrum.

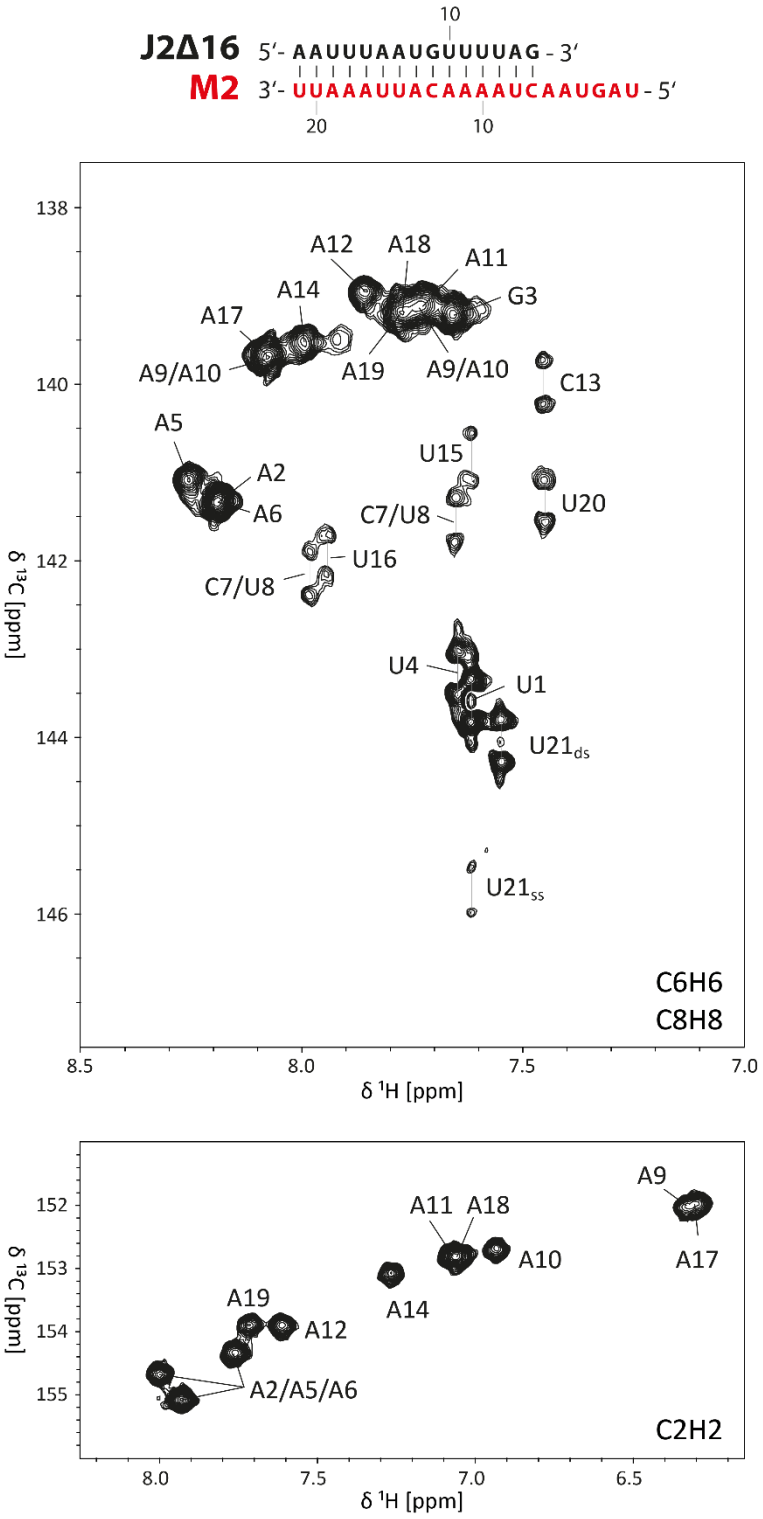

Figure 32. C6H6, C8H8 and C2H2 resonance regions of the $^{1}H^{13}C$ HSQC spectrum of selectively ^{13}C labelled J2Δ16-M2 (M2 fully ^{13}C labelled). Spectrum was recorded in pH 8.3 NMR buffer with 100 µM RNA at 288 K and 700 MHz with 52 scans. The resonance assignment of M2 is annotated in the spectrum. Where an unambiguous assignment was not possible, the labels indicate the possible assignments separated by slashes.

4.3.5 J2Δ14-M2

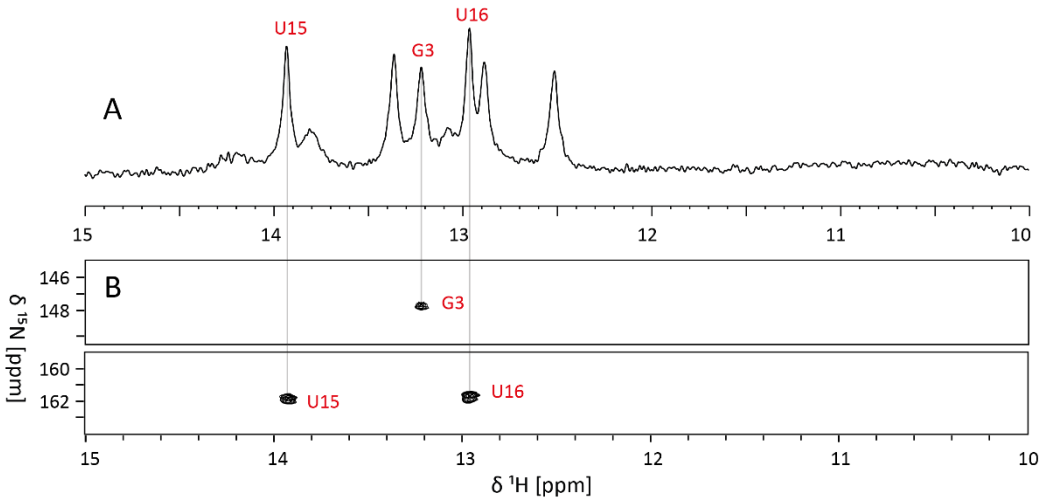

Figure 33. Imino resonance assignment of selectively ^{15}N labelled RNA duplex J2Δ14-M2 (100 µM, M2 full ^{15}N labelled) in 25 mM KCl, 25 mM Tris, 4.5 mM $MgCl_2$, 100 µM DSS and 8% D_2O at pH 8.3. Imino resonance assignments are colour coded according to J2Δ8 (black) and M2 (red) single strand. A. Imino region of ^{1}H 1D spectrum recorded at 288 K with 256 scans at 600 MHz. B. ^{1}H,^{15}N-HSQC spectrum recorded at 288 K with 32 scans at 600 MHz.

The RNA duplex J2Δ14-M2 was designed to include a 5' overhang construct with the same duplex length as J2Δ8-M2 of 13 bp in the set of RNA substrates. Although theoretical calculations predicted a significantly lower duplex stability of -9.82 kcal mol^{-1} compared to -15.50 kcal mol^{-1} for J2Δ8-M2, J2Δ14-M2 was tested spectroscopically, nonetheless. Figure 33 shows the ^{1}H 1D and ^{1}H,^{15}N-HSQC spectrum of a 100 µM ^{15}N labelled sample (^{15}N M2) measured in 25 mM KCl, 25 mM Tris, 4.5 mM $MgCl_2$ at 288 K. In the imino proton region six narrow as well as two broad resonances could be observed between 12.4 and 14.0 ppm. The ^{1}H,^{15}N-HSQC unexpectedly showed two uridine and one guanine signal. Since the ^{15}N labelled M2 strand contains merely one guanine at position 3 ($G3_{M2}$), it could be the only nucleotide causing this signal. However, the proposed hetero duplex of J2Δ14 and M2 did not include a base paired $G3_{M2}$, which pointed towards a substantial deviation from the predicted hetero duplex formation. This assumption was corroborated by the inconclusive duplex formation results in native polyacrylamide gel electrophoresis, where the observed duplex band exhibited an unexpected running behaviour. The two observable uridine residues on the other hand could be assigned to $U15_{M2}$ and $U16_{M2}$, as their chemical shifts matched the resonances of J2Δ16-M2's ^{1}H,^{15}N-HSQC. A side-by-side comparison of the ^{1}H imino proton spectra of J2Δ14-M2 and J2Δ16-M2 at the same buffer conditions (Figure 34) revealed three more resonances ($U5_{J2Δ14}$, $U8_{J2Δ14}$ and $G9_{J2Δ14}$), that could potentially be assigned to the hetero duplex, completing a stretch of five stable base pairs in the centre of the duplex. It was questionable whether the single stranded 5' overhang of M2 could potentially fold back and form a single base pair between $G3_{M2}$ and $C7_{M2}$, while the central part of the heteroduplex was stable. Both the hetero duplex resonances and the GC base pair also persisted at a lower temperature (278 K).

With those conflicting results we refrained from a full imino proton assignment via $^{1}H,^{1}H$-NOESY and decided to dismiss J2Δ14-M2 from the set of RNA test substrates.

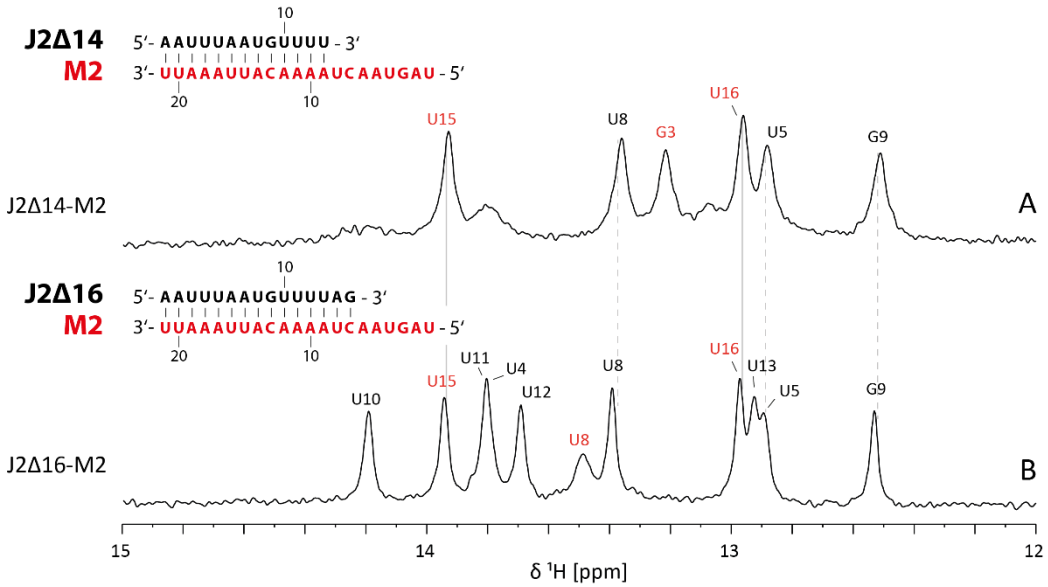

Figure 34. Imino region of ^{1}H NMR spectra of RNA duplex J2Δ14-M2 compared to J2Δ16-M2. Imino resonance assignments are colour coded according to J2Δx (black) and M2 (red) single strand. A. 100 µM sample in 150 mM KCl, 25 mM Tris, 4.5 mM $MgCl_2$, 100 µM DSS and 8% D_2O at pH 8.3 recorded at 288 K with 256 scans at 600 MHz. B. 100 µM sample in 150 mM KCl, 25 mM Tris, 4.5 mM $MgCl_2$, 100 µM DSS and 8% D_2O at pH 8.3 recorded at 288 K with 128 scans at 600 MHz.

4.3.6 J2-M2

The RNA duplex J2-M2 was designed as one of the two blunt end constructs. Therefore, the hairpin and single strand extensions of J2h-M2 were omitted to result in a 21 Nt blunt end hetero duplex. For the initial analysis of the imino proton pattern both a ^{1}H 1D and $^{1}H,^{15}N$-BEST-TROSY experiment were recorded on a 100 µM ^{15}N labelled sample (^{15}N M2) in 150 mM KCl, 25 mM KH_2PO_4/K_2HPO_4, 1.5 mM $MgCl_2$ at pH 6.2 at 298 K (Figure 35).

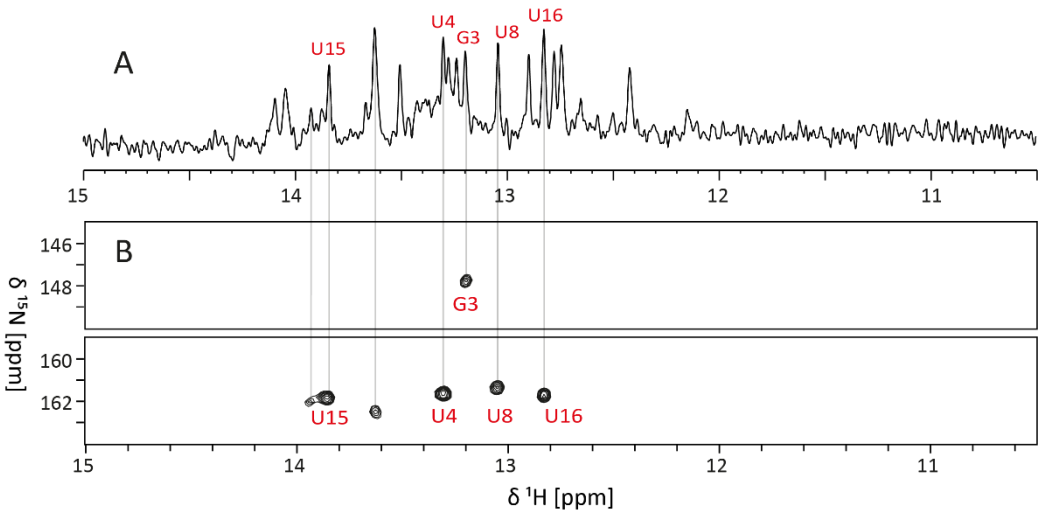

Figure 35. Imino resonance assignment of selectively ^{15}N labelled RNA duplex J2-M2 (100 µM, M2 full ^{15}N labelled) in 150 mM KCl, 25 mM KH_2PO_4/K_2HPO_4, 1.5 mM $MgCl_2$ and 7% D_2O at pH 6.2. Imino resonance assignments are colour coded according to J2 (black) and M2 (red) single strand. A. Imino region of ^{1}H 1D spectrum recorded at 298 K with 512 scans at 800 MHz. B. ^{1}H,^{15}N-BEST-TROSY spectrum recorded at 298 K with 128 scans at 800 MHz. The reduced signal to noise of the ^{1}H 1D spectrum was attributed to a residual acrylamide impurity in the sample that decreased the effective receiver gain of the measurement.

Despite the moderate signal-to-noise in the depicted ^{1}H spectrum, 16 distinct peaks could be identified in the imino proton region between 12.3 and 14.2 ppm. The BEST-TROSY spectrum gave rise to seven signals, one guanine and six uridines, of which five could be assigned by spectral overlap with J2Δ8-M2 and J2Δ16-M2 to $G3_{M2}$, $U4_{M2}$, $U8_{M2}$, $U15_{M2}$ and $U16_{M2}$. NMR measurements in pH 8.3 buffer conditions illustrated the high degree of chemical shift agreement between the blunt end duplex and the double stranded segments of J2Δ8-M2 and J2Δ16-M2 even further. It was possible, as highlighted in Figure 36, to assign all the remaining imino proton resonances directly by superimposing ^{1}H 1D spectra of the referred constructs. Here, the sequence of imino residues from $U4_{J2}$ to $U10_{J2}$ could be mapped to J2Δ16-M2 while all residues from $U11_{J2}$ to $G3_{M2}$ aligned with J2Δ8-M2. The terminal and next-to-terminal imino protons of $U21_{M2}$, $U20_{M2}$, $U20_{J2}$ and $U1_{M2}$ were again not detectable due to high solvent exchange rates which led to peak broadening beyond detectability. With only minor peak overlap in the ^{1}H spectrum as well as homogenous duplex formation J2-M2 was determined to be a suitable RNA substrate for RhlB studies.

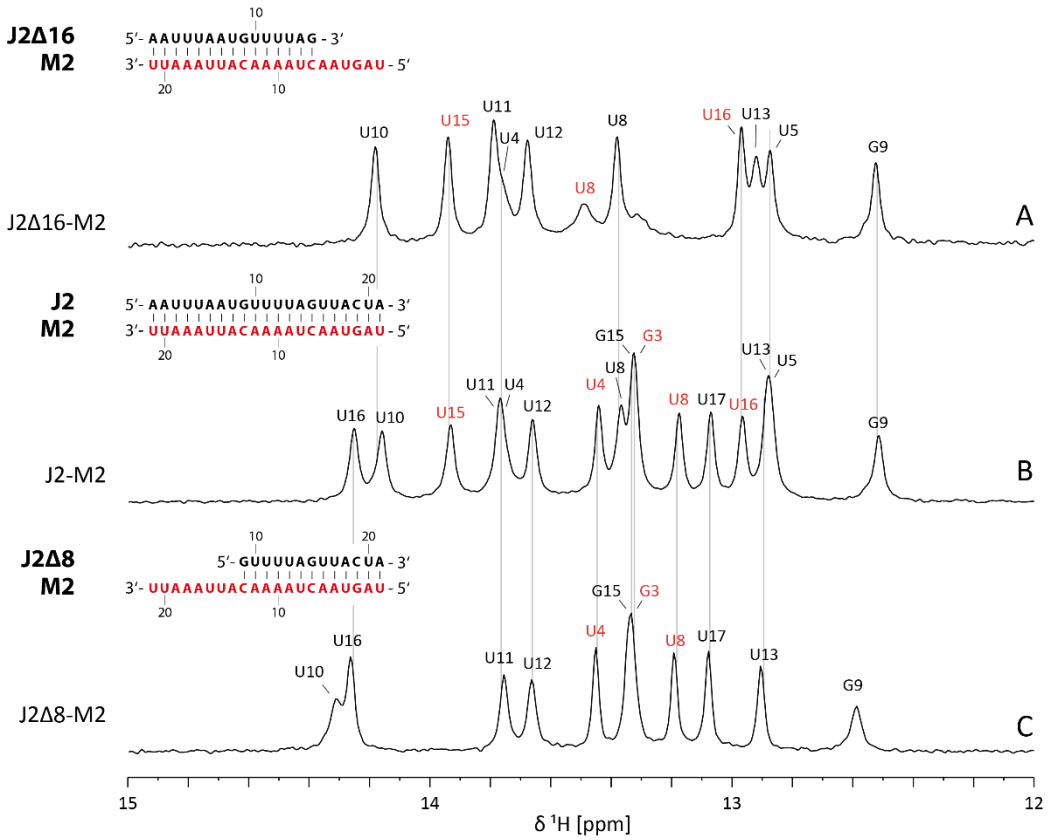

Figure 36. Imino region of ^{1}H NMR spectra of RNA duplex J2-M2 compared to J2Δ16-M2 and J2Δ8-M2. Imino resonance assignments are colour coded according to J2/J2Δ8/J2Δ16 (black) and M2 (red) single strand. A. 100 µM sample in 150 mM KCl, 25 mM Tris, 4.5 mM $MgCl_2$, 5 mM DTT, 100 µM DSS and 15% D_2O at pH 8.3 recorded at 288 K with 512 scans at 700 MHz. B. 100 µM sample in 150 mM KCl, 25 mM Tris, 4.5 mM $MgCl_2$, 5 mM DTT, 100 µM DSS and 10% D_2O at pH 8.3 recorded at 288 K with 128 scans at 700 MHz. C. 100 µM sample in 150 mM KCl, 25 mM Tris, 4.5 mM $MgCl_2$, 5 mM DTT, 100 µM DSS and 8% D_2O at pH 8.3 recorded at 288 K with 512 scans at 600 MHz.

4.3.7 J2Δ8-M3

The shorter blunt end construct J2Δ8-M3 resembled a truncated version of J2Δ8-M2, where M2 is reduced by 8 Nt to result in a 13 Nt hetero duplex. With the assignment of J2Δ8-M2 at hand, the imino proton assignment time for this construct could be shortened by directly aligning both ^{1}H 1D spectra. Figure 37 shows the spectral overlap in the corresponding ^{1}H spectra of 100 µM J2Δ8-M3 in the 150 mM KCl, 25 mM Tris, 4.5 mM $MgCl_2$, pH 8.3 recorded at 288 K side by side with J2Δ8-M2 measured under the same conditions. Nine of the expected 13 imino resonances could be observed for J2Δ8-M3, of which eight could be directly matched from the J2Δ8-M2 duplex. As with previous constructs, $U20_{J2\Delta8}$ and $U1_{M3}$ could not be detected due to peak broadening. Approaching the 5’ end of J2Δ8, the stabilizing effect of the 3’ single stranded overhang became apparent: while nucleotides $U10_{J2\Delta8}$ and $G9_{J2\Delta8}$ formed stable base pairs in the duplex with M2, neither $G9_{J2\Delta8}$ nor $U10_{J2\Delta8}$ could be detected in the case of M3. Even $U11_{J2\Delta8}$ was significantly broadened in the construct without the single strand extension.

Nevertheless, the construct was determined to be fully suitable for further NMR experiments.

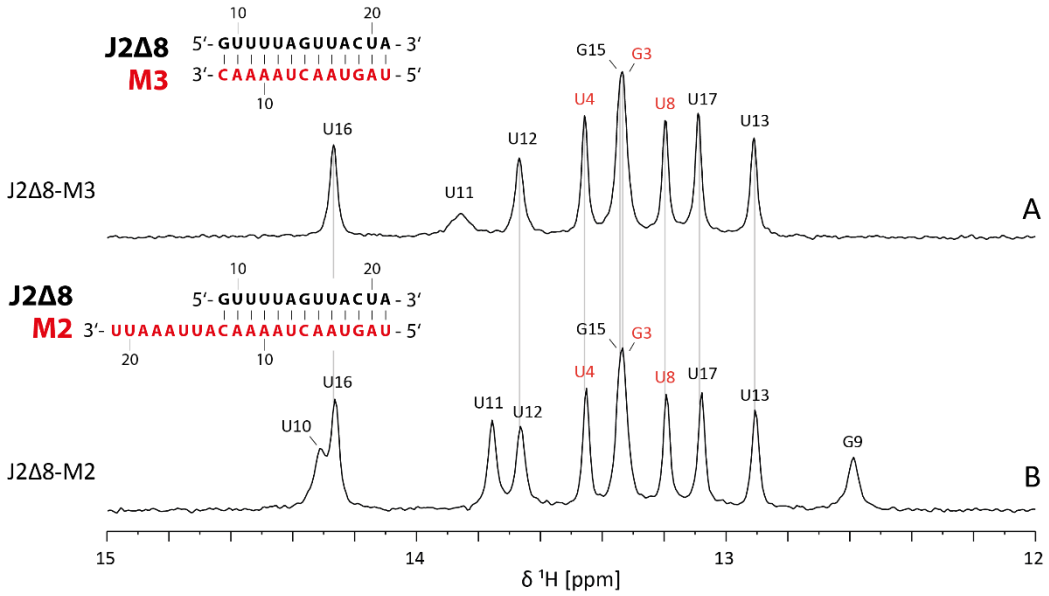

Figure 37. Imino region of ^{1}H NMR spectra of RNA duplex J2Δ8-M3 compared to J2Δ8-M2. Imino resonance assignments are colour coded according to J2Δ8 (black) and M2/M3 (red) single strand A. 100 µM sample in 150 mM KCl, 25 mM Tris, 4.5 mM $MgCl_2$, 200 µM DSS and 10% D_2O at pH 8.3 recorded at 288 K with 512 scans at 700 MHz. B. 100 µM sample in 150 mM KCl, 25 mM Tris, 4.5 mM $MgCl_2$, 5 mM DTT, 100 µM DSS and 8% D_2O at pH 8.3 recorded at 288 K with 512 scans at 600 MHz.

4.3.8 M2 Single Strand

While an RNA single strand (M2) was included in assay-based studies of RhlB's RNA substrate preferences, utilizing it for NMR experiments was not planned, as the single strand was not expected to give rise to imino proton resonances in the ^{1}H spectrum. But since secondary structure predictions of M2 suggested the possible formation of short stem-loop structures (see appendix 8.2.2), an ^{1}H 1D spectrum was recorded nonetheless with a 300 µM ^{15}N labelled sample in 50 mM KCl and 25 mM KH_2PO_4/K_2HPO_4 at pH 6, followed by an ^{1}H,^{15}N-HSQC and ^{1}H,^{1}H-NOESY experiment. The corresponding spectra, depicted in Figure 38, show the clear detection of two narrow imino proton resonances as well as one broad signal between 12.3 and 13.3 ppm. Through the 2D experiments the two intense signals could be identified to be a guanine and a uridine, that are in neighbouring base pairs based on the NOE cross peak. The guanine could be unambiguously assigned to $G3_{M2}$ and considering a base pair formation with $C13_{M2}$ would result in a far more favourable loop size than with $C7_{M2}$, the neighbouring uridine had to be assigned to $U4_{M2}$. While the formation of a GU base pair involving $G3_{M2}$ was in principle possible, too, it could be discarded in this case, as non-canonical base pairs like GU or UU would have a more upfield chemical shift around 10-12 ppm. This assignment was in agreement with theoretical secondary structure predictions for the M2 strand using the web application mFold. Here, the energetically most stable conformation

assumed the same two base pairs $G3_{M2}$-$C13_{M2}$ and $U4_{M2}$-$A12_{M2}$ (see appendix 8.2.2). The transition to pH 8.3 buffer conditions did result in severe peak broadening for $U4_{M2}$, while $G3_{M2}$ remained stable (Figure 39). Despite the limited number of reporter signals it was concluded to add the M2 single strand to the set of RNA substrates further investigated via NMR spectroscopy.

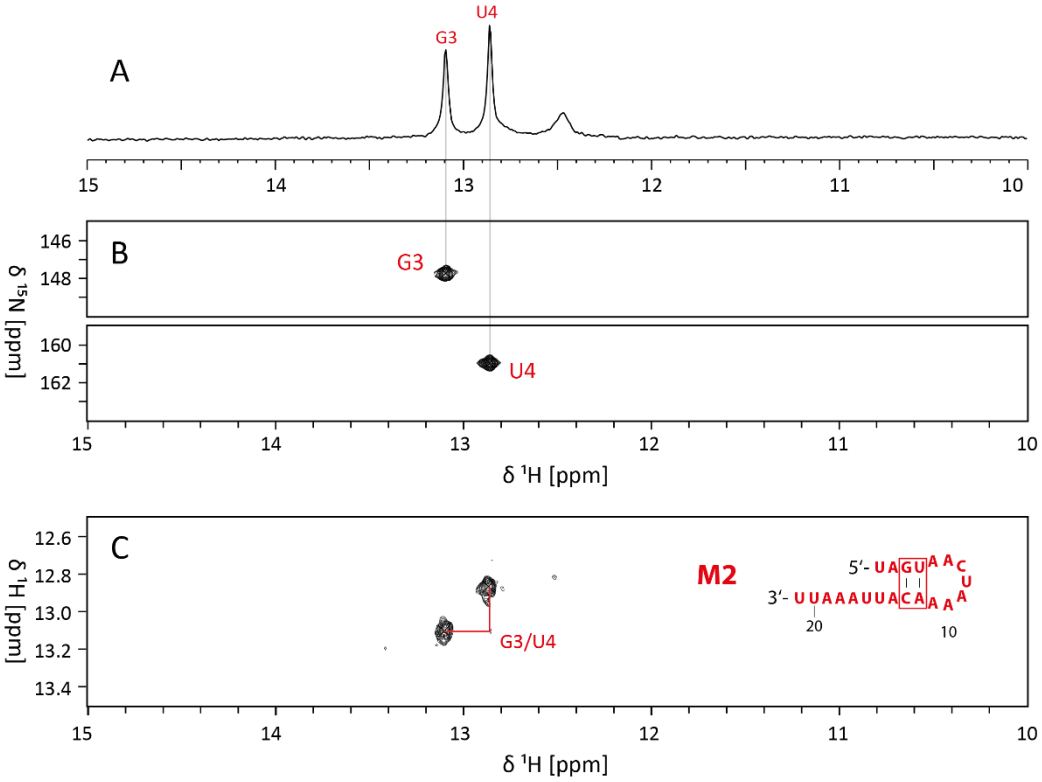

Figure 38. Imino resonance assignment of RNA single strand M2. A. Imino region of ^{1}H 1D spectrum of 300 µM unlabelled sample in 50 mM KCl, 25 mM KH_2PO_4/K_2HPO_4, and 10% D_2O at pH 6.2 recorded at 288 K with 256 scans at 700 MHz. B. $^{1}H,^{15}N$-HSQC spectrum of 160 µM selectively ^{15}N labelled sample (M2 G/U ^{15}N labelled) in 50 mM KCl, 25 mM KH_2PO_4/K_2HPO_4, and 10% D_2O at pH 6.2 recorded at 283 K with 16 scans at 800 MHz. C. Imino region of $^{1}H,^{1}H$-NOESY spectrum of 300 µM unlabelled sample in 50 mM KCl, 25 mM KH_2PO_4/K_2HPO_4, and 10% D_2O at pH 6.2 recorded at 288 K with 256 scans at 700 MHz. The sequential walk is indicated by red lines.

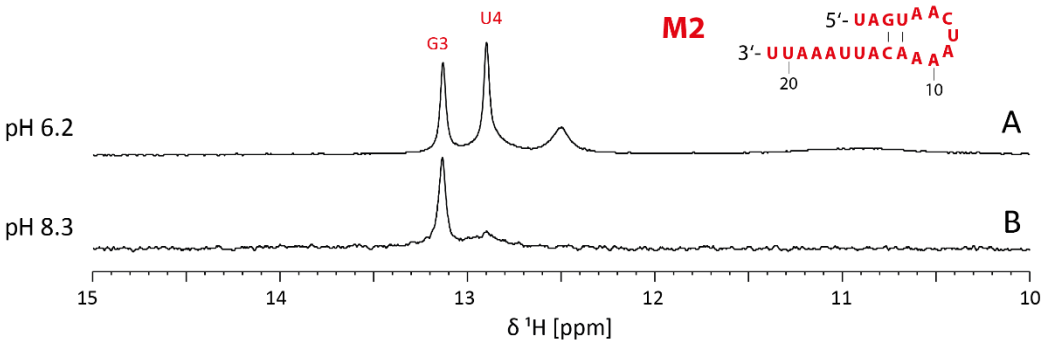

Figure 39. Imino regions of ^{1}H NMR spectra of RNA single strand M2 at different buffer conditions. A. 300 µM sample in 50 mM KCl, 25 mM KH_2PO_4/K_2HPO_4, and 10% D_2O at pH 6.2 recorded at 288 K with 256 scans at 700 MHz. B. 100 µM sample in 150 mM KCl, 25 mM Tris, 4.5 mM $MgCl_2$, 100 µM DSS and 7% D_2O at pH 8.3 recorded at 288 K with 320 scans at 700 MHz.

4.4 Discussion

In this chapter the design, production, and biochemical as well as NMR spectroscopic characteristics of 7 short RNA single strands and 6 heteroduplexes were analysed to determine their suitability as RNA substrates for NMR spectroscopic investigations of the RNA unwinding mechanism of DEAD-Box helicase RhlB. The aim of these investigations was to identify a set of short RNA duplex constructs with different strand properties such as blunt end, 3' single strand or 5' single strand overhang, which provide a sufficient set of assignable ^{1}H imino proton reporter resonances under the required buffer and measurement conditions to allow subsequent interaction studies with the helicase *via* NMR.

The overall synthesis and purification strategy of the RNA sequences via *in vitro* transcription was successful, although the unstructured nature of the <15 Nt RNA single strands made some of the constructs, in particular J2Δ16 and M2, more prone to degradation during purification compared to larger RNA constructs that adopt a stable secondary structure. The thermal stability of all investigated duplexes was between 37.5°C and 58.7°C and was approximately 5°C lower than predicted but could be increased with the addition of Mg^{2+} ions. It could be shown, that apart from J2Δ14-M2, all constructs adopted the proposed duplex with minimal to no alternate populations (less than 5 %). A multitude of NMR measurement conditions were examined thoroughly to balance requirements for temperature, pH and buffer conditions of RNA and helicase. A pH of 8.3 together with temperatures below 288 K, ionic strength of 150 mM and an organic Tris buffer system have been identified to be most favourable. While the specific requirement of NMR experiments at pH 8.3 increased the peak width and therefore reduced the overall resolution of the spectra compared to pH 6.2 measurements, the loss of individual signal intensity due to increased solvent exchange was considered acceptable.

In total 7 individual RNA sequences were successfully characterized, and NMR spectroscopic experiments confirmed that 4 of the planned 6 double-stranded RNA constructs as well as hairpin-forming M2 single strand were suitable for further investigations involving RhlB.

To reduce the net amount of individual RNA sequences to be synthesized, a toolbox approach was utilized, where the sequence of the M2 strand was kept constant, while the counter-strand sequences were adapted to produce duplexes with a variety of strand properties. This not only reduced the preparation time over all the constructs but also significantly facilitated the NMR assignment through common imino proton and nucleobase resonances. The NMR spectroscopic analysis however also revealed a drawback to this approach: if one RNA duplex exhibits loss of imino proton signals due to high solvent exchange rates of instable base pairs, this feature is also present in other

constructs with the same sequence. Due to the necessity of a high AU content almost all the constructs exhibited line broadening beyond detectability for the terminal AU base pairs. This inconvenience could be overcome for future projects by curating the RNA sequence more towards imino proton assignment by incorporating for example terminal GC base pairs.

Chapter 5 Structural and Functional Impact of RNase E on RhlB's Helicase Reaction

5.1 Expression and Purification of RhlB, RNase E (628-843) and RNase E (694-790)

5.1.1 Expression and Purification of RhlB

For the purification of RhlB a strategy combining Ni-NTA affinity chromatography and size exclusion chromatography was utilized. RhlB was therefore expressed with a C-terminal His_6-tag that would subsequently be cleaved off via TEV protease digestion.

The protein was expressed in *E. coli* BL21 cells as described in chapter 3.5.3 and initial expression tests informed the decision to perform large scale expressions at 37°C in TB medium with induction by 1 mM IPTG at an OD_{600} of ~1.5 and subsequent growth at 20°C for 16 h before harvest. The cells were lysed via high-pressure homogenization, centrifuged and the supernatant treated with 3% polyethyleneimine (PEI) to precipitate endogenous nucleic acids. The protein was subsequently purified from lysate in three purification steps exemplified in Figure 40. Firstly, the His_6-tagged protein was immobilized on the Ni-NTA column and eluted with an increasing gradient of up to 500 mM imidazole. Corresponding elution fractions containing the 48.8 kDa sized His_6-RhlB were identified (fractions 7-10) with 10% SDS-PAGE, pooled and the His_6-tag cleaved off by incubation with TEV protease. Simultaneous dialysis against buffer A reduced the imidazole concentration back to the initial 10 mM. A second Ni-NTA affinity chromatography was performed to remove both the cleaved tag and other protein impurities that coeluted from the first Ni-NTA column. Here, the tag-free helicase did not bind to the column and eluted from the column right away (fractions 2-8) while the cleaved His_6-tag together with highly affine proteins remained on the column and were eluted afterwards with an increased imidazole gradient. As shown in Figure 40B in the elution fraction 17 a small percentage of helicase did remain uncleaved. This fraction was also separated due to its affinity to the column. Since RhlB was essentially pure after the second Ni-NTA affinity chromatography the final size exclusion chromatography (SEC) was only used to inspect RhlB's folding homogeneity and to eventually remove oligomerized or aggregated protein. The concentrated RhlB solution was therefore loaded onto a Superdex 200 column and eluted with 1.2 column volumes SEC buffer. Figure 40C nicely exemplifies the even and homogeneous elution profile of RhlB with a single peak and merely a minute fraction of protein aggregation in the column's exclusion volume (fraction 19). The fractions 43 to 53 contained high amounts of pure protein and thus were combined for further experiments. Expressions in TB medium yielded up to 43 mg protein per litre medium.

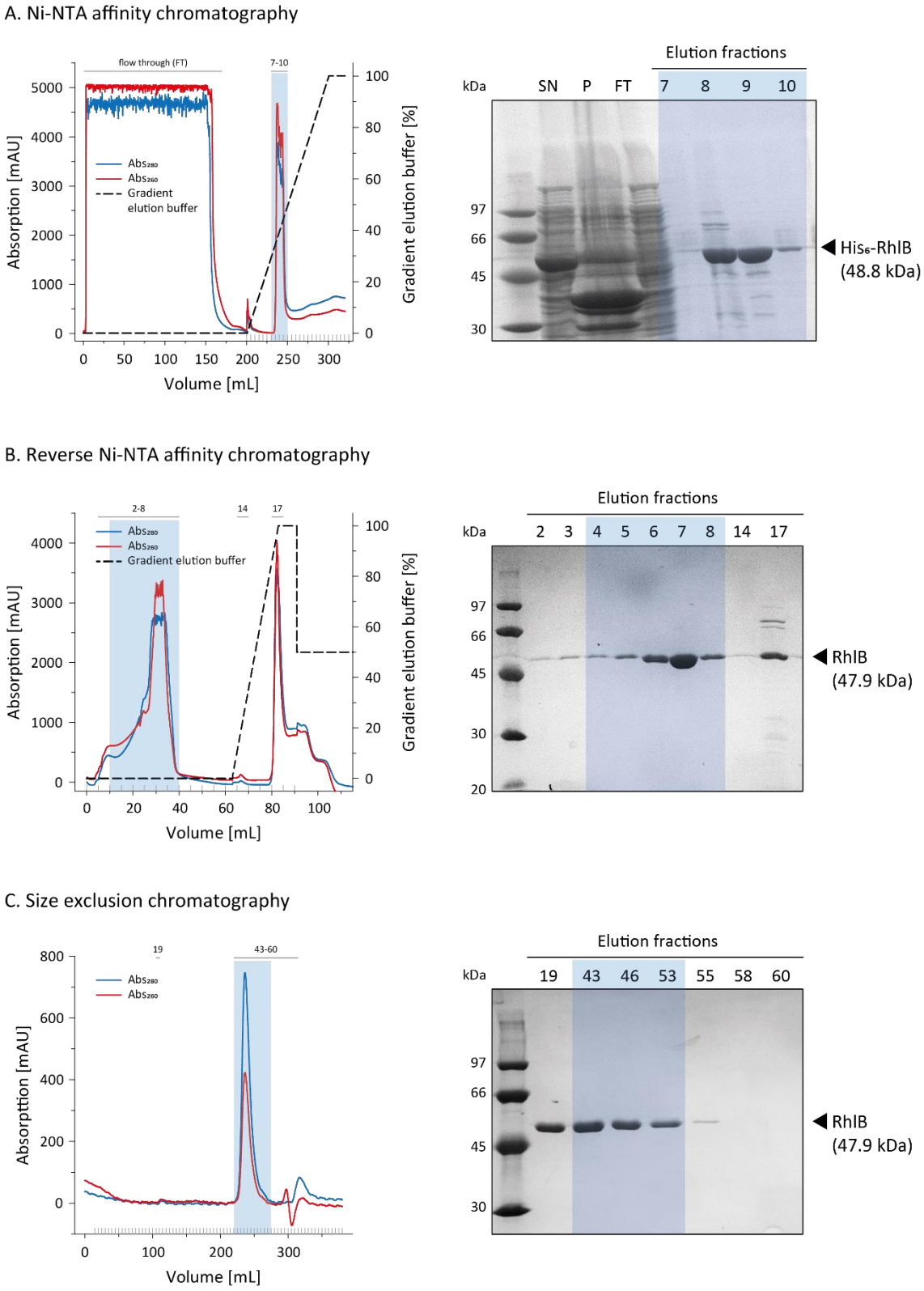

Figure 40. Exemplary purification of RhlB with chromatograms and 10% SDS-PAGEs for Ni-NTA (A.), reverse Ni-NTA (B.) and size exclusion chromatography (C.). The chromatograms show absorption at 260 nm (red) and 280 nm (blue) as well as the percentage of buffer B for gradient elution (dashed lines) as a function of elution volume. Numbered 5 mL fractions investigated via SDS-PAGE are also indicated in the top part of the chromatograms and fractions chosen for further purification are highlighted in blue. A. SDS-PAGE includes lysis fractions of supernatant (SN) and pellet (P) as well as the flow through (FT) of the lysate on the Ni-NTA column.

Although all measures were taken to work RNase free during protein purifications, residual RNases endogenous to the *E. coli* host cells could remain in the sample and cause a freshly purified RhlB probe to degrade substrate RNA. Since catalytic amounts of RNases are already sufficient to induce RNA degradation, a visual assessment via SDS-PAGE did not provide the necessarily control for RNase impurities. An RNase test was therefore conducted for all freshly purified RhlB batches, to check for possible RNase contaminations in presence of an RNA strand. If the respective protein sample showed a reduction in RNA band intensity in an RNA-PAGE after incubation with the single stranded RNA (here M2 was used) compared to a negative control, RNase induced degradation was assumed. Several RhlB samples did in fact show a positive RNase test, meaning RNA degradation was observed (see Figure 41B). To separate RhlB from possible RNases a heparin affinity chromatography was performed. It was hypothesised that heparin, a highly sulphated and densely positively charged polysaccharide that can function as a nucleic acid mimetic, would bind both the helicase as well as possible RNases, as it has been shown for DNA binding proteins in the past[166–168]. A slow gradual increase of the salt concentration in the buffer would then elute both proteins in different fractions and thereby separate them, assuming a difference in their respective affinity for the column.

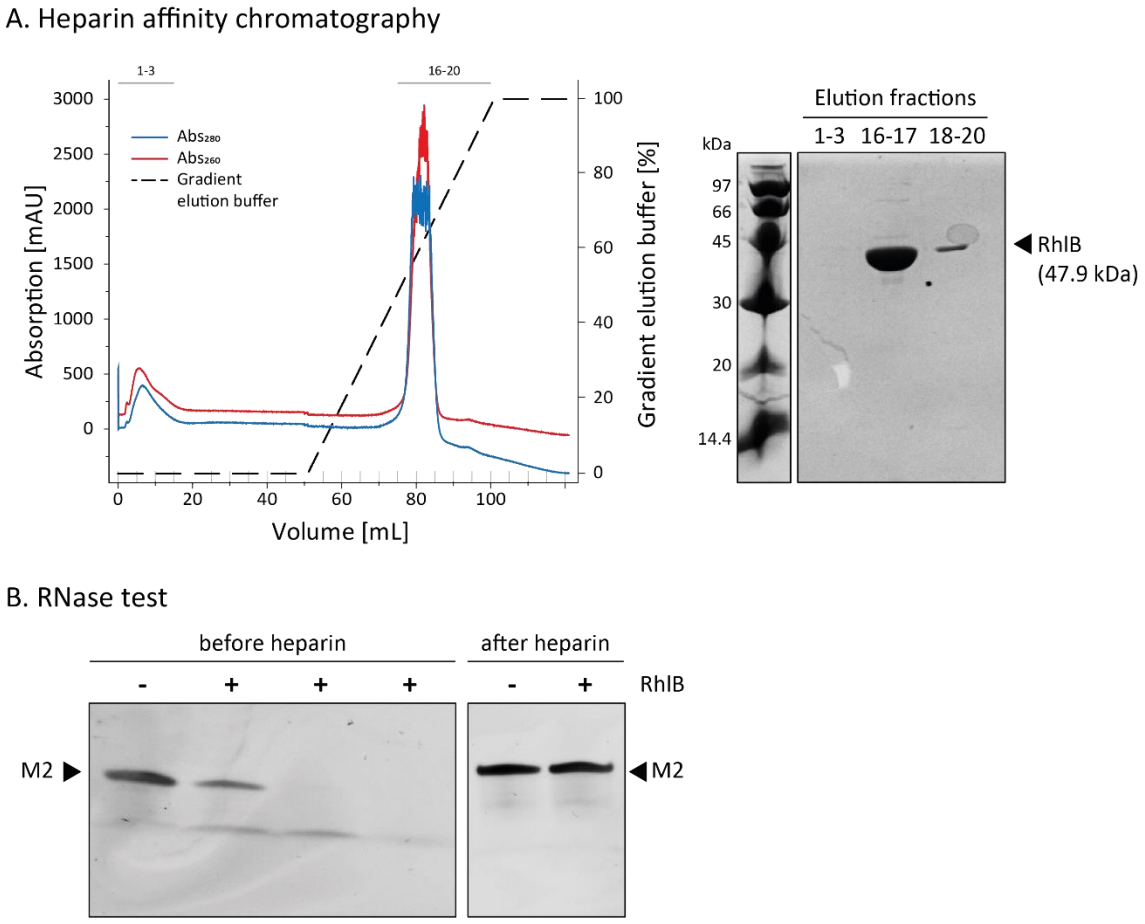

Figure 41. RhlB's RNase contamination analysis and subsequent purification via Heparin affinity chromatography. A. Heparin affinity chromatogram showing absorption at 260 nm (red) and 280 nm (blue) as well as gradient of elution buffer as a function of elution volume. 10% SDS-PAGE was used to determine elution fractions that contained RhlB. B. RNase contamination test as described in section 3.5.5 for three RhlB samples before and one sample after heparin affinity chromatography. Incubated RNA samples were analysed via denaturing RNA-PAGE for reduced RNA band indicative of RNase induced degradation.

The corresponding heparin affinity chromatogram in Figure 41A shows the small absorption increase directly after injection (at 0 mL) and a more pronounced peak during the elution with approximately 50% elution buffer. An SDS-PAGE confirmed the presence of RhlB in the elution fractions 16-20 and a subsequent RNase test with those fractions exhibited no more RNase activity, confirming the successful removal of RNase contaminations.

5.1.2 Expression and Purification of RNase E (628-843) and RNase E (694-790)

Both RNase E fragments were expressed with an N-terminal His_6-tag. Preliminary test expressions highlighted the effect the two different expression vectors had on the optimized expression conditions. While RNase E (628-843) expression with the pET11A vector was optimally induced with 1 mM IPTG at OD_{600} of 1.5 in TB medium, 0.4 mM IPTG were sufficient for induction of RNase E (694-790) in the pET21A vector, and higher

concentrations even reduced the total protein yield. It proved beneficial for both protein fragments to stop the expression 2.5-3 h after induction at 20°C, because prolonged expression periods showed increased protein degradation. Purifications of RNase E fragments (628-843) and (694-790) were successfully executed analogue to RhlB and are exemplified in Figure 42 and Figure 43 and yielded pure protein in both cases.

For RNase E (628-843), the Ni-NTA chromatography included a LiCl buffer washing step after application of the lysate to the column to remove protein bound RNA. The SDS-PAGE fractions of this washing step did not show any protein being stripped of the column but instead a strong absorption peak at 260 nm indicative of eluted nucleic acids. The subsequent size exclusion chromatography unfortunately revealed a persisting fraction of nucleic acids with a 260 nm absorption peak (fractions 1-8) which indicated that the RNA removal was not exhaustive. During purification it also became apparent that RNase E (628-843) was slightly unstable and tended to slowly fragment over time. This behaviour was also reflected in RNase E (628-843)'s high instability index computed by the protParam tool of the Expasy web suite that classified the protein as unstable. To reduce any protein degradation to a minimum, the protein was constantly stored on ice during purification and flash frozen at -80°C for long term storage. Expressions in TB medium nonetheless provided good yields of up to 62 mg protein per litre medium.

The SDS-PAGEs of RNase E (628-843) and (694-790) also revealed the peculiar running behaviour of both proteins in the gel, which will be inspected in detail in a following section of this chapter.

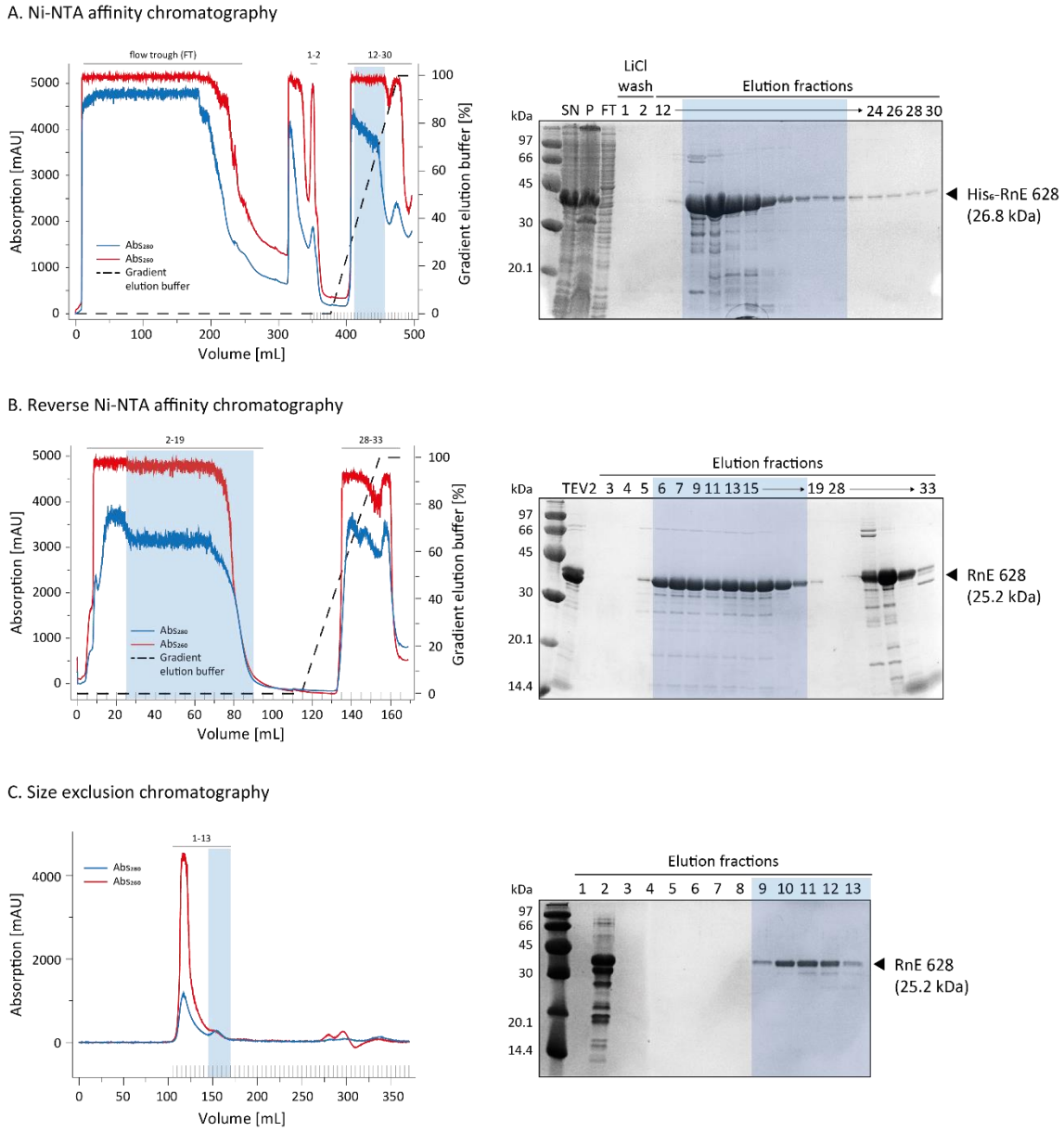

Figure 42. Exemplary purification of RNase E (628-843) with chromatograms and 15% SDS-PAGEs for Ni-NTA (A.), reverse Ni-NTA (B.) and size exclusion chromatography (C.). The chromatograms show absorption at 260 nm (red) and 280 nm (blue) as well as the percentage of buffer B for gradient elution (dashed lines) as a function of elution volume. Numbered 5 mL fractions investigated via SDS-PAGE are also indicated in the top part of the chromatograms and fractions chosen for further purification are highlighted in blue. A. SDS-PAGE includes lysis fractions of supernatant (SN) and pellet (P) as well as the flow through (FT) of the lysate on the Ni-NTA column. B. SDS-PAGE includes TEV protease digested sample before reverse Ni-NTA chromatography (TEV).

The purification of RNase E (694-790) was carried out with the same chromatographic steps as the other proteins but required some buffer optimization. The original buffer compositions that were utilized for RhlB and RNase E (628-843) resulted in a significant degradation of RNase E (694-790). This was prevented by slight adjustment of the pH from 8.3 to 8.0 and a decrease of the ionic strength of the buffer from 500 mM NaCl to 200 mM NaCl and 100 mM KCl, as it was recommended by Worrall *et al.*[17]. Figure 43 shows an exemplified purification under those optimized buffer conditions.

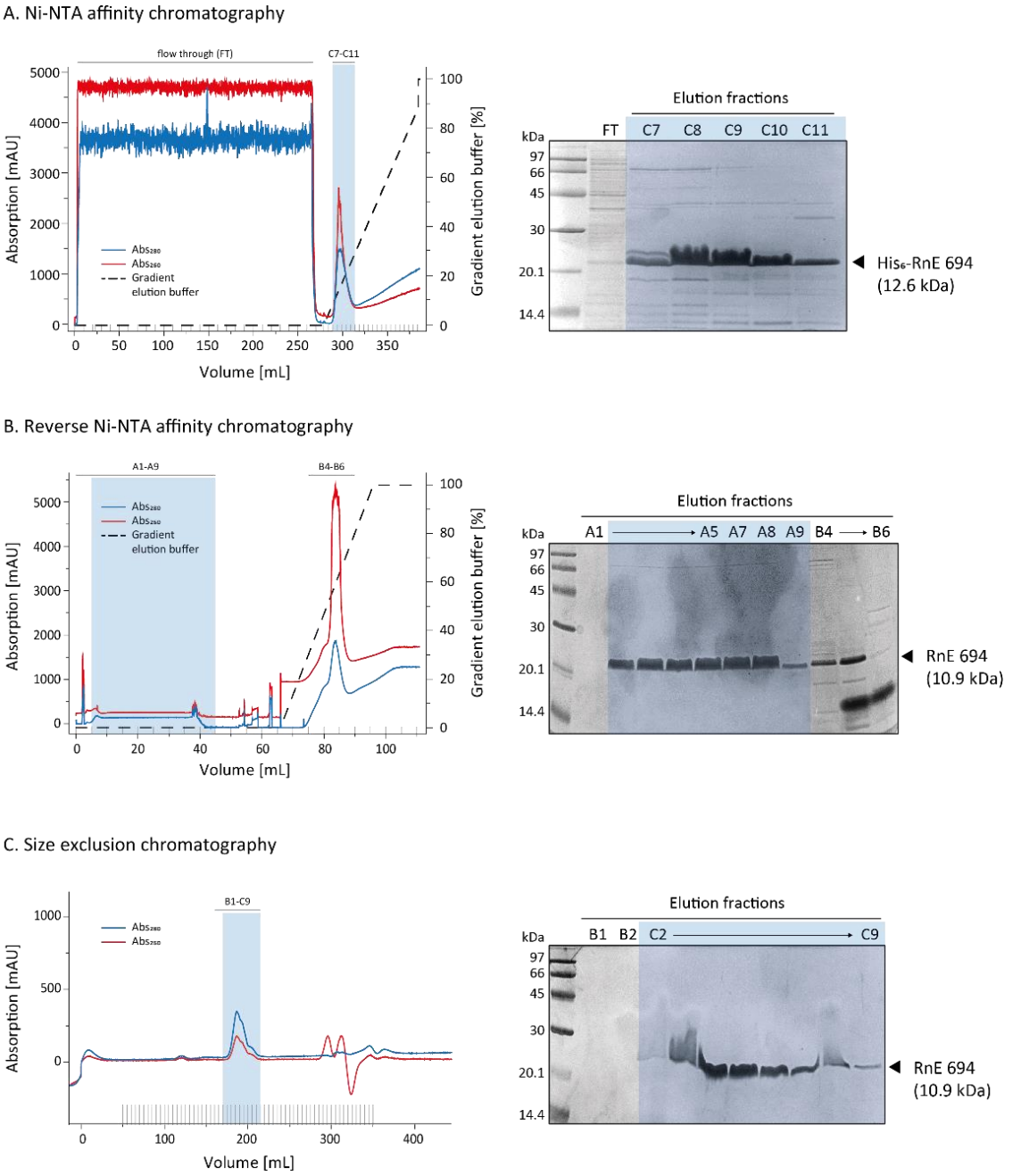

Figure 43. Exemplary purification of RNase E (694-790) with chromatograms and 15% SDS-PAGEs for Ni-NTA (A.), reverse Ni-NTA (B.) and size exclusion chromatography (C.). The chromatograms show absorption at 260 nm (red) and 280 nm (blue) as well as the percentage of buffer B for gradient elution (dashed lines) as a function of elution volume. Numbered 5 mL fractions investigated via SDS-PAGE are also indicated in the top part of the chromatograms and fractions chosen for further purification are highlighted in blue. A. SDS-PAGE includes flow through (FT) of the lysate on the Ni-NTA column. C. Lopsided running behaviour in the SDS-PAGE is attributed to partially overloaded lanes.

The His_6-tagged protein was successfully eluted from the Ni-NTA column in high concentration and residual protein impurities could be completely removed by the second Ni-NTA chromatography following the TEV cleavage (Figure 43B). The size exclusion chromatography gave rise to one elution peak containing exclusively RNase E (694-790) and sufficient amounts of protein could be yielded (up to 27 mg protein per litre medium).

5.1.3 Folding Behaviour of RNase E Fragments

It became apparent during purification of RNase E (628-843) and RNase E (694-790) that both protein fragments exhibit an unexpected SDS-PAGE running behaviour. As indicated by arrows in the SDS-PAGEs in Figure 44A and B did both proteins migrate significantly higher in the gel as their calculated mass suggested: RNase E (628-843), which has a molecular mass of 25.2 kDa, migrated at the height of a protein with a size of approximately 35-40 kDa and RNase (694-790) ran approximately 10 kDa higher than its molecular mass of 10.9 kDa suggested. MALDI-TOF mass spectrometry measurements were performed to verify the mass: A sample of His_6-tagged RNase E (628-843) measured 26865.3 g mol^{-1} (theoretical mass 26775.3 g mol^{-1}) and a ^{13}C labelled sample of RNase E (694-790) measured 11628.3 g mol^{-1} (theoretical mass 11595.0 g mol^{-1}). The additional mass observed could be attributed to the binding of water or buffer ions, which is typically observed in mass spectrometric analysis of proteins. The mass spectrometry measurement therefore confirmed the correct molecular mass of the purified fragments. It was consequently suspected that both proteins are disordered, and that their unfolded and extended structure caused their retarded running behaviour in the gel.

As pointed out previously, the CTD of RNase E is proposed to be natively disordered and interspersed with short linear segments that exhibit a strong folding propensity[83,127]. To corroborate this for our specific protein fragments, we utilized the webtool PONDR® VL-XT (Predictor of naturally disordered regions). This software provides a disorder/order score for protein chains based on different attributes of the amino acid sequence such as hydropathy and sequence complexity and utilizes three neural networks that were trained with experimental data of proteins with disordered regions. PONDR® outputs are between 0 and 1 for each amino acid and are then smoothed over a sliding window if 9 amino acids. If the residue value exceeds or matches a score of 0.5, the residue is considered disordered[169,170]. The results of those predictions are plotted in Figure 44C. Both RNase E (628-843) and RNase E (694-790) were predicted to be completely disordered with only the terminal amino acids showing a score below 0.5. Those results became even more striking when opposed with the PONDR plot of DEAD-Box helicase RhlB, a member of a protein family who's conserved RecA-like domain structure is well documented throughout literature[15,21,113,171]. And in fact, for RhlB the ordered regions dominate the sequence.

Another type of folding analysis by Uversky showed that an overall prerequisite for intrinsically disordered proteins (IDP) is a combination of low mean hydrophobicity (also called hydropathy) and high net charge. Those two features lead to a low tendency for protein compaction and strong electrostatic repulsion and ultimately result in an absence of a defined and compact structure[172,173]. Figure 44D shows the charge-hydropathy plot for all three investigated proteins. Both RNase E protein fragments clearly fall onto the side of natively unfolded proteins due to their high charge density and scarcity of hydrophobic residues, while RhlB resides in the area designated to folded proteins. A set

of 54 disordered and 105 ordered proteins was also referenced in this plot to highlight the accuracy with which a protein can be assigned to either natively folded or unfolded states.

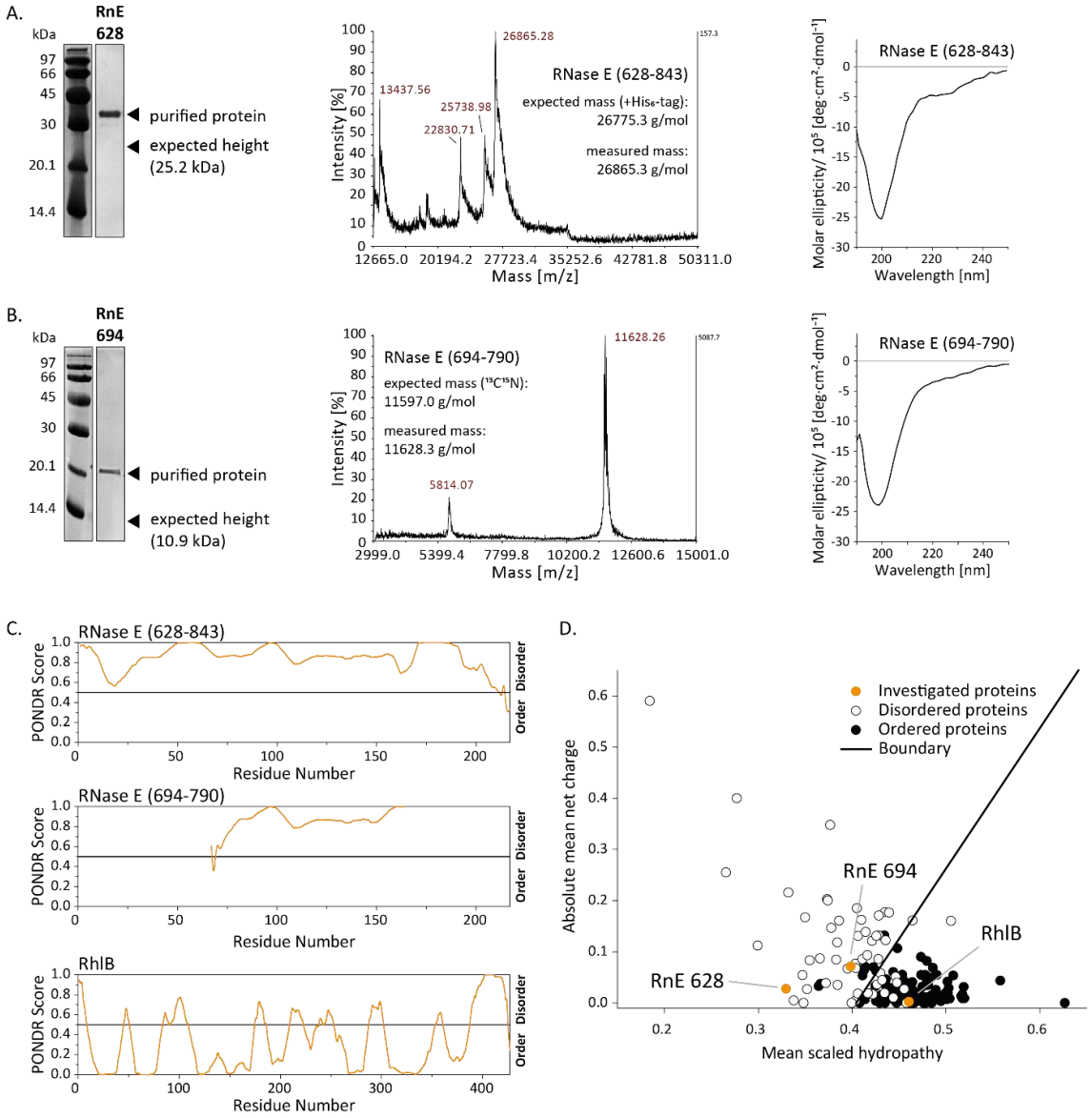

Figure 44. Secondary structure and folding analysis of RNase E (628-843) and RNase E (694-790). A. SDS-PAGE, mass spectrometric analysis and CD spectrum measured for RNase E (628-843). Highlighted in 10% SDS-PAGE with arrows are the expected and actual running height of the purified protein. MALDI TOF mass spectrometry result acquired for His_6-tagged protein. Expected and measured mass, as noted in the figure, are in overall good agreement. CD spectrum shows molar ellipticity for wavelengths of 190-250 nm. B. SDS-PAGE, mass spectrometric analysis and CD spectrum measured for RNase E (694-790). Highlighted in 15% SDS-PAGE with arrows are the expected and actual running height of the purified protein. MALDI TOF mass spectrometry result acquired for $^{13}C,^{15}N$ labelled protein. Expected and measured mass, as noted in the figure, are in overall good agreement. CD spectrum shows molar ellipticity for wavelengths of 190-250 nm. C. Prediction of naturally disordered regions of RNase E (628-843), RNase E (694-790) and RhlB as a function of their respective amino acid residue sequence predicted using the PONDR® VL-XT webtool (pondr.com). If a residue value exceeds or matches a score of 0.5, the residue is considered disordered. D. Charge-hydropathy plot of RNase E (628-843), RNase E (694-790) and RhlB in relation to several exemplary naturally unfolded (white circles) and folded (black circles) proteins. Plotted is the absolute mean net charge against the mean scaled hydropathy. Mean net charge is defined as number of net charges divided by number of amino acids and hydrophobicity was calculated using the Expasy Protscale tool with a Kyte & Dolittle approximation, a window size of 5 and normalized to a scale of 0 to 1. The black line represents the border between extended IDPs and proteins with a defined structure as set by Uversky [172,173].

To support those predictions with experimental data we measured CD spectra for both RNase protein constructs. Since the different secondary structure elements (α-helices, β-sheets, random coils) exhibit characteristic ellipticity profiles in the range of 190 to 250 nm (see Figure 13) and the spectrum of a protein is a linear combination of its individual secondary structure elements, it is possible to derive the overall secondary structure composition of a protein from the CD spectrum. Figure 44A and B show the corresponding spectra for RNase E (628-843) and RNase E (694-790). Both proteins exhibit a CD spectrum with a significant ellipticity decrease around 200 nm that is characteristic for a peptide chain in random coil conformation. RNase E (628-843) shows an additional small dip at 220 nm, indicating a small fraction of β-helical fold in the structure. We can therefore assume that they both adopt a predominantly disordered structure.

These results were later confirmed conclusively for RNase E (694-790) with an $^{1}H,^{15}N$ correlation spectroscopy experiment of $^{13}C,^{15}N$ labelled RNase E (694-790) measured in 150 mM KCl and 25 mM HEPES at pH 6.5. At those conditions but also in NMR buffer at pH 8.3 the protein showed only a small chemical shift dispersion from 8.7 to 8.1 ppm for the backbone amide resonances, which is associated with a disordered protein conformation (see Figure 45)[174]. The NMR spectroscopic analysis of RNase E (694-790) will be discussed in more detail in section 5.2.3.

Taken together, it can be confirmed from the experimental data that both RNase E fragments adopt a predominantly unfolded structure, although a small fraction of α-helix or β-sheet could not be excluded for RNase E (628-843) based on the CD spectra.

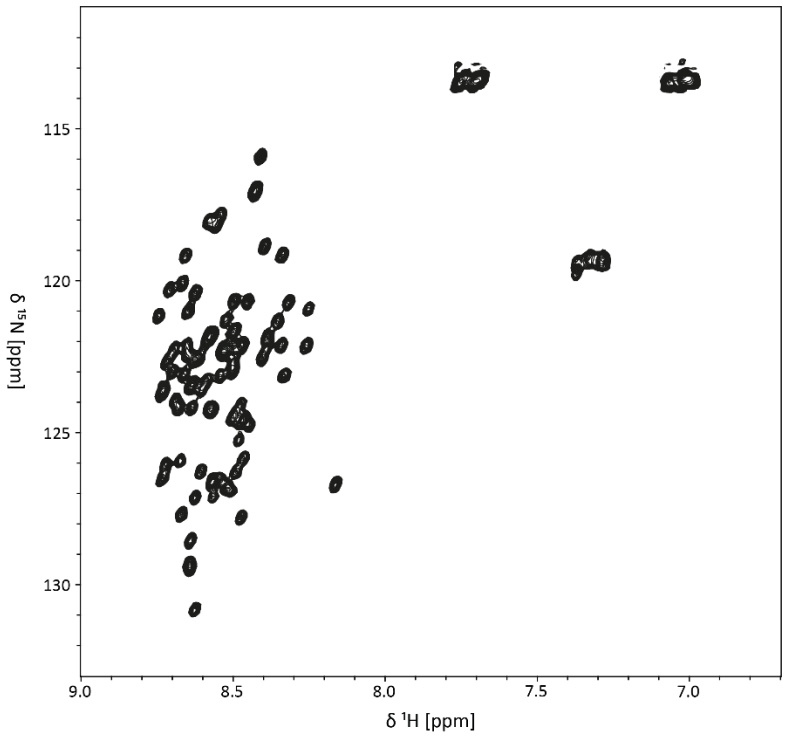

Figure 45. ^{1}H, ^{15}N HSQC spectrum of 50 µM $^{13}C,^{15}N$ labelled RNase E (694-790). The spectrum was recorded with 24 scans in 150 mM KCl, 25 mM HEPES, 100 µM DSS and 10% D_2O (pH 6.5) at 278 K and 800 MHz. The small chemical shift dispersion of backbone amide resonances points towards a predominantly disordered protein structure.

5.1.4 RNA Binding

Before progressing to further investigations, it had to be confirmed whether – under the established purification protocol – RhlB has adopted a functional fold that could bind RNA to undergo the unwinding reaction. In addition, the binding ability of the larger RNase E fragment (628-843) encompassing the RNA binding sites RBD and AR2 was to be verified. To assess the RNA binding ability of both proteins electrophoretic mobility shift assays (EMSA) were performed. Those assays utilize the difference in electrophoretic mobility of unbound RNA vs protein-bound RNA, where the band of the less mobile complex probe will be "shifted" up on the gel compared to the free RNA.

Figure 46A shows the EMSA of the RNA construct J2h-M2 both alone and mixed with RhlB or RhlB/RNase E (628-843). As we can see, the band of the main and secondary conformations of the free RNA duplex (gel lane 1) are strongly decreased in the presence of RhlB while an RNA population is appearing in the gel pocket (gel lane 3-5). The Coomassie staining of the same gel shows that at the given conditions RhlB does not migrate into the gel but remains in the gel pocket. Despite extensive efforts to adjust the EMSA conditions in a way that allows RhlB to migrate into the gel, the protein remained in the pocket. Nonetheless, the stained RNA in the pocket provides initial evidence that RNA is bound by the protein. To exclude the possibility that the RNA was degraded, a separate sample with both RNA and RhlB was shortly heated to 95°C to denature the protein before loading it to the gel (gel lane 6). This sample resulted in full recovery of the – now homogenously folded - free RNA, therefore further corroborating that the RNA was bound to the helicase. 35 µM RhlB (23-fold excess over RNA) was not sufficient to fully bind the J2h-M2 duplex, at a higher protein concentration of 160 µM however complete RNA binding was achieved (Figure 46B).

The addition of ATP or non-hydrolysable analogue AMP-PNP (3 mM) to RhlB was also examined. While the addition of ATP would allow the helicase to undergo a complete helicase reaction, ATP hydrolysis and therefore RNA release is inhibited with AMP-PNP and so it is expected to form a ternary long-lived complex with RNA and RhlB as has been reported for other DEAD-Box helicases[122]. In the EMSA the sample containing AMP-PNP (gel lane 4) showed approximately the same amount of complex-bound RNA (84%) as the sample without any nucleotide (gel lane 3, 81%). This suggests that either AMP-PNP is not bound by the helicase or the binding of both RNA and AMP-PNP does not significantly affect the helicases affinity for RNA, as has been demonstrated for some members of the DEAD-Box helicase family (see section 1.3.3). With the addition of ATP (gel lane 5), less bound RNA was detected (59%), indicating that in the presence of ATP the helicase can complete the full unwinding reaction cycle and release some of the RNA duplex again.

In the same EMSA, the sample containing the RhlB/RNase E (628-843) complex showed full RNA binding (gel lane 7). While we cannot distinguish from this gel, which protein is binding the RNA substrate, the EMSA in Figure 46C definitively confirmed that RNase E (628-834) fully binds the RNA duplex at a concentration of >100 µM. Interestingly, heat

denaturation did not release the free RNA population from the RhlB/RNase E (628-843) complex (Figure 46gel lane 8) as it did for RhlB alone. This nicely highlights how the unstructured RNase E (628-843) binds RNA rather trough electrostatic interactions within its arginine-rich RNA binding sites than through secondary structures. Hence its RNA binding is not affected by heat-induced unfolding. A subsequent PCI extraction of the EMSA sample allowed a full recovery of the free RNA, verifying that RNA was not degraded (data not shown).

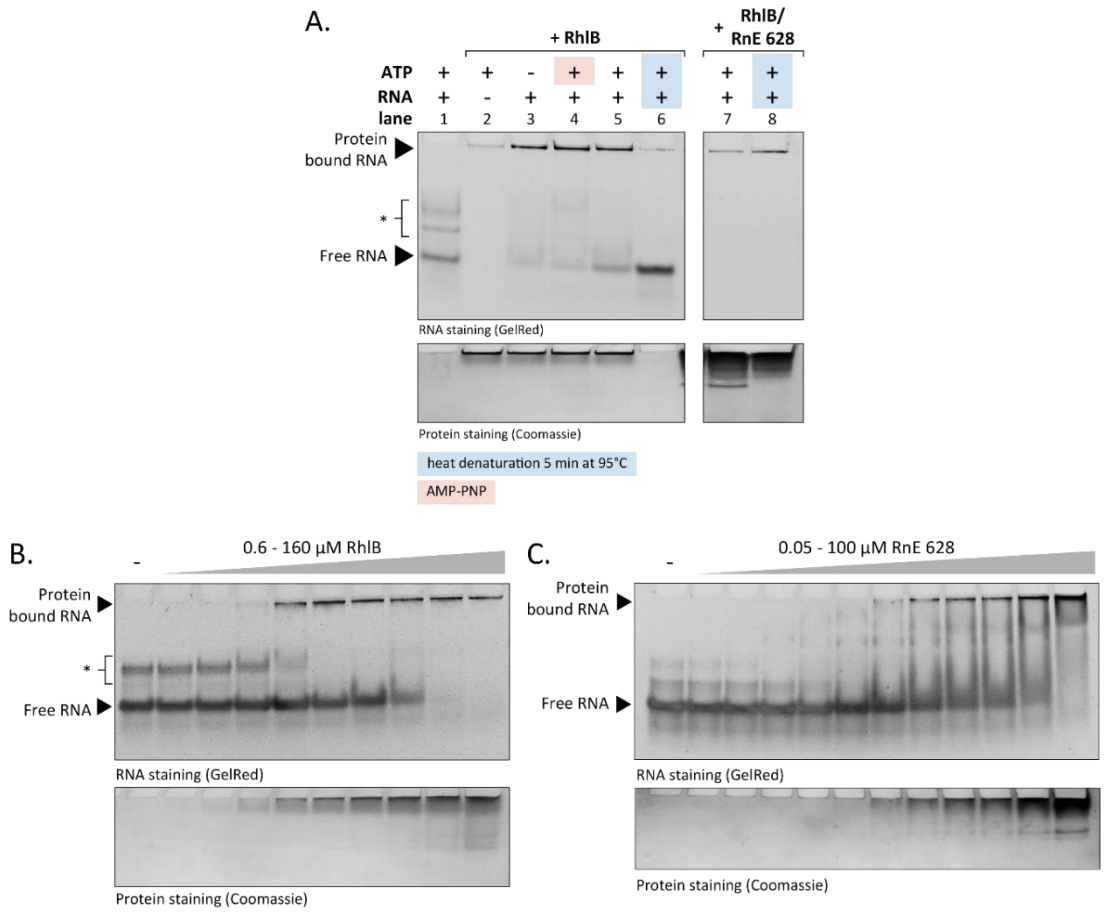

*Figure 46. Electrophoretic mobility shift assays (EMSAs) of RhlB and RNase E (628-843) with J2h-M2 RNA substrate. A. EMSA of 1.5 µM J2h-M2 mixed with or without 35 µM RhlB, 140 µM RNase E (628-843) and 3 mM ATP/AMP-PNP. Samples were mixed in EMSA buffer and incubated for 15 min at 4°C before 5 µL of each sample was loaded onto the discontinuous gel (6% + 10%). Gel was run at 0.5 W and 4°C for 5h and stained with GelRed and Coomassie. Samples for lanes 6 and 8 were heated to 95°C for 5 min right before loading onto the gel. Main conformation of J2h-M2 and protein-bound RNA are indicated with arrows and minor folding populations of J2h-M2 were marked with an asterisk * (refolding of the duplex was skipped prior to performing the EMSA). B. EMSA of 1.5 µM J2h-M2 with increasing concentration of RhlB (0.6 – 160 µM) in absence of ATP, procedure conducted as above (A). C. EMSA of J2h-M2 with increasing concentration of RNase E (628-843) (0.05 – 100 µM), procedure conducted as above (A).*

5.2 RhlB-RNase E Interaction

The next step was to investigate the interaction of RhlB with RNase E (694-790) using NMR spectroscopy. Here the overall dynamic, structure and conformation of both individual proteins were examined and the effects of the protein interaction on both binding partners was studied. For RNase E (694-790), we inspected whether the binding to RhlB would induce the formation of secondary structures accompanying an induced fit disorder-to-order transition, as it has been proposed for this RNase E fragment[7,19,83]. For that, both proteins were expressed ^{15}N labelled (and ^{13}C labelled in case of RNase E (694-790)) using M9 minimal medium and purified as before. Since RhlB with its 47.9 kDa is of considerable size with respect to NMR spectroscopic analysis, a full assignment would require a more elaborate assignment strategy and involve single domain constructs and selective isotope labelling. However, the focus of this thesis was kept on the RNA centred NMR analysis, and the full backbone assignment of both proteins was therefore considered beyond the scope of this project.

5.2.1 Probing RhlB-RNase E Binding

Firstly, the binding of RhlB to both RNase E fragments had to be confirmed. RhlB has been previously shown to form a 1:1 complex with RNase E (628-843) with an association constant in the region of 50 nM and was successfully copurified with different fragments of RNase E including (694-790)[17,19,20].

To attest the binding of RhlB to the minimal binding fragment RNase E (694-790) we performed analytical size exclusion chromatography of each protein individually and in 1:1 mix. The respective chromatograms of 100 µM protein samples are depicted in Figure 47A. RhlB eluted at a volume of 15.1 mL which is in very good agreement with the calibration curve, from which the 47.9 kDa monomeric protein was calculated to elute at around 15.1 mL. Interestingly, an elution volume of 15.1 mL could also be observed for RNase E (694-790). Considering the previously discussed unstructured nature of the RNase E fragment, this early elution it is not surprising and can be ascribed to the disordered nature of the 10.9 kDa protein (expected elution volume of globular protein 17.9 mL). In the chromatogram of both proteins, we observed a distinct peak shift towards a smaller elution volume (13.1 mL) which is indicative of complex formation. The corresponding SDS-PAGE confirmed the presence of RhlB and RNase E (694-790) in the elution peak. For the interaction with RNase E (628-843) we used a pull-down assay. Therefore His-tagged RNase E (628-843) was immobilized on a Ni-NTA affinity chromatography and untagged RhlB was subsequently loaded onto the column. If both proteins interact, RhlB would remain on the column despite not being tagged and both proteins would be eluted together as soon as an increased concentration of elution buffer displaced RNase E (628-843) from the column. In the chromatogram of this assay, shown

in Figure 47B, a large elution peak could be detected, which contained both protein fractions. With that, the binding of both RNase E fragments to RhlB could be confirmed.

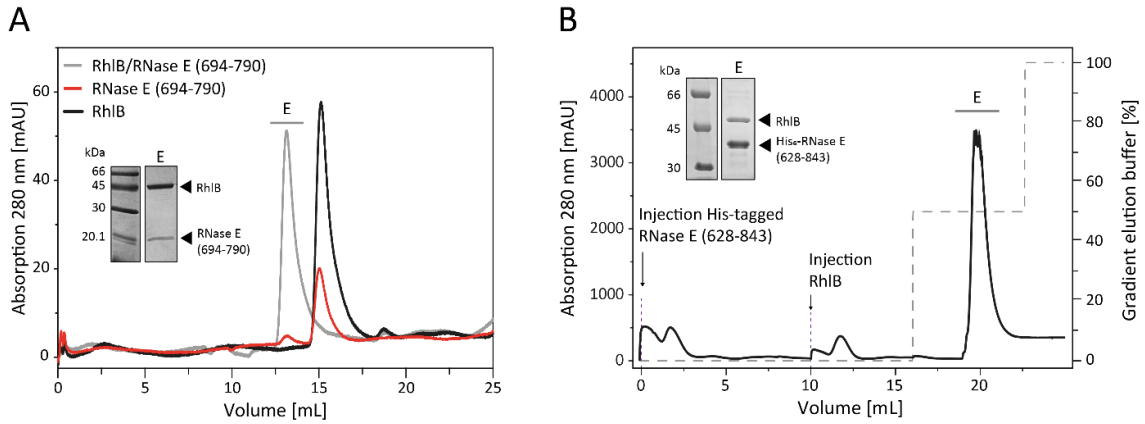

Figure 47. Binding analysis of RhlB with RNase E (628-843) and RNase E (694-790). A. S200 analytical size exclusion chromatograms of 100 µM RhlB, RNase E (694-790) and a 1:1 mix of both proteins. A clear shift in peak profile toward a higher molecular weight species could be observed for the mixed sample. (B) Pull-down assay using His_6-tagged RNase E (628-843) with untagged RhlB. Both proteins were sequentially loaded onto a 5 mL Ni-NTA affinity column and co-eluted using a gradient of elution buffer. The presence of both proteins in the elution fraction as verified via SDS-PAGE.

5.2.2 NMR Spectroscopic Analysis of ^{15}N RhlB

Before studying the interaction of RhlB with RNase E (694-790) we inspected the overall structure and dynamic of ^{15}N labelled RhlB alone using 2D ^{1}H,^{15}N correlation spectroscopy. It was, against our expectations, not possible to obtain a sufficiently resolved spectrum of the helicase in pH 8 NMR buffer (150 mM KCl, 25 mM Tris, 5 mM DTT, pH 8.3). Under those conditions and at different temperatures we were only able to the detect side chain amide signals and a few individual backbone resonances with both 400 µM and 100 µM samples as represented in panel A of Figure 50. This apparent instability of RhlB could however be overcome by increasing the salt concentration to 450 mM KCl (panel B of Figure 50). It appeared that this instability of RhlB at 150 mM KCl was a phenomenon limited to the ^{15}N labelled protein though, since analytical size exclusion chromatograms revealed a second elution peak for the labelled protein, that was not present in any unlabelled sample of the helicase (Appendix Figure 69). The chromatograms also showed that an increased salt concentration produced homogenously folded ^{15}N RhlB matching the elution volume of the unlabelled protein. Based on the elution volume of this second population (13.1 mL), the formation of a full aggregate could be excluded, since aggregates would be eluted with the columns void volume of 8 mL and re-solubilization of the aggregate at higher salt concentration was considered unlikely. Although the calculated elution volume of 13.6 mL for a globular RhlB dimer was not matched exactly, the formation of an elongated or otherwise dynamic dimer was plausible. It was suspected that the altered composition of the M9 minimal medium caused the conformational heterogeneity of the helicase. In consideration of

these results was the NMR spectroscopic inspection of labelled RhlB performed in NMR buffer containing an increased KCl concentration of 450 mM.

Both ^{15}N HSQC and BEST-TROSY experiments were recorded for ^{15}N labelled RhlB. While the BEST-TROSY delivered spectra with a higher resolution, the ^{15}N HSQC exhibited a better sensitivity and therefore allowed for the detection of more resonance signals. Since this analysis did not aim at a resonance assignment, the ^{15}N HSQC was chosen for the subsequent inspection. The well dispersed backbone resonances of the ^{15}N HSQC spectrum (Figure 48) indicated that the helicase is folded with a defined structure, but an increased number of resonances around 8 ppm (^{1}H) also points towards structural disorder within the protein. Referencing the structure prediction model of RhlB (Figure 9) this disordered part is most likely to be attributed to the flexible C-terminal extension (CTE) as well as outward facing loop regions within the core domains of the protein.

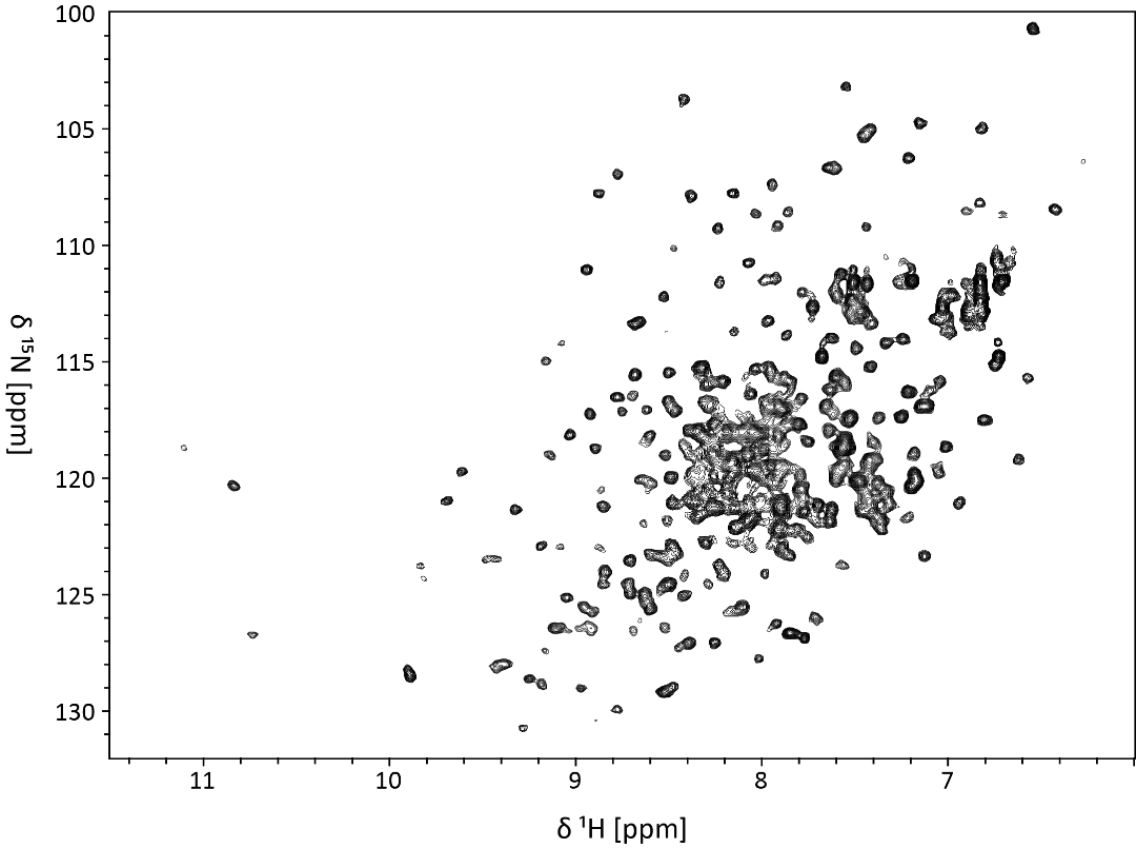

Figure 48. 2D ^{1}H,^{15}N HSQC spectrum of ^{15}N labelled RhlB. 100 µM sample in 450 mM KCl, 25 mM Tris, 5 mM DTT, 100 µM DSS and 10% D_2O at pH 8.3 were recorded at 305 K with 104 scans at 800 MHz. Resonances were referenced to DSS.

A peak count yielded approximately 261 backbone resonances out of 404 theoretically expected signals. The lack of these 35% of resonances and the varying signals intensity that can be observed throughout the visible backbone resonances provide evidence that parts of the protein are dynamic and the corresponding resonances in a slow or intermediate conformational exchange which causes them to be broadened beyond detectability. This behaviour could be attributed to the flexible loops connecting secondary structure elements in the core domains and in the C-terminal extension.

Furthermore, in absence of any substrate RhlB is expected to populate conformational ensembles in which the two core domains are separate and have some independent mobility. The three resonances at around 11 ppm (^{1}H) were assumed be the three tryptophan side chain resonances, of which one is located in the N-terminal and two in the C-terminal domain of the protein. The fact that they are all visible suggests that both domains are at least in part folded and rules out an alternative interpretation, where one domain of the helicase is folded while the other domain adopts a molten globule state and completely vanishes from detection, as it has been reported for other multi-domain proteins[175]. To give a more conclusive interpretations for a protein of this size, measurements with a fully deuterated sample would be required to eliminate spin relaxation.

5.2.3 NMR Spectroscopic Analysis of ^{15}N RNase E (694-790)

Next, the structure of RNase E (694-790) was analysed. Tests at different temperatures revealed that both BEST-TROSY and ^{15}N HSQC experiments yielded the highest sensitivity at 278 K. Panel A of Figure 49 shows the corresponding NMR spectrum of the BEST-TROSY. The poor chemical shift dispersion in the proton dimension, which is typical for proteins in an unfolded or coiled coil conformation, supports the reasoning for the natively disordered structure of the protein. After identifying that the RNase E fragment exhibits an unstructured nature via CD spectroscopy and analytical gel filtration chromatography, those results were confirmed via NMR spectroscopy.

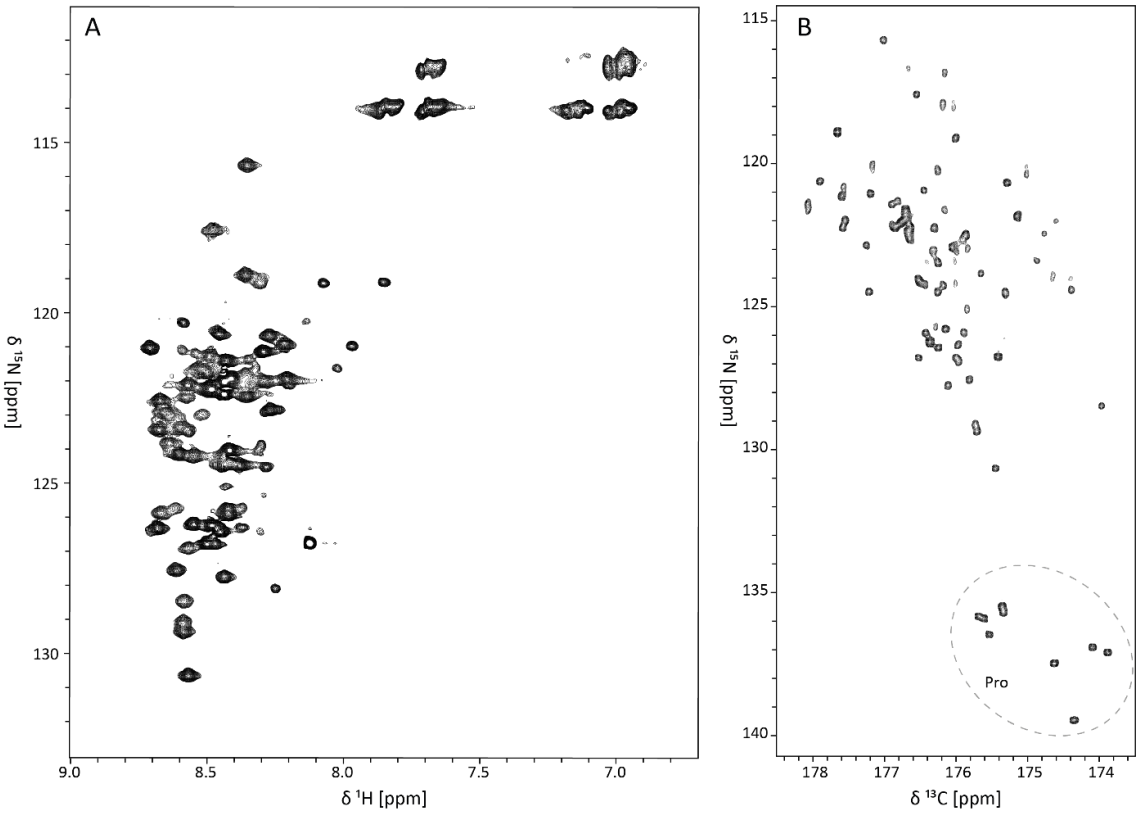

Figure 49. 2D $^{1}H,^{15}N$ and $^{13}C,^{15}N$ correlation spectra of $^{13}C,^{15}N$ labelled RNase E (694-790). Measurements were performed with a 350 µM protein sample in 450 mM KCl, 25 mM Tris, 100 µM DSS and 10% D_2O at pH 8.3 were recorded at 278 K at 600 MHz. Resonances were referenced to DSS. A. BEST-TROSY experiment was recorded with 92 scans. B. (HACA)CON experiment was recorded with 32 scans. Proline C-N correlations are highlighted.

From the 88 theoretically detectable backbone amide resonances - the number of prolines and the N-terminus subtracted from the total number of amino acids - 82 individual signals could be observed. Since a significant number of resonance signals did exhibit a reduced signal intensity indicative of conformational dynamics, it is possible that the remaining residues are either in conformational exchange and therefore broadened beyond detectability or concealed by spectral overlap.

Intrinsically disordered proteins are also typically investigated at mildly acidic pH conditions to yield NMR spectra with higher sensitivity. Similar measurements with RNase E (694-790) at pH 6.5 did indeed lead to $^{1}H,^{15}N$ spectra with an overall more consistent peak intensity and a significantly higher sensitivity especially considering the reduced protein concentration of only 50 µM compared to 350 µM for comparable spectra at pH 8 (Figure 45). Unfortunately, interaction studies with RhlB at a pH of 6 were not viable, since RhlB quickly precipitated at this pH. NMR conditions were therefore limited to pH 8.3. To nonetheless maximise the number of observable resonances and to improve the poor ^{1}H amide chemical dispersion, the BEST-TROSY was complemented by a (HACA)CON experiment. This carbon-detected C-N correlation experiment was developed by Bermel *et al.* and has been used regularly to study protein systems that lack a defined structure[154,155,176]. The correlation of the carbonyl carbon with the amide of the following amino acid yields a fingerprint spectrum similar to that of an $^{1}H,^{15}N$ HSQC and excitation via Hα additionally excludes side chain resonances (compared to conventional

CON experiment). Unlike $^{1}H,^{15}N$ correlation experiments, the (HACA)CON also allows for the detection of proline resonances. Figure 49 panel B displays the (HACA)CON spectrum of RNase E (694-790) under pH 8 conditions. The striking increase in resolution is clearly visible and the additional low field shifted proline residues can easily be identified. Together with an improved chemical shift dispersion, a total of 91 out of 98 backbone resonances could be detected even under pH 8 conditions, including all nine proline residues.

5.2.4 NMR Spectroscopic Analysis of RhlB/RNase E (694-790) Complex

After assessing the NMR profile of the individual proteins, the effect of their interactions was investigated for both sides. Therefore 1 or 1.2 equivalents of unlabelled protein were added to the corresponding ^{15}N or $^{13}C^{15}N$ labelled interaction partner.

At first, we had a look at RhlB's point of view. 100 µM ^{15}N labelled RhlB were mixed with an equivalent amount of unlabelled RNase E (694-790) and the resulting $^{1}H,^{15}N$ correlation spectra were compared to spectra of isolated ^{15}N RhlB at the same measurement conditions. This meant measuring at an elevated ionic strength of 450 mM KCl to ensure a comparability to the other spectra but experiments at 150 mM salt concentration were recorded as well. Here, the most apparent effect of the complex formation on the helicase's stability was observed. As illustrated in Figure 50, the presence of RNase E (694-790) did result in an overall well dispersed $^{1}H,^{15}N$ HSQC spectrum of folded RhlB at 450 mM KCl but even at a lower salt concentration of 150 mM KCl. The binding of the RNase fragment seemed to prevent the formation of the potentially dimerized or misfolded RhlB species, possibly by blocking a dimerization surface on the helicase. Alternatively, the high charge density of the RNase fragment could be promoting solubility of RhlB analogue to an increased ion concentration in the buffer. A direct comparison of the $^{1}H,^{15}N$ HSQC spectra of RhlB/RNase E (694-790) at low and high salt concentrations showed only minute differences in chemical shift (Appendix Figure 71) between the two conditions.

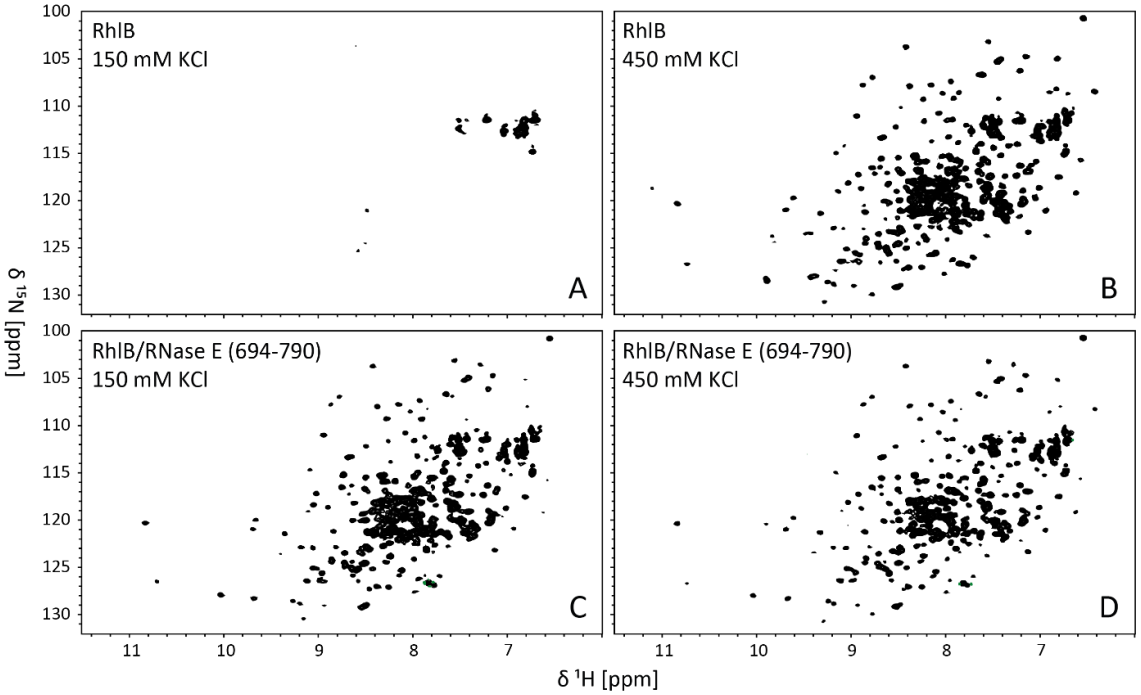

Figure 50. 2D $^{1}H,^{15}N$ HSQC spectra of ^{15}N labelled RhlB at different salt concentrations and in presence and absence of 1 equivalent unlabelled RNase E (694-790). Measurements were performed at 305 K with a 100 µM protein sample in 150 or 450 mM KCl, 25 mM Tris, 5 mM DTT, 100 µM DSS and 10% D_2O (pH 8.3). A. Spectrum was recorded with 104 scans at 600 MHz. B.-D. Spectra were recorded with 104 scans at 800 MHz.

For a more detailed spectra analysis we compared the $^{1}H,^{15}N$ HSQC of RhlB with a sample containing RhlB and one equivalent of RNase E (694-790) at 450 mM, as depicted in Figure 51. In presence of the RNase E fragment the overall peak intensity was reduced by approximately 10-15%. The spectra also showed distinct chemical shift perturbations (CSPs) in 48 amide resonances upon binding of the RNase E fragment. Since RhlB's putative RNase E binding site encompasses only 30 amino acids (AA 368-399), we can assume that additional section of the protein experienced structural changes. This finding corresponds well with experiment by Bruce *et al.*, who also observed structural changes in additional sections of RhlB's CTD beyond the interaction site with RNase E[127].

While 261 amide resonances could be identified for RhlB alone, only 230 signals were detected for the complex. Although it must be acknowledged that the helicase might not be fully saturated with 1 equivalent of interaction partner, the disappearance of 31 residue signals upon binding suggests those peaks were in slow to intermediate conformational exchange.

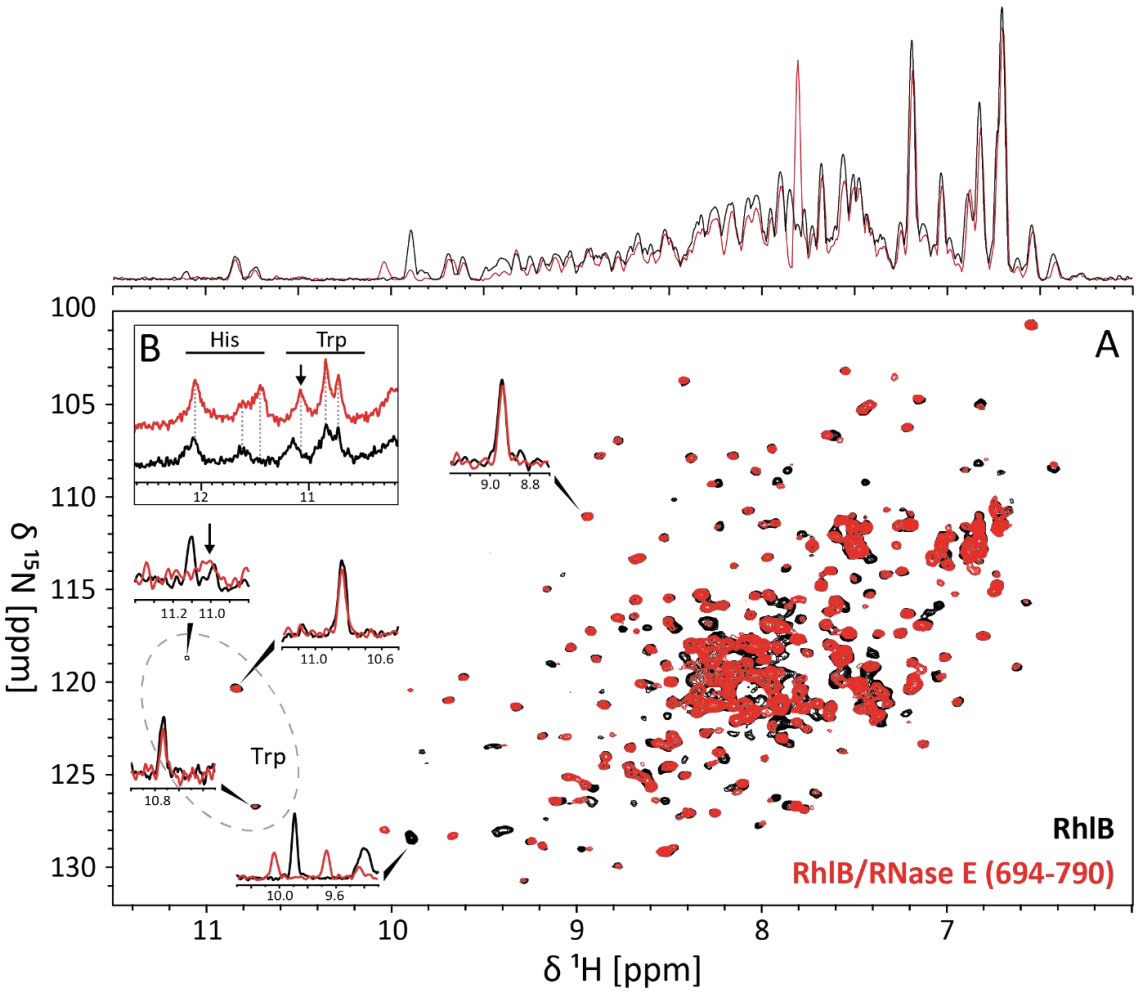

Figure 51. Overlay of RhlB with RhlB/RNase E (694-790) complex. A. Overlay of ^{1}H, ^{15}N HSQC spectra of 100 µM ^{15}N labelled RhlB in presence and absence of 1 equivalent unlabelled RNase E (694-790). 1D projections show exemplary peaks with no or strong CSP as well as all three tryptophan side chain resonances. Both spectra were measured with 104 scans in 450 mM KCl, 25 mM Tris, 5 mM DTT, 100 µM DSS and 10% D_2O (pH 8.3) at 305 K and 800 MHz. B. Overlay of ^{1}H 1D spectra with jump-return-echo water suppression of 100 µM unlabelled RhlB in presence and absence of 1 equivalent unlabelled RNase E (694-790). Closeup of 13-10 ppm region shows tryptophan side chain resonances as well as supposable histidine side chain resonances. Both spectra were measured with 16 scans in 150 mM KCl, 25 mM Tris, 5 mM DTT, 100 µM DSS and 10% D_2O (pH 8.3) at 305 K and 600 MHz.

This also applied to one of the three tryptophan side chain resonances, as it completely vanished in the 2D spectrum, while the other two were not affect at all by the interaction. The 1D projection of the corresponding tryptophan hints on the appearance of a new tryptophan peak in the bound conformation, and with the sensitivity enhancement of an ^{1}H 1D experiment with jump-return water suppression the chemical shift perturbation for the one tryptophan was confirmed (Figure 51 panel B). Those 1D spectra also revealed two more significantly low-field shifted amide resonances, which were proposed to be protonated histidine side chain amide protons and which also experienced CSPs during binding. As explained in more detail in section 1.3.4, Bruce *et al.* identified four regions in the CTD of RhlB along with the direct binding site, that are structurally impacted by binding of RNase E.

Of these four only the region corresponding to residues 251-263 contains a tryptophan residue. It is therefore most likely that the shifted tryptophan residue is W255.

Next, the protein-protein interaction was analysed from the perspective of RNase E (694-790). Therefore, the $^{13}C,^{15}N$ labelled protein was titrated with up to 2 equivalents of unlabelled helicase. Optimizations of measurement conditions were still ongoing at the point of recording, so that the titration was performed at 289 K and recorded with regular $^{1}H,^{15}N$ HSQC instead of with a BEST-TROSY at 278 K. The spectra of each titration step are overlayed in Figure 52. Despite the suboptimal resolution, the titration did unambiguously reveal that RNase E (694-790) did not undergo a disorder-to-order conformation upon binding to the helicase since the chemical shift dispersion of the RNase remained very narrow upon addition of the helicase. Using 1D projections for two example peaks it could be excellently shown that the chemical shift perturbations upon complex formation followed a slow exchange regime. With the improved resolution of a BEST-TROSY at 278 K (Figure 53 panel A) a total number of 66 peaks were identified for RNase E (694-790) in complex with the helicase (1:1 stoichiometry), which was 6 resonances fewer than for the isolated RNase E fragment. This reduced number of resonances indicated again that the complex was not fully saturated, and a certain number of resonances were in conformational exchange. We therefore increased the RhlB concentrations to 1.2 equivalents for the recording of the (HACA)CON experiment. The excellent chemical shift dispersion of the C-N correlation experiment revealed that significantly more resonances vanished upon complex formation than previously assumed from the BEST-TROSY. A total of 56 resonances could be detected for the (HACA)CON experiment with RhlB while 91 signals were observed for RNase E (694-790) alone. This difference of 34 signals appears to reflect the binding site of the RNase E fragment more accurately than the BEST-TROSY, since the two proposed amino acid sequences directly interacting with RhlB were defined to span 48 residues[18]. Since those amino acid stretches were identified by limited proteolysis it is possible that the actual number of interacting amino acids could be even smaller. In fact, our data indicate that only two of the four proline residues encompassed in Chandran's proposed interaction site are actually binding, since the (HACA)CON spectrum clearly shows the disappearance of only two prolines.

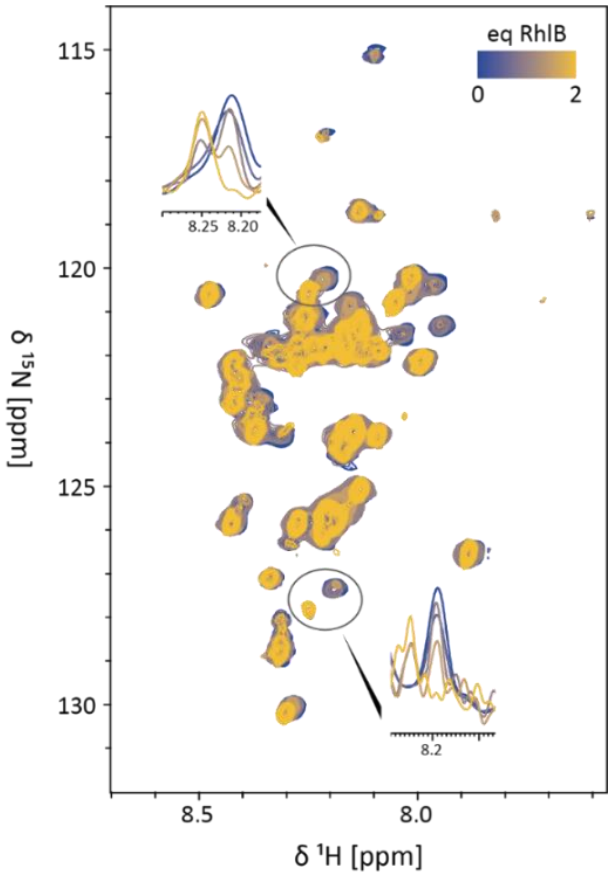

Figure 52. Overlayed $^{1}H,^{15}N$ HSQC spectra of ^{15}N labelled RNase E (694-790) titrated with up to 2 equivalents of unlabelled RhlB. The measurements were performed with 100 µM RNase E (694-790) in 150 mM KCl, 25 mM Tris, 5 mM DTT, 100 µM DSS and 10% D_2O (pH 8.3). Spectra for the following titration steps were recorded (equivalents of RhlB): 0, 0.25, 0.5, 1, 2. All experiments were recorded with 24 scans at 289 K and 900 MHz.

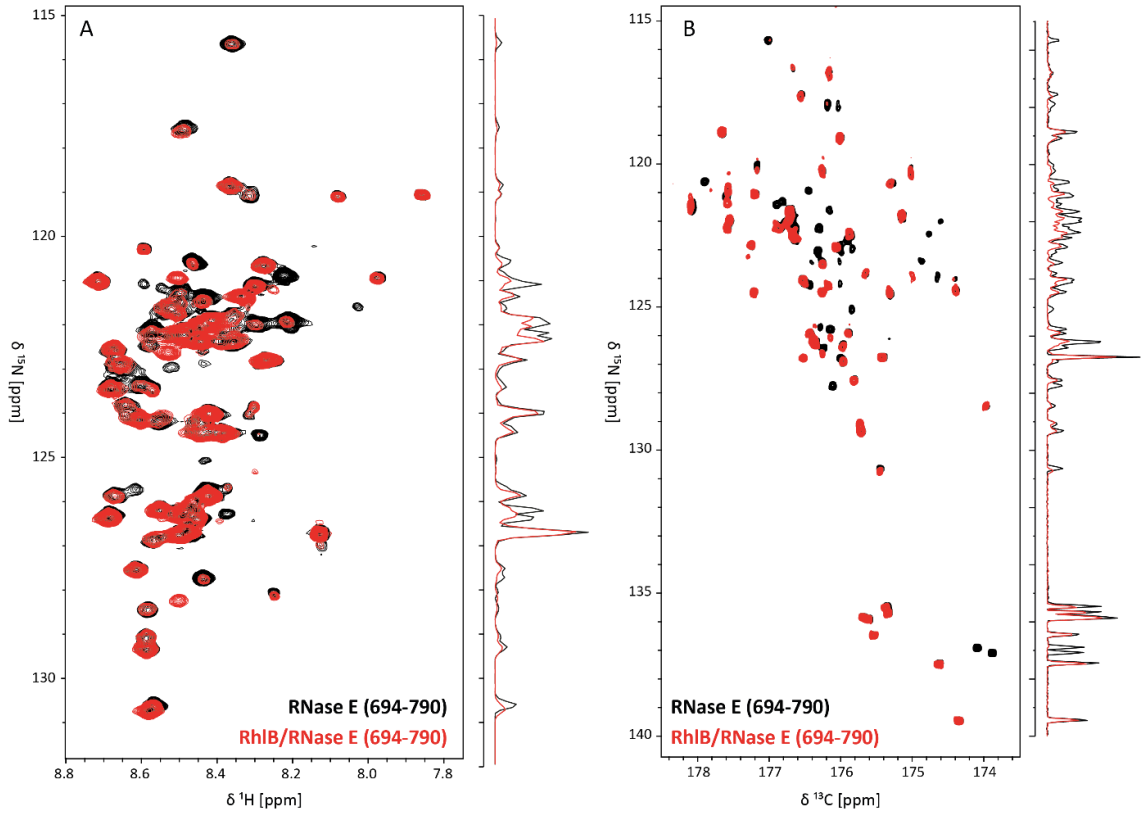

Figure 53. Overlay of BEST-TROSY and (HACA)CON spectra of 350 µM $^{13}C^{15}N$ labelled RNase E (694-790) and RhlB/RNase E (694-790) complex. Measurements are recorded with 350 µM RNase E (694-790) and 1 (A.) or 1.2 (B.) equivalents RhlB in 150 mM KCl, 25 mM Tris, 5 mM DTT, 100 µM DSS and 10% D_2O (pH 8.3) at 278 K and 600 MHz. A. BEST-TROSY spectra were measured with 104 scans. B. (HACA)CON spectra were measured with 32 scans.

In summary, we could successfully demonstrate the complex formation of RhlB and RNase E (694-790) via NMR spectroscopy. We showed with a straightforward analysis that RhlB adopts a well-defined structure that is interspersed with flexible disordered loop sections in both domains. We encountered unexpected differences between unlabelled and labelled helicase in their tendency towards non-native dimer formation, which we could resolve by adaptation of the buffer's ionic strength. The interaction with RNase E (694-790) did not significantly alter the secondary structure composition or dynamic of RhlB but induced several CSPs including a section that presumably encompasses Trp255. The BEST-TROSY spectra of RNase E (694-790) showed that the protein maintained it disordered structure upon binding of the helicase and did not follow a disorder-to-order transition reminiscent of an induced fit. Full titration experiments also showed that both proteins interact in a slow-exchange regime. The carbon-detected C-N correlation experiment proved to be particularly beneficial in the assessment of the complex involving a natively disordered protein as it provided an improved chemical shift dispersion over a BEST-TROSY. More in-depth analysis of the protein-protein interaction would require full resonance assignments of both proteins, which was beyond the scope of this thesis. After the structural analysis of the complex, we wanted to investigate the effect of both RNase E fragments on RhlB's reaction cycle.

5.3 ATP Hydrolysis of RhlB

Apart from assessing the structural influence of the protein-protein interactions of RhlB and RNase E (694-790) we primarily wanted to investigate the regulatory impact both RNase E fragments have on RhlB's reaction mechanism. The ATP turnover rate has become one of the standard measures to determine the catalytic activity of a DEAD-Box helicase, as the hydrolysis of one ATP can be directly correlated to a full reaction cycle and a variety of straight-forward spectrophotometrically assays can nowadays be utilized to determine the amount of consumed ATP even with very low sample concentrations[16,17,106,120,177]. Previous studies by Luisi and Carpousis on the RhlB's ATPase activity showed that wildtype RhlB in isolation has a barely detectable ATPase activity and that the helicases ATP turnover rate in complex with either RNase E (628-843) or RNase E (694-790) was increased up to 15-fold[6,17,20]. The focus of our investigations was set on the interplay between these activating RNase E fragments and RNA substrates with specific features such as orientation of single strand extension and duplex length. As opposed to preceding studies on RhlB that only relied on bulk RNA from *S. cerevisieae* for ATPase assays we individually tested five different RNA substrates with defined structures: Two blunt end duplexes with lengths of 13 and 21 Nt, 13 and 15 Nt long duplexes with 3' or 5' single strand extension, respectively, and a 21 Nt single strand. The Molecular Probes™ EnzChek™ Phosphate Assay Kit, as previously described by Worrall *et al.*, was used to assess the catalytic performance of RhlB[17]. For each measurement, a sample containing RNA substrate, RhlB ± 1 equivalent RNase E fragment and the kit components were preincubated in reaction buffer at room temperature and the reaction started by addition of an access of ATP. The linear increase in absorption at 360 nm, which correlated with the release of free phosphate after ATP hydrolysis was measured for 300 sec (Figure 54A). The assays were in part performed by Laura Raschke under my supervision.

Figure 54B shows the ATP turnover rates determined in triplicates for RhlB, RhlB/RNase E (694-790) and RhlB/RNase E (628-843) for all investigated RNA constructs as well as a negative control, where RNA was omitted. For all measurements with RhlB, only minimal turnover rates were detected, which did not exceed the control measurement without RNA. The variations between RNA constructs were not significant, given that the measured values were at the assay's lower sensitivity limit. These results are also in agreement with findings by Worrall *et al.*, who reported similar turnover rates (< 0.5 mol Pi/ min/ mol helicase) for RhlB in presence and absence of bulk yeast RNA substrate[17].

Upon addition of the minimal binding fragment RNase E (694-790), an increase in RhlB's ATP turnover rate could be detected for all RNA constructs, but notable differences in the degree of activation were observed for the individual RNA substrates. While measurements with the 3' single stranded RNA showed only a minute rate increase of 0.2 mol Pi/ min/ mol helicase, both the 5' extended and the single strand construct exhibited a three times higher rate increase (0.6 mol Pi/ min/ mol helicase). Both blunt

end RNA duplexes performed similar regarding turnover rate and measurements displayed an activation between that of the 3' and 5' extended constructs. Notably, the two constructs with which RhlB exhibited the strongest increase in ATPase turnover rate both feature a single stranded 5'end. This differential stimulation of RhlB's ATP turnover rate induced by the allosteric binding of RNase E (694-790) strongly supports the assumption of a preference for specific substrate features like 5' over 3' overhang. These findings are also corroborated by Chandran *et al.*'s research on RhlB in complex with a larger RNase E fragment, where a strong increase in RNA unwinding activity was observed for 5' tailed RNA substrates whereas 3' tailed substrates showed only minimal unwinding[18]. Whether RhlB has an inherent preference for specific RNA features that are merely amplified by RNase E (694-790)'s binding or whether the RNase E fragment induces these preferences, will be further assessed in the following chapters.

For measurements with RhlB/RNase E (628-843), the larger RNase E fragment extended by two flanking RNA binding sites, even higher ATP turnover rates were observed for RhlB and the differences in the degree of activation for individual RNA constructs were even more pronounced. The highest rate was detected for measurements using the RNA single strand with 3.24 mol Pi/ min/ mol helicase, while the 13 Nt blunt end duplex construct even exhibited a decrease in turnover rate. Both constructs with single strand overhangs showed considerably higher ATP hydrolysis rates than the blunt end constructs of comparable length.

While the rate activation observed by RNase E (694-790) could be attributed to allosteric effects on RhlB, a more complex interplay must be considered for the larger RNase E fragment. While it is beyond the explanatory power of this experimental setup, the following possible interpretations are given for the increased ATPase turnover rate for RhlB in complex with RNase E (628-843): The RNA binding regions of RNase E (AR2, RBD) might either recruit more substrate to the complex by binding additional RNA molecules, thereby increasing the availability of RNA substrate for the reaction, or function as scavenger of RNA single strands released during the reaction cycle. Considering that in context of the *in vivo* degradosome complex the RNA binding domains of RNase E are attributed with keeping unwound RNA secondary structure elements separated for further degradation, the latter appears to be more likely. Either way, the results revealed that the propensity of the RNA binding sites AR2 and RBD to bind or release the RNA substrates is not equal for all substrates, as the differences in rate activation indicate. The mechanism of RNA binding for by RBD and AR2 has not been fully examined, yet. Even though the high density of positively charged amino acids in both binding sites is indicative of an electrostatically driven interaction with the RNAs, the degree in rate stimulation did not follow the substrates charge density, thereby implying a more complex mode of interaction.

Taken together, the ATPase assay experiments revealed that in presence of binding partner RNase E, RNA substrate features have a significant influence on the overall ATP turnover rate of the helicase, revealing the strongest stimulation for 5' single strand tailed

constructs, as RhlB exhibited the highest degree of activation with both the 5' tailed duplex and the single strand. This finding, that was previously missed due utilization of RNA substrates with unidentified composition, gives a useful rational for similar preferences observed previously in RhlB unwinding experiments[18].

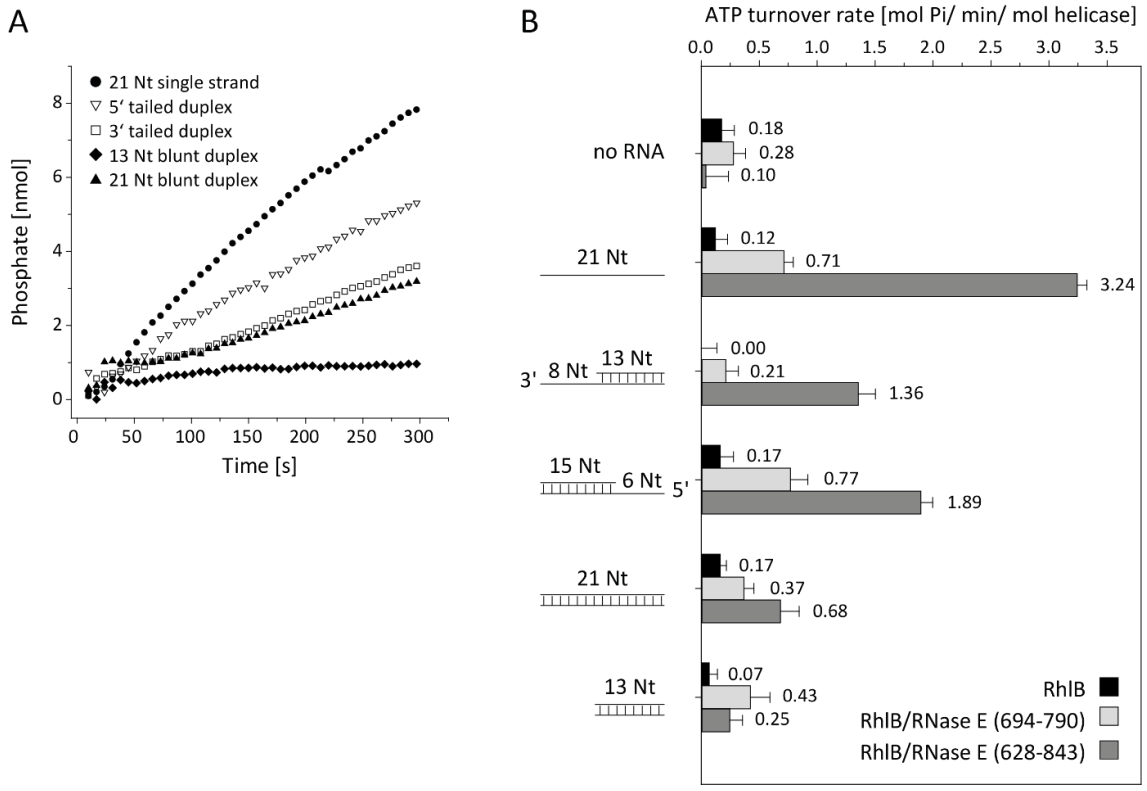

Figure 54. ATPase activity of RhlB with varying RNA substrates and RNase fragments as determined with Phosphate Assay. (A) Representative activity profile of RhlB in complex with RNase E (628-843) and single stranded (●), 5' overhang (▽), 3' overhang (□) 13 Nt (◊), or 21 Nt blunt end (▲) RNA substrates. (B) Bar diagram of ATP turnover rates of RhlB in presence and absence of RNase E (694-790) or (628-843) as well as different RNA ubstrates.

5.4 Analysis of RhlB's ATP Hydrolysis Rate under NMR Conditions

After the initial assessment of RhlB's catalytic performance under the influence of both RNase E and different RNA substrates, we wanted to gain a deeper understanding of the underlying structural processes by utilizing NMR spectroscopy. To verify, whether the molecular effects previously observed in the ATP turnover experiments could be reproduced under conditions required for NMR measurement, we performed real-time NMR mixing experiments with the detection of the ATP hydrolysis through consecutive ^{31}P 1D NMR spectra. To be able to detect RNA imino proton resonances in the subsequent RNA centred experiments, RNA substrate concentrations were increased to 100 µM and a four-fold excess of helicase over RNA was used to ensure complete substrate binding. For the experimental setup, depicted in Figure 55 panel A, 300 µL of a pre-equilibrated solution of 5' tailed RNA substrate J2Δ16-M2 and either RhlB or a RhlB/RNase E complex in buffer were prepared in a Shigemi NMR tube, which was inserted with capillary containing 40 µL ATP solution. The ATP turnover reaction was initiated upon injection of the ATP solution and the progress of ATP hydrolysis monitored via consecutive ^{31}P 1D spectra. Panel B shows the chemical shift of ATP's three phosphate resonance signals in a ^{31}P 1D spectrum directly after mixing and the phosphate resonances of the hydrolysis products ADP, free phosphate and even AMP after 14 h of reaction, verifying that ATP is in fact hydrolysed in the NMR tube. Control experiments of ATP in reaction buffer confirmed that no spontaneous ATP hydrolysis took place over a period of 24 h under the utilized buffer conditions and that the conversion of ATP to ADP observed in the mixing reaction could therefore unambiguously be attributed to catalytic reaction by RhlB (See Appendix Figure 73). The resonance integral of ATPα over the course of 12.8 h was used to calculate the total amount of ATP hydrolysed, from which the ATP turnover rate constants were obtained via double exponential fit (Figure 55 panel C). The measurement with RNase E (694-790) was discontinued at 124 min due to technical difficulties and fitted for that duration. Comparing the hydrolysis reaction of RhlB with either RNase E fragment demonstrated a 4-fold accelerated ATP turnover rate in presence of RNase E (694-790) and 1.5-fold accelerated rate for the RNase E (628-843) construct, therefore demonstrating that the activating effect of RNase E could in fact be reproduced under NMR conditions. While the previous enzymatic phosphate assay showed an even higher rate increase for the RhlB/RNase E (628-843) complex in comparison to RhlB/RNase E (694-790), this could not be reproduced under NMR conditions. This deviation could presumably be ascribed to the altered protein:RNA ratio from 1:2.5 in the phosphate assay to 4:1 in the NMR experiments. Changing the ratio was required to ensure complete binding of RNA to the helicase since the RNA concentrations for this experiment are approaching RhlB's K_D. The excess of proteins over RNA created a situation, where the RNA binding regions of RNase E (628-843) were likely to compete with RhlB for the RNA substrate, leading to a decreased amount of RNA being bound by the helicase. Based on this finding it was decided to omit the larger RNase E fragment from subsequent

structural NMR investigations in section 5.6, as conformational changes in the RNA substrate could not unambiguously be attributed to the helicase.

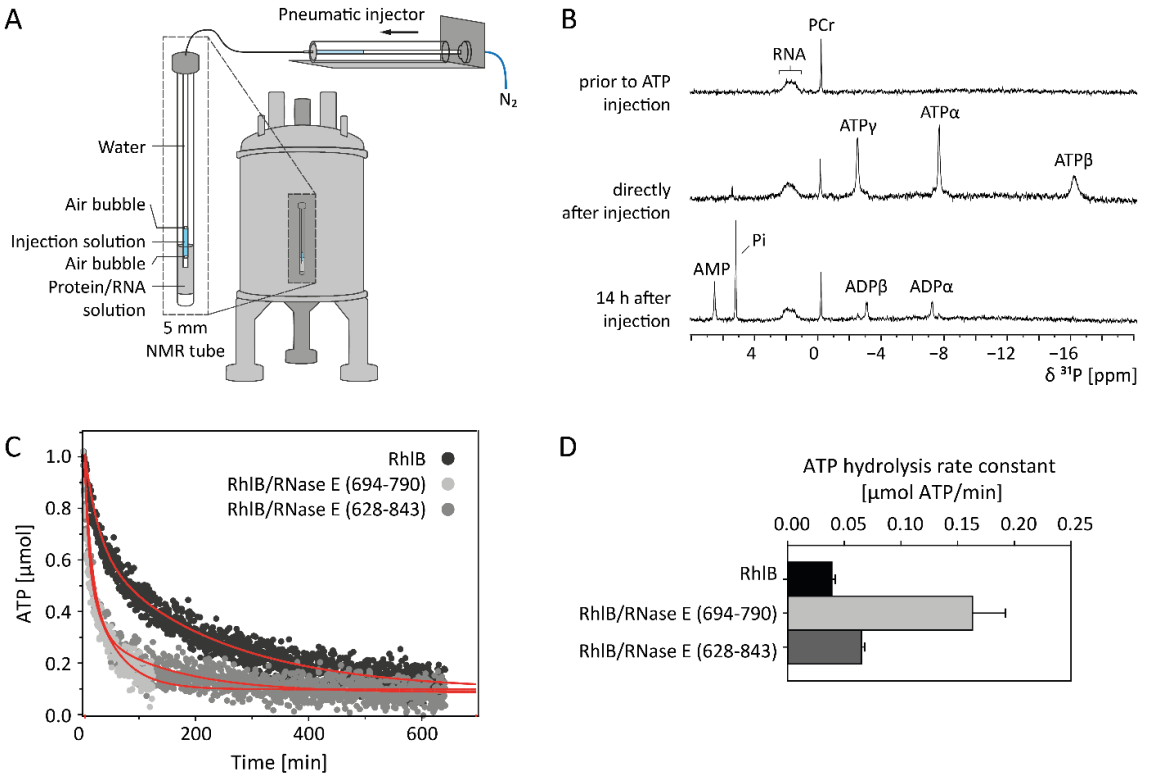

Figure 55. ATP hydrolysis kinetics measured by ^{31}P NMR real-time mixing experiments. (A) Schematic of a real-time NMR mixing setup including NMR tube with injection insert. Injection of ATP solution is triggered by pneumatic piston in direct response to electronic signal of pulse sequence command. The 5 mm NMR Shigemi tube containing 300 µL of a pre-equilibrated protein/RNA mix (100 µM 5'-tailed RNA duplex J2Δ16-M2, 400 µM RhlB or RhlB/ RNase E complex (1:1), 100 µM DSS, 6% D_2O,150 mM KCl, 25 mM Tris/HCl, 5 mM DTT, 4.5 mM $MgCl_2$ at pH 8.3) was mixed with 40 µL of injection solution (25.5 mM ATP, 100 µM DSS, 6% D_2O, 150 mM KCl, 25 mM Tris/HCl, 5 mM DTT, 4.5 mM $MgCl_2$ at pH 8.3) to result in a final ATP concentration of 3 mM. (B) 1D ^{31}P NMR spectra of the protein/RNA reaction mix prior to ATP injection, directly after injection (4.2 min) and 14 h after injection. The appearance of ATP resonance signals confirms the successful injection process. Spectra were recorded with 128 scans and referenced to phosphocreatine (PCr). Peak assignments: P_i, orthophosphate; AMP, α-phosphate group of adenosine monophosphate; ADPα/ ADPβ, α- or β-phosphate groups of adenosine diphosphate; ATPα/ ATPβ/ ATPγ, α-, β- or γ-phosphate groups of adenosine triphosphate. (C) Kinetic NMR data of RhlB induced ATP hydrolysis in ^{31}P real-time NMR experiment. Shown are total amounts of ATP over time for reactions with RhlB alone and in complex with RNase E (694-790) or RNase E (628-843). The amount of ATP was determined from ^{31}P peak integral of ATPα and curves were fitted with double exponential curve fit. Experiments were recorded for 12.8 h as pseudo-2D with 32k scans and ATP injection after 128 scans. The measurement with RNase E (694-790) was discontinued at 124 min and fitted for that duration. (D) Bar diagram of ATP turnover rate constants k_1 extracted from double exponential fits.

5.5 NMR Spectroscopic Investigation of RhlB's RNA Binding Affinity

Since our ATPase assays clearly revealed a RNA substrate-depending ATP turnover rate of RhlB in presence of RNase E, we wanted to further investigate where this RNA differentiation originates and identify whether the RNA substrate preferences observed in chapter 5.3 are inherent to RhlB and only enhanced by the RNase E fragments or whether RNase E (628-843) and (694-790) induce those preferences in the first place. As previously explained, first evidence of RNA substrate preference for the RhlB/RNase E (628-843) complex has been demonstrated in RNA unwinding experiments by Chandran *et al.*, where the complex showed an increased unwinding activity towards a 5' tailed over a 3'-tailed duplexes. However, the low inherent unwinding activity of RhlB by itself prevented them from obtaining substrate specific unwinding activities in absence of RNase E[18]. We encountered the same inability to obtain RNA specific results for RhlB alone in our ATPase assay experiments. We therefore wanted to instead focus on the initial binding step of RNA to the helicase in absence of ATP. While it has to be pointed out, that for some DEAD-Box helicases the binding of ATP and RNA have been shown to be cooperative and so the absolute affinities might change with the addition of ATP [90,106,114,117,178], binding of RNA alone should give a valuable insight into possible substrate preferences of the helicase without undergoing the full multi-step reaction cycle and provide an accurate read-out even in absence of RNase E fragments.

We investigated the binding of our RNA constructs to RhlB by performing ^{1}H 1D NMR titration experiments. By monitoring the intensity and the overall changes of RNA imino proton resonances upon stepwise addition of up to 4 equivalents of protein we would not only be able to determine the apparent K_D for each substrate but could also detect possible conformational rearrangements accompanied by changes in base pairing within the RNA substrate during binding. For each measurement 100 µM RNA were prepared in pH 8 NMR buffer. RhlB, RNase E, or a RhlB/RNase E complex was added stepwise with ^{1}H 1D measurements of the RNA resonance signals at every titration step, exemplified in Figure 56 B. The apparent K_Ds were then calculated by plotting the normalized reciprocal intensity of the non-overlapping RNA's imino proton resonances against protein concentration and fitting globally with a ligand binding function.

The 5' single strand extended construct was used for the first set of titrations to evaluate the effect of both RNase E fragments on the affinity of RhlB (See Figure 56A and Table 11): For all titration experiments a homogenous intensity decrease accompanied by peak broadening was observed with increased protein concentration for all imino proton resonances, which is indicative of binding to the large protein as this affects the tumbling speed of the RNA and therefore the linewidth of the peaks. No chemical shift perturbations were observed for any of the imino protons resonances. The 5'-OV construct J2Δ16-M2 was bound by RhlB with a low micromolar apparent K_D of 46.6 µM. This affinity was increased by about a factor of three to 14.5 µM when RNase E (694-790)

was bound to the helicase. Since RNase E (694-790) itself did not exhibit RNA-binding capabilities, as demonstrated in the control titration omitting RhlB (Figure 56A), this change in affinity can be directly attributed to the allosteric binding of the RNase E fragment to the helicase. In complex with the larger RNase E (628-843) fragment the affinity towards the RNA substrate was increased even further to 5.1 µM. As expected from the additional RNA binding domains AR2 and RBD, titrations with only RNase E (628-843) showed that the larger RNase E fragment itself does bind the 5' tailed RNA with an overall apparent K_D of 162.7 µM. This implies that the apparent K_D observed for the RhlB/RNase E (628-843) complex is presumably a product of both the helicase and RNases RNA binding sites, which unfortunately impedes a clear readout of the helicase's individual RNA affinity in this complex. Nonetheless this illustrates how the two individual effects of allosteric activation and additional RNA binding regions both contribute to RhlB's interaction with RNA.

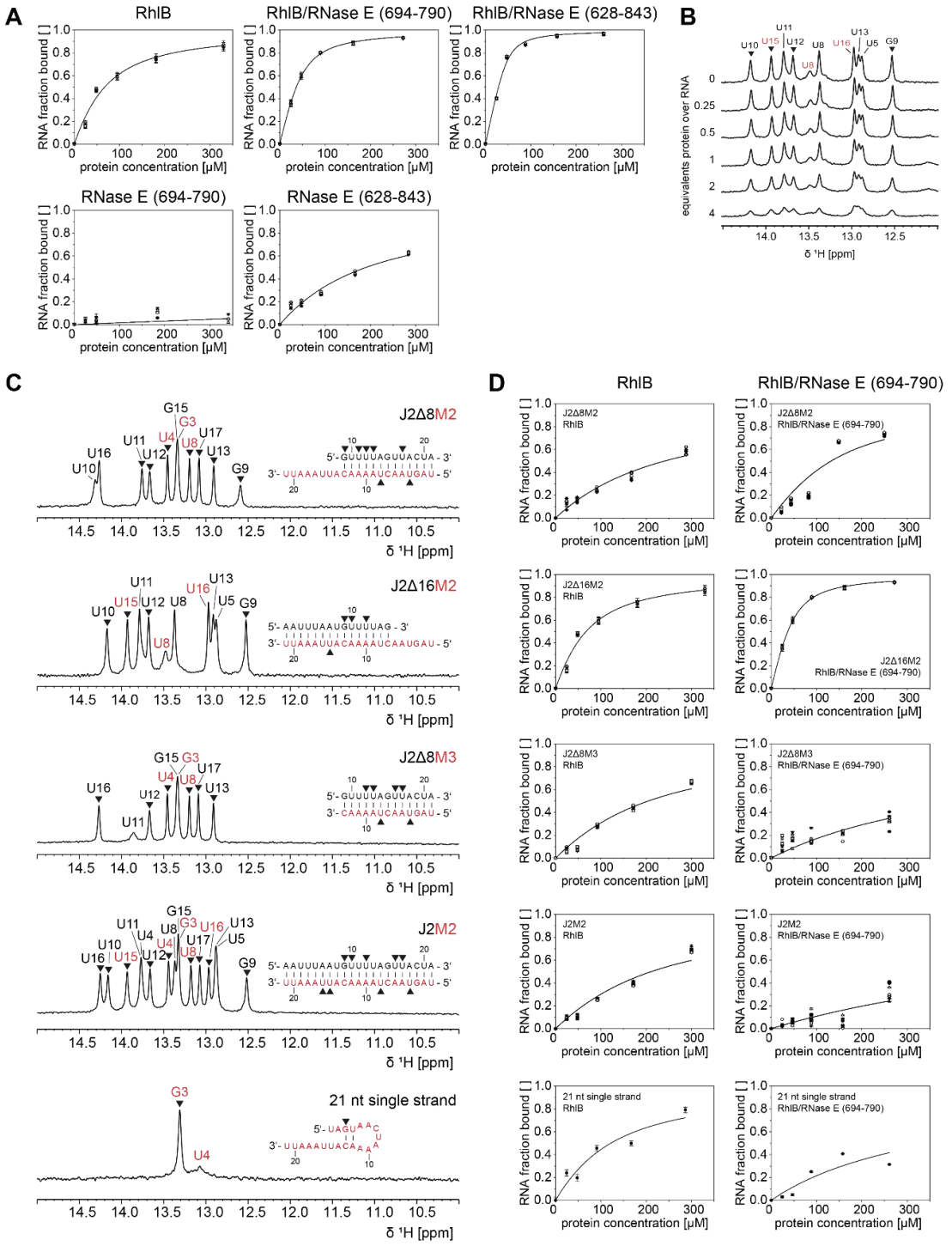

Figure 56. ¹H NMR analysis of RhlB's binding affinity towards different RNA substrates in presence and absence of RNase E (694-790) or (628-843). (A). RNA binding curves of 5'-OV RNA substrate for titrations with RhlB alone, with RhlB in complex with RNase E (694-790) or RNase E (628-843), and with RNase E (694-790) or RNase E (628-843) alone. (B) Exemplary imino proton region of the 1D ¹H NMR spectra of the 5'-OV RNA substrates with stepwise titration of up to 4 equivalents RhlB. (C) Assigned imino proton region ¹H NMR spectra for all RNA substrates including oligonucleotide sequences. It is noted that the 21 Nt single strand forms two weak base pairs under NMR conditions. Resonances assigned to bottom or top strand are colour coded in red and black, respectively, and only peak intensities of non-overlapping resonances (annotated with ▼) were used for K_D calculations. (D) RNA binding curves of all RNA substrates for titrations with RhlB and RhlB/RNase E (694-790). Binding affinities were determined by plotting normalized reciprocal peak intensity against protein concentration and fitting globally with one site ligand binding function [179].

Table 11. Measured apparent K_D of RhlB and RhlB/RNase E complexes towards different RNA substrates calculated from global fit of NMR titration experiments. Affinities are given in [µM] and the error given are the mean SD of the global curve fit.

RNA substrate	RhlB	RhlB/ RNase E (694-790)	RhlB/ RNase E (628-843)	RNase E (694-790)	RNase E (628-843)
J2Δ16-M2 (5'-tailed duplex)	46.6 ± 6.8	14.5 ± 0.7	5.1 ± 0.8	n.d. [a]	162.7 ± 30.9
J2Δ8-M2 (3'-tailed duplex)	216.3 ± 33.7	82.3 ± 27.3	-	-	-
J2Δ8-M3 (13 Nt blunt duplex)	154.9 ± 30.8	487.5 ± 40.5	-	-	-
J2-M2 (21 Nt blunt duplex)	161.7 ± 30.3	n.d. [a]	-	-	-
M2 (21 Nt single strand)	94.9 ± 59.2	364.5 ± 1.0	-	-	-

[a] not determinable

Next, we wanted to examine whether RNA substrates with other features are bound with similar affinity as the 5' tailed duplex and whether RNase E (694-790) also increases RhlB's affinity towards those constructs. The same experimental setup was utilized. The K_Ds of all RhlB titrations summarized in Table 11 and plotted in Figure 56D clearly show significant differences in affinity towards the individual RNA constructs: While the RNA with 5' single strand extension was bound with an apparent K_D of 46.4 µM, the affinity for the 3' extended duplex construct of the same length was 4.6 times lower with 216.3 µM. Interestingly, this affinity is even weaker than that of the helicase towards the single strand, which was measured with 94.9 µM. RhlB's affinity towards both blunt end constructs appear to be between that of the single strand and the 3'-tailed RNA with the difference in duplex length not having a significant effect on the overall affinity (13 Nt duplex 154.9 µM, 21 Nt duplex 161.7 µM). Those results are clearly revealing an inherent binding preference of RhlB towards RNA with a single stranded 5' end, even in absence of any RNase E fragment.

Nonetheless, the following considerations have to be made for the 21 Nt single strand here: while M2 was used as the single stranded RNA substrate for both the ATPase assay and the NMR experiments, initial ^{1}H 1D NMR measurements in chapter 4.3.8 revealed two adjacent base pairs forming a small hairpin with both dangling 3' and 5' ends. Although those two base pairs presented the opportunity to also measure the K_D of M2

with this NMR setup, it gave rise to the question whether M2 is actually recognized as a single strand by RhlB.

Most striking was the subsequent observation that the affinity of RhlB did not increase for all RNA constructs with the addition of RNase E (694-790): While RhlB's apparent K_D towards the 3' tailed RNA was reduced by a factor of 2.5 to 82.3 µM when RNase E (694-790) was present, a 3-4 fold decrease in affinity was measured for both the single strand and the blunt-ended constructs, with the 21 Nt blunt end substrate J2-M2 binding so weak that a K_D could not be determined reliably. Unfortunately, some of the RNA constructs could not be titrated to full saturation even at maximum protein concentrations of 250-300 µM. Their relative affinity towards the protein complex nonetheless demonstrates that the allosteric binding of RNase E (694-790) has a direct effect on RhlB's RNA substrate affinities, favouring duplex substrates with single stranded extensions.

While there is no direct comparison of the measured affinities of RhlB for these specific RNA substrates in the literature, we could corroborate the range of the determined affinities in part with electrophoretic mobility shift assays (EMSAs) performed on the constructs JM2h, J2Δ8-M2 and J2-M2 (See Figure 46 and Appendix Figure 70). The differences in experimental setup and concentrations did cause a variance in the exact K_D values, the assays did nonetheless corroborate that RhlB's and RNase E (628-843)'s affinities for the RNA substrates lie in the 100 µM range. Chandran *et al.* however did assess the affinity of RhlB and RhlB/RNase E (628-843) to a fragment of the 9S precursor rRNA[18]. They, too, did observe an increase in affinity in presence of the RNase E fragment, even though their affinities measured by filter binding assay were ranging between 25 and 193 nM. One could argue that differences in experimental conditions and the cross-linking step included in the filter binding assay might lead to divergent values, the large discrepancy in measured affinities is nonetheless peculiar.

With regards to the conformation and base pairing status, the homogenous broadening of the imino proton resonances upon protein addition was observed for all other constructs for both RhlB and the RhlB/RNase E (694-790) complex, too, with no disappearance or formation of individual imino proton signals and no chemical shift perturbations. However, the fact that the allosteric interaction of RNase E (694-790) evidently changed RhlB's affinity towards its RNA substrates, provides a strong indication that RhlB's RNA interaction site is altered in some way. In the following chapter we want to further investigate how the presence of RNase E (694-790) affects the RNA binding process of RhlB.

5.6 Structural Changes in RhlB's RNA Binding Pocket introduced by RNase E

After revealing that the allosteric interaction with RNase E (694-790) directly affected the binding affinity of RhlB for its RNA substrates, we wanted to gain further insight into whether RNase E causes structural changes in the RNA binding pocket of the helicase. We continued to use NMR resonances of the RNA to provide an informative basis since the RNA conformation during binding should be sensitive to conformational changes in the binding cavity of the protein.

So far, neither selective peak appearance/ disappearance of RNA imino protons nor chemical shift perturbations indicative of conformational change could be observed for any of the tested RNA substrates in section 5.5. Nonetheless, there are multiple possible explanations for a structural change in the binding pocket despite the lack of observable alterations in the imino proton pattern: On the one hand, the structural changes might affect the fraying ends of the RNA duplexes, where the increased solvent exchange caused line broadening of the resonance signals beyond detectability. On the other hand, if we refer to the published crystal structures of other RNA bound helicases, we can see that the amino acids in the binding pocket directly coordinate only the RNA's backbone phosphate and 2'OH groups of the sugar moiety. It is therefore possible, that the imino protons are too distant from these nuclei to be sensitive to changes in the chemical surrounding. We therefore decided to switch the focus from the imino protons to the aromatic resonances of the nucleobases while increasing the spectral resolution with 2D NMR experiments. The chemical shifts of nucleobase resonances C2H2, C6H6 and C8H8 are not only well dispersed in a ^{13}C HSQC, which facilitates resonance assignment, they are also sensitive to their chemical surrounding and the RNAs conformation, making them suitable reporter signals in close proximity to the proposed interaction point. While the resonances of the sugar moieties would be even better reporter resonances in terms of distance to the coordinating amino acids in the binding pocket, their spectral overlap - particularly between C2' and C3' – would complicate a continuous resonance tracking during titration and impede the resonance assignment.

We performed 2D NMR titration experiments analogue to the ones described in section 5.5 with selectively ^{13}C labelled RNA substrate, stepwise addition of up to 2 equivalents of RhlB or RhlB/RNase E (694-790) complex and recording of ^{13}C HSQC spectra for each titration step. The 5' overhang duplex J2Δ16-M2 was used as RNA substrate, as this construct exhibited the most favourable affinity values for the RhlB/RNase E (694-790) complex to ensure complete binding. To further reduce the spectral overlap, only the 21 Nt M2 strand was ^{13}C labelled.

First, a stepwise addition of up to two equivalents of RhlB to the ^{13}C labelled 5'-tailed RNA duplex was recorded. The experiment showed a homogenous intensity decrease throughout all monitored nucleobase resonances matching the behaviour of the

previously tracked imino protons. Overlayed in Figure 57A are the first and last titration step of this experiment with 1D projections of each titration step for some exemplary resonances. At 2 equivalents of protein only the most intense resonances are still visible, while all other peaks are broadened beyond detectability. For a stepwise progression of all resonances, see appendix Figure 74. No significant chemical shift perturbations could be observed for any nucleobase resonance, which indicated that there is no notable change in chemical surrounding or in conformation experienced by the nucleobase resonances upon binding to the helicase.

Significant differences were detected for the same titration experiment with RhlB/RNase E (694-790) complex. While the initial nucleobase resonances of the 5' tailed duplex decreased evenly as before, a set of new nucleobase resonances arose with increasing protein concentration, implying the formation of a second conformationally distinct population of nucleotides in the M2 strand. Their high intensity, narrow line width and relative peak position within the respective nucleobase chemical shift regions indicate that these resonances map to a more dynamic and flexible conformation, presumably with unpaired residues. Subsequent denaturing gel electrophoresis of the fully titrated NMR sample ruled out any degradation of the RNA substrate as the cause for these sharp resonances (See Appendix Figure 75). As annotated in Figure 57B with black arrows, a total of nine distinct new resonances could be identified. Based on the chemical shift region and multiplicity of the peaks, three could be assigned to be C2H2 resonances of adenine residues, three were identified as pyrimidine C6H6 resonances and three as purine C8H8 resonances. Taken together, those nucleobase resonances resemble a set of six nucleotides, of which three are adenine residues and the remaining three are pyrimidines. When trying to identify those six nucleotides within the sequence of the 21 Nt M2 strand (as only M2 is ^{13}C labelled), the following assumption was made: It is known from various crystal structures of other DEAD-box helicases with single stranded RNA substrates that the amino acid coordination within the RNA binding pocket encompass 5 consecutive nucleotides[15,21]. We therefore assume that the sequence of six unpaired nucleotides represent a single binding position within the binding pocket of RhlB. If we furthermore take the apparent flexibility and dynamic of the peak population into account, a terminal position within the RNA sequence is more likely, where the ends can fray and produce a partially single stranded conformations more easily than in the middle of the strand. Consequently, the most reasonable sequence that fulfils those criteria and matches the required nucleotides are the six nucleotides of the 3' end of the M2 strand U_{21} to U_{15} (3'-UUAAAU…).

To corroborate this assumption, we compared the ^{13}C HSQC of this newly arising peaks with the spectra of two constructs for which the 3' end of M2 is also unpaired: the 3' tailed duplex J2Δ8-M2 and the lone M2 strand. In both constructs, as illustrated in panel D of Figure 57, the 3' terminal eight nucleotides (A_{14} to U_{21}) are unpaired. In Figure 57C, their respective ^{13}C HSQC spectra are overlayed with the titration start and endpoints from Panel B. Noticeably, the new peak population (coloured in blue and annotated with

black arrows) has a striking chemical shift overlap with the unpaired 3' tail of M2 and J2Δ8-M2. The C6H6 nucleobase resonances of the newly arising pyrimidines are in a significant agreement with the single stranded resonances of U_{21}, U_{20} and U_{16} as assigned for the 3' tailed duplex construct. For the C8H8 resonances the chemical shift alignment with unpaired A_{17}, A_{18} and A_{19} is not as clear for all three resonances but we can still see a considerable spectral overlap for two of the three purine residues. For the C2H2 resonances we can also see are clear agreement for two of the three adenines residues. The third resonance signal is slightly shifted but nonetheless resonates at a chemical shift typically reserved for adenines devoid of stable base pair formation. The small changes in chemical shift were acceptable and somewhat likely considering that we do compare spectra of free RNA with RhlB bound RNA. The experiments also nicely highlighted the strengths of 2D heteronuclear NMR spectroscopy to elucidate conformational changes that are not detectable via 1H 1D NMR experiments due to peak broadening as the nucleobases produced a clearer readout of the RNA duplex conformation even for the blunt end.

With that we conclude that the second population that is forming upon binding of the RNA to the RhlB/RNase E (694-790) complex represents a partial duplex opening encompassing the 3' terminal nucleotides U_{21} to U_{16} of the ^{13}C labelled M2 strand, while the remaining duplex stays intact. Based on the dynamic appearance of these newly arising resonances in the M2 strand of the J2Δ16-M2 duplex, we could furthermore assume, that the interaction with RhlB that introduces this partially single stranded conformation is taking place at the 6 terminal nucleotides of the 5' end of the counter strand J2Δ16.

To further validate this assumption, attempts have been made to investigate the RNA duplex with inverse labelling (^{13}C labelled J2Δ16 strand) to directly observe how the counter strand is affected by the binding to RhlB/RNase E (694-790). Unfortunately, *in vitro* transcription and purification of the labelled J2Δ16 strand was challenging and, in the end, not successful.

The results presented in this chapter nonetheless provide novel insight into the unique stimulating interplay between RhlB and RNase E: The allosteric activation of RhlB by RNase E (694-790) introduced a conformational transition in a stretch of 6 nucleotides of the 5'-tailed duplex substrate from a base paired to an unpaired conformation. The remaining duplex however remained intact, as ^{1}H NMR titration spectra did not show a complete loss of imino proton resonances even at 4 equivalents of protein over RNA. Those data furthermore represent the first evidence of the DEAD-Box helicase structurally changing the RNA duplex substrate even in absence of ATP.

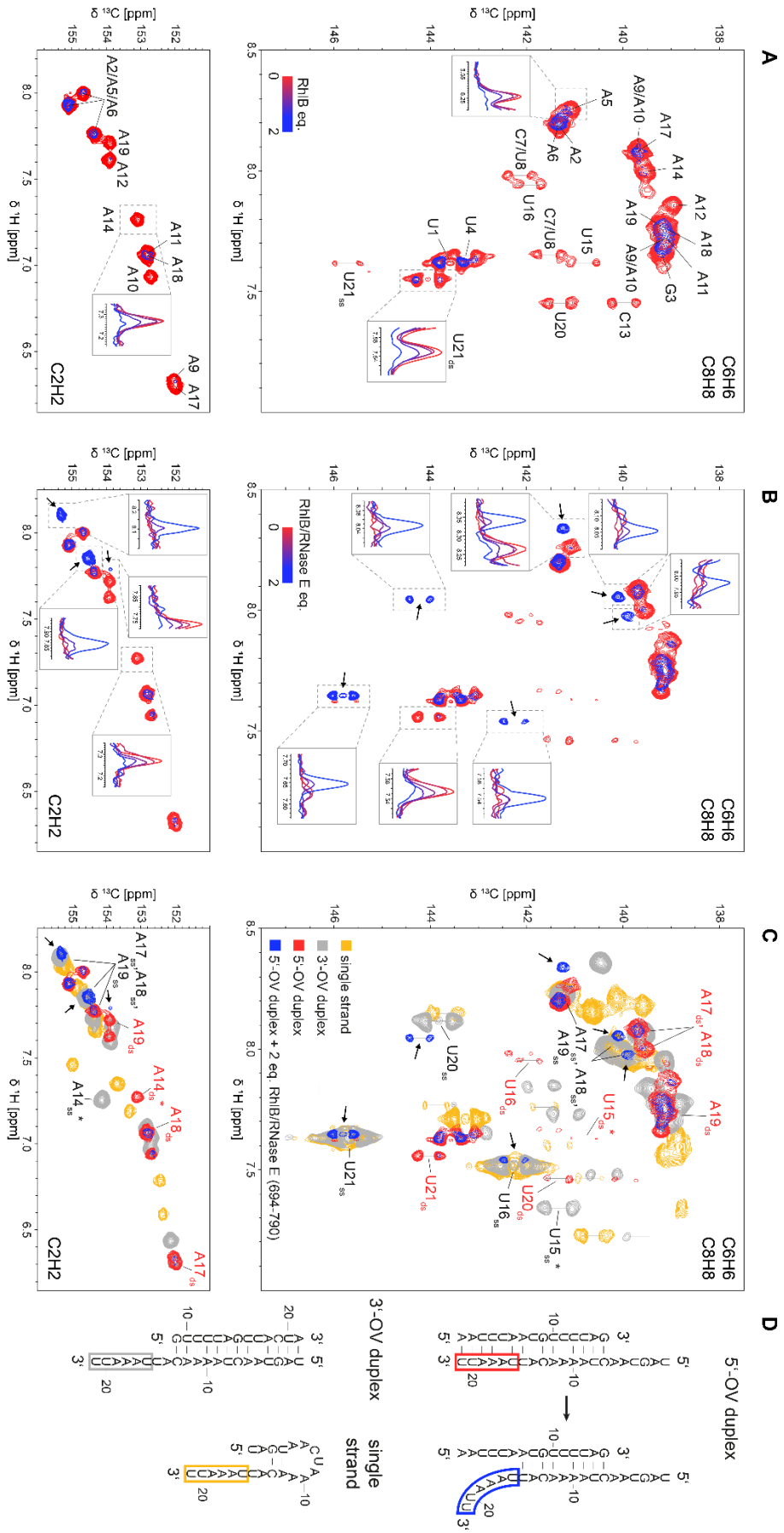
A
B
C
D
C6H6
C8H8
C2H2
δ ¹³C [ppm]
δ ¹H [ppm]
RhlB eq.
RhlB/RNase E eq.
single strand
3'-OV duplex
5'-OV duplex
5'-OV duplex + 2 eq. RhlB/RNase E (694-790)
5'-OV duplex
3'-OV duplex
single strand

Figure 57. Conformation of 5' overhang RNA substrate during RhlB binding in absence and presence of RNase E (694-790). Superposition of ^{13}C HSQC spectra of the 5'-overhang RNA duplex titrated with up to 2 equivalents of RhlB (A) or RhlB/RNase E (694-790) (B). Overlayed are the spectra of the aromatic regions for 0 (red) and 2 (blue) equivalents protein over RNA with exemplary 1D projections for each titration step to illustrate the overall signal intensity decrease. Spectra were recorded with 52 scans at 700 or 800 MHz for RhlB and RhlB/RNase E (694-790), respectively, at 288 K. Titrations were performed with 100 µM 5'-OV RNA substrate, that was exclusively ^{13}C labelled in the 21 Nt strand. (B) Newly arising signals are indicated with black arrows and additional 1D projections for each titration step. (C) Superposition of ^{13}C HSQC spectra from (B) with spectra of 100 µM 3'-OV duplex (grey) and 100 µM 21 Nt single strand construct (yellow). ^{13}C HSQC of the 3'-OV duplex was recorded at 288 K and 600 MHz with 32 scans, while the 21 Nt single strand was recorded at 288 K and 700 MHz with 52 scans. Newly arising signals from (B) were again indicated with arrows and nucleotides corresponding to the base paired and unpaired conformation of U21 to U16 were labelled accordingly, as assigned for the 5' and 3' tailed duplex. U15 and A14 are marked with an asterisk (*), as they represent the limit in the RNA sequence were a spectral alignment of 3'-OV RNA and single strand was present based on structural similarity. For the full resonance assignment of both 5'-OV and 3'-OV duplexes, see Figure 29 and Figure 32. (D) Secondary structures of 3'-OV, 5'-OV and 21 Nt single stranded RNA constructs as determined by NMR spectroscopy. Relevant nucleotides U21 to U16 are highlighted in the corresponding strands.

5.7 Discussion

The research presented in this chapter intended to investigate the potential interplay between the RNA substrate selectivity of *E. coli* DEAD-Box helicase RhlB and its degradosome complex partner RNase E. To gain novel insight into the helicases reaction mechanism the experimental focus was predominantly put on the NMR spectroscopic analysis of the RNA substrate. This unique approach came not only with the requirement for large quantities of both isotope-labelled and unlabelled RNAs and proteins but the differing demands for measurement conditions also had to be thoroughly weighed up to find the most suitable compromise while also making efforts to maximise comparability between experimental techniques.

Firstly, both RhlB and the RNase E fragments (628-843) and (694-790) were successfully synthesized and purified before the protein interaction could be demonstrated for both complexes. CD spectra and ^{15}N HSQC experiments provided evidence that RNase E (694-790) adopts a disordered structure, which has previously only been predicted, and that the RNase E fragment remains unstructured when binding to RhlB. Considering that RNase E (694-790) only includes half of the segment predicted to form coiled coils (AS 685-712; see Figure 10), it is plausible that the complete section is required to achieve structure. Under NMR conditions RhlB adopts a stable fold with disordered or flexible regions, that results in a reduced number of detectable resonances (only 64%). These data agree with the predicted structural model of RhlB, which also features the highly unstructured elongated C-terminal extension as well as several flexible loop regions. Upon binding of RNase E (694-790) well-dispersed NMR spectra at significantly lower ion concentrations could be recorded for RhlB than for the isolated protein, indicating an overall stabilizing effect of RNase E (694-790) on the otherwise dynamic open conformation of the helicase without altering its global structure.

Measurement of RhlB's ATP turnover rate using RNA substrates with distinct strand features revealed the following: while in absence of either RNase E fragment no ATPase activity was detected for RhlB with any of the investigated RNA constructs, RNA strand features have a significant influence on the overall ATP turnover rate of the helicase when bound to RNase E. The addition of RNase E (694-790) caused a significant stimulation of ATPase activity (up to 6-fold higher) for all RNA substrates, with the highest absolute rate for the 5' tailed RNA duplex J2Δ16-M2 and the single strand M2. The addition of the larger RNase E fragment (628-843) encompassing two RNA binding sites resulted in an even larger boost of ATPase activity for all constructs with single stranded features, especially for 5' tailed RNA duplex J2Δ16-M2 and the single strand M2. The data suggested that those flanking RNA binding sites of RNase E indirectly contribute to the helicase's reaction cycle, possibly through recruiting of specifically single stranded RNA substrate or capturing of released single strand. Not only are those new findings in great agreement with the RhlB's 5'-tailed RNA preferences previously reported by Chandran *et*

al. in unwinding experiments, but it also highlights the complex interplay of the degradosome components[18].

After successfully providing evidence that the allosteric activation of RhlB's ATP turnover rate could be achieved under NMR spectroscopic conditions using ^{31}P kinetic mixing experiments, the influence of RNase E on RhlB's binding to the distinct RNA substrates in absence of ATP was determined using ^{1}H NMR titration experiments, where up to 4 equivalents of RhlB or RhlB/RNase E complex was added to the RNA. No chemical shift perturbations (CSPs) were observed for the imino protons for any of the titrations, but homogeneous intensity decrease and peak broadening for all measured RNA imino protons could be detected, indicating the formation of a higher molecular weight complex and thus the protein binding. Here the experimental data provided the first evidence that RhlB exhibits an intrinsic preference in form of an increased affinity for a 5'-tailed duplex over 3' tailed or blunt ended constructs. Moreover, we could demonstrate that the allosteric binding of RNase E (694-790) selectively increases the affinity of RhlB for 5' and 3' single strand tailed duplexes while decreasing the affinity for blunt end duplexes. It was suspected that an alteration in RNA affinity must be accompanied with a change of RhlB's interaction with the RNA. To assess this assumption the titrations with RhlB and RhlB/RNase E (694-790) were repeated with isotope-labelled 5'-tailed duplex J2Δ16-M2 using ^{13}C HSQC, thereby shifting the focus onto the nucleobase resonances of the RNA. With those experiments it could be further revealed that, while upon binding to RhlB no CSPs or other indications of a conformational change could be detected for the RNA, the binding of RhlB in complex with RNase E (694-790) introduced the formation of six new nucleotide resonances in the labelled M2 strand. By referencing ^{13}HSQC spectra of both the M2 single strand as well as the 3'-tailed duplex J2Δ8-M2, the resonances could be assigned to an unpaired conformation of M2's 3' end. Also considering the sharp linewidth and therefore dynamic nature of the new resonances, the data indicate that RhlB/RNase E (694-790) introduced a partial strand opening of six nucleotides on the blunt end of the RNA substrate, as the remaining base pairs were still intact. The data therefore provide the first evidence of a DEAD-Box helicase that – upon allosteric activation of its interaction partner – alters the conformation of its substrate RNA even in absence of ATP.

Due to the narrow linewidth and implied dynamic of the detected unpaired M2 resonances, we infer that the grip of the helicase is on the counter strand J2Δ16. The length of the partial strand opening of 5-6 nucleotides matches the length of nucleotides coordinated in the RNA binding pocket of RhlB. We therefore postulate that the allosteric effect of RNase E (694-790) alters the grip on the RNA duplex in the RNA binding pocket in a way that destabilizes the duplex. One must consider that we do not know whether all RNA substrates would undergo partial duplex opening in presence of RhlB/RNase E (694-790) or whether this is merely a consequence of the already instable blunt end of the J2d16-M2 duplex specifically. Nonetheless, we can conclusively say that the grip on the

substrate RNA is altered in a way that it increases the affinity for the RNA substrate while simultaneously destabilizing the duplex.

How can the results presented here be brought together into a cohesive picture that builds on the published literature? The K_D values measured for RhlB show the helicases intrinsic preference for duplexes with 5' single strands over 3'-tailed or blunt end duplexes. This poses the question how this selectivity can be achieved given that the RNA binding site of DEAD-Box helicases is reportedly sequence unspecific[14]. Here the CTE of RhlB most likely comes into play. Firstly, it has been shown that RhlB lacking its C-terminal extension binds RNA more weakly than full-length RhlB, indicating that the basic tail plays a role in RNA association[18]. Secondly, X-rax crystallography data of Mss116p, another DEAD-Box helicase featuring a flexible CTE, have demonstrated that the protein's CTE assists in RNA binding by providing additional binding contacts with the RNA outside of the RNA binding pocket[22]. A similar mechanism might be at play for RhlB, were the CTE provides specific recognition for the 5' single stranded element of the bound RNA, while the double stranded section is bound by the RNA binding site. An RNA substrate that lacks the single stranded section or features a single strand with the wrong polarity is consequently bound less tightly.

However, in absence of the degradosome complex partner, RhlB's ATP turnover and unwinding activity is barely above the detection threshold and so any preferences for specific RNA substrates are at this point unnoticed and without effect. As the NMR data have shown, allosteric binding of RNase E (694-790) does not only increase the helicases affinity for non-blunt duplexes but simultaneously decreases the affinity for blunt-end duplexes. The increase in affinity for the 5' tailed duplex is accompanied by a change in the RNA binding pocket that leads to a destabilization and partial opening of the duplex. We therefore propose that the allosteric stimulation of RNase E tightens the grip on the RNA in the binding pocket of RhlB, thereby increasing the affinity and that these altered interaction contacts force the bound part of the duplex into a conformation that destabilizes the base pairing. The fact that blunt end constructs even experience a reduced binding affinity indicates that both the alignment of the RNA substrate through the 5' single strands interaction with the CTE as well as the allosteric alteration in the binding pocket are required for full beneficial stimulation of RhlB's activity. Since it was demonstrated by Stampfl *et al.* and others that the duplex stability directly correlates with the helicases unwinding rate, even a partial destabilization could significantly increase RhlB's activity[92,124]. We do however have to acknowledge that the allosteric activation of RNase E is most likely not limited to the findings reported for the RNA binding pocket but also extends to the ATP binding site. Whether this is the case and RNase E also directly affects the hydrolysis reaction remains to be elucidated. It is indicated by the findings from Bruce *et al.* who reported changes in RhlB's surface exposure for motif VI and V_a, which are involved in ATP binding and hydrolysis as well RNA-ATP communication[127].

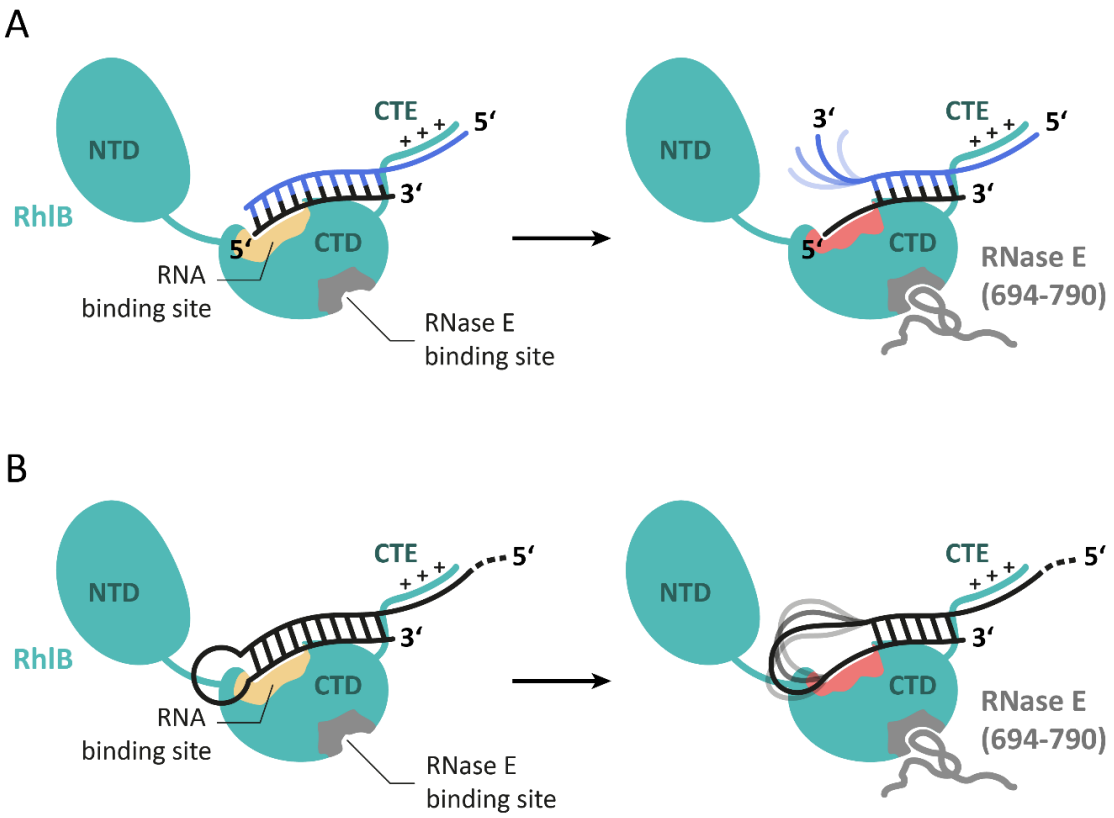

Figure 58. Proposed model of RhlB's interaction with a 5' tailed substrate RNA in absence and presence of RNase E (694-790). A. Schematic representation of RhlB in open conformation with bound 5'-tailed RNA duplex J2Δ16-M2 in the absence of ATP. Binding sites for RNase E (grey) and RNA substrate (yellow) are highlighted within the C-terminal domain (CTD). The C-terminal extension (CTE) of RhlB interacts with the 5' single stranded extension of the RNA. Upon allosteric binding of RNase E (694-790) the interaction site is altered (red) in a way that the 5' terminal nucleotides of the shorter strand (black) are bound more tightly. This leads to a separation of the terminal base pairs and leaves the 3' end of the 21 Nt strand (blue) dynamic and flexible. B. Schematic representation of the same model in vivo: the bound double strand is part of a stem loop structure in a longer RNA transcript.

Finally, we want to put the results into context of the degradosome *in vivo*. As is known from literature, RhlB is required in the degradosome to unfold secondary structure elements in the RNA strand to facilitate the full degradation by 3'-to-5' exonuclease PNPase[4,10]. As PNPase degrades the RNA in a 3'-to-5' direction and encounters a stable secondary structure within the RNA strand it pauses on the 3' single strand downstream of the stable RNA duplex. The preference for single strands with a 5' polarity in RhlB therefore allows the helicase to bind upstream from the secondary structure element where the 5' single strand is available. This interaction is most likely aided or mediated by the RNA binding sites RBD and AR2 of RNase E. To ensure RhlB does not unfold RNA transcripts outside of the regulated context of the degradosome, RhlB's activity is only boosted once it is bound to its complex partner RNase E. While the helicase presumably recognizes the 5' single strand of its substrate via its CTE and positions the target duplex in the binding pocket, RNase E allosteric effects on the RNA binding pockets destabilize the bound double strand, thereby allowing the helicase to unwind the substrate easier and more quickly. Figure 58 summarizes this proposed interaction both *in vitro* as it was

observed with duplex J2Δ16-M2 as well as *in vivo*, where the double strand would be extended into a stem loop structure.

Chapter 6 Concluding Remarks and Perspective

This PhD thesis aimed to contribute to the understanding of the functional interplay of DEAD-Box helicase RhlB and endoribonuclease RNase E, essential protein components of the multi-enzyme complex degradosome in *E. coli*. It is their interaction and the subsequent activation of the DEAD-Box helicase that drives forward the uninterrupted and consistent degradation of mRNA transcripts *in vivo*. The utilization of NMR spectroscopy and oligonucleotide substrates with specific strand properties enabled a unique RNA-centred approach to this insufficiently understood protein interaction. The conducted investigations determined that RhlB has an innate preference for 5'-tailed duplex substrates, that originates in an increased RNA affinity and is reflected in both the helicases ATPase as well as unwinding activity[18]. It could further be revealed that the allosteric binding of RNase E causes a change in RhlB's RNA binding pocket. This leads to a change in the helicases RNA affinity towards the physiologically favoured ss-tailed substrates but also introduces a partial strand opening in the 5' tailed duplex substrate as observed with 2D NMR spectroscopy. These findings provide intriguing new insights into the complex interplay between DEAD-Box helicases and their cognate interaction partners, as base pair opening as a consequence of partial duplex destabilization has so far not been reported for DEAD-Box helicases in absence of ATP. For future studies, more physiological complex compositions would be an interesting strategy to follow. Performing 2D NMR titration experiments with longer and hairpin-forming RNAs rather than short hetero-duplexes would better mimic *in* vivo mRNA transcripts and provide further insight into the duplex destabilization caused by RNase E's allosteric binding.

However, the focus on an RNA-centred characterization of this RNP complex also presented its own set of challenges, starting with the numerous compromises between RNA and protein measurement conditions and the consequent trade-off in resolution and sensitivity of NMR spectra over sheer feasibility. Consequently, we see this approach only rarely in literature and instead protein-centred structural studies build the majority[180]. With the ambition to incorporate more and larger components of a multi-protein complex, the large molecular mass makes structure determination by NMR spectroscopy laborious if not impossible. It is this size limitation that in the past decades has advanced cryogenic electron microscopy (cryo-EM) as a more favourable technique for the structure determination of large macromolecular structures such as RNP complexes. At the same time, cryo-EM's major drawback is its inability to capture structural dynamic. NMR spectroscopy on the other hand can provide an assortment of experiments that allow the measurement of such molecular dynamics and is therefore an essential technique to answer specific functional questions. So, in order to further advance both structural and functional studies of macromolecular complexes in the future, a combination of both techniques should be aspired, where NMR spectroscopy can provide insight into the structural dynamics and cryo-EM facilitates the determination of the 3-dimensional structure. Because the reality is, within the cell proteins involved in the most

essential metabolic processes are barely working in isolation and to understand these intricate machineries we need to shift our focus more towards the investigation of the full complexes. This way we might one day fully comprehend the interlaced communication pathways and regulatory networks of the dynamic processes that take place *in vivo*.

Chapter 7 References

[1] C. M. Arraiano, J. M. Andrade, S. Domingues, I. B. Guinote, M. Malecki, R. G. Matos, R. N. Moreira, V. Pobre, F. P. Reis, M. Saramago, I. J. Silva, S. C. Viegas, *FEMS Microbiol. Rev.* **2010**, *34*, 883–923.

[2] B. Py, H. Causton, E. Mudd, C. Higgins, *Mol. Microbiol.* **1994**, *14*, 717–729.

[3] R. McLaren, S. Newbury, G. Dance, H. Causton, C. Higgins, *J. Mol. Biol.* **1991**, *221*, 81–95.

[4] B. Py, C. Higgins, H. Krisch, A. Carpousis, *Nature* **1996**, *381*, 169–72.

[5] K. McDowall, S. Cohen, *J. Mol. Biol.* **1996**, 349–355.

[6] N. Vanzo, Y. Li, B. Py, E. Blum, C. Higgins, L. Raynal, H. Krisch, A. Carpousis, *Genes Dev.* **1998**, *12*, 2770–2781.

[7] A. Leroy, N. Vanzo, S. Sousa, M. Dreyfus, A. Carpousis, *Mol. Microbiol.* **2002**, *45*, 1231–1243.

[8] H. Strahl, C. Turlan, S. Khalid, P. Bond, J. Kebalo, P. Peyron, L. Poljak, M. Bouvier, L. Hamoen, B. Luisi, A. Carpousis, *PLOS Genet.* **2015**, *11*, e1004961.

[9] A. J. Callaghan, M. J. Marcaida, J. A. Stead, K. J. McDowall, W. G. Scott, B. F. Luisi, *Nature* **2005**, *437*, 1187–1191.

[10] G. Coburn, X. Miao, D. Briant, G. Mackie, *Genes Dev.* **1999**, *13*, 2594–603.

[11] M. Fairman-Williams, U. Guenther, E. Jankowsky, *Curr. Opin. Struct. Biol.* **2010**, *20*, 313–24.

[12] J. Caruthers, D. McKay, *Curr. Opin. Struct. Biol.* **2002**, *12*, 123–133.

[13] M. Singleton, M. Dillingham, D. Wigley, *Annu. Rev. Biochem.* **2007**, *76*, 23–50.

[14] A. Henn, M. Bradley, E. De La Cruz, *Annu. Rev. Biophys.* **2012**, 247–269.

[15] P. Linder, E. Jankowsky, *Nat. Rev. Mol. Cell Biol.* **2011**, *12*, 505–16.

[16] T. Bizebard, I. Ferlenghi, I. Iost, M. Dreyfus, *Biochemistry* **2004**, *43*, 7857–7866.

[17] J. Worrall, F. Howe, A. McKay, C. Robinson, B. Luisi, *J. Biol. Chem.* **2008**, *283*, 5567–5576.

[18] V. Chandran, L. Poljak, N. Vanzo, A. Leroy, R. Miguel, J. Fernandez-Recio, J. Parkinson, C. Burns, A. Carpousis, B. Luisi, *J. Mol. Biol.* **2007**, *367*, 113–32.

[19] A. Callaghan, J. Aurikko, L. Ilag, G. Grossmann, V. Chandran, K. Kühnel, L. Poljak, A. Carpousis, C. Robinson, M. Symmons, B. Luisi, *J. Mol. Biol.* **2004**, *340*, 965–79.

[20] V. Khemici, I. Toesca, L. Poljak, N. Vanzo, A. Carpousis, *Mol. Microbiol.* **2004**, *54*, 1422–30.

[21] T. Sengoku, O. Nureki, A. Nakamura, S. Kobayashi, S. Yokoyama, *Cell* **2006**, *125*, 287–300.

[22] A. Mallam, M. Del Campo, B. Gilman, D. Sidote, A. Lambowitz, *Nature* **2012**, *490*, 121–125.

[23] B. Y. M. Nirenberg, P. Leder, M. Bernfield, R. Brimacombe, J. Trupin, F. Rottmant, C. O. Neal, *Biochemistry* **1965**, *53*, 1161–1168.

[24] M. D. Berg, C. J. Brandl, *RNA Biol.* **2021**, *18*, 316–339.

[25] D. P. Burma, A. K. Srivastava, S. Srivastava, D. Dash, D. S. Tewari, B. Nag, *J. Biosci.* **1985**, *8*, 757–766.

[26] M. V. Iorio, D. Palmieri, *Front. Oncol.* **2019**, *9*, 1–2.

[27] S. Valadkhan, A. Mohammadi, Y. Jaladat, S. Geisler, *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 11901–11906.

[28] A. Fire, S. Xu, M. Montgomery, S. Kostas, S. Driver, C. Mello, *Nature* **1998**, *391*, 806–811.

[29] S. M. Hammond, E. Bernstein, D. Beach, G. J. Hannon, *Nature* **2000**, *404*, 293–296.
[30] A. Reining, S. Nozinovic, K. Schlepckow, F. Buhr, B. Fürtig, H. Schwalbe, *Nature* **2013**, *499*, 355–359.
[31] R. Rieder, K. Lang, D. Graber, R. Micura, *Chembiochem* **2007**, *8*, 896–902.
[32] J. Barrick, R. Breaker, *Genome Biol.* **2007**, *8*, R239.
[33] W. Gosda, U. Wallwitz, H. Schmidt, *Die Kult.* **1965**, *13*, 177–237.
[34] N. Leontis, J. Stombaugh, E. Westhof, *Nucleic Acids Res.* **2002**, *30*.
[35] W. Fuller, A. Hodgson, *Nature* **1967**, *215*, 817–821.
[36] S. Nozinovic, B. Fürtig, H. Jonker, C. Richter, H. Schwalbe, *Nucleic Acids Res.* **2010**, *38*, 683–694.
[37] J. Buck, J. Noeske, J. Wöhnert, H. Schwalbe, *Nucleic Acids Res.* **2010**, *38*, 4143–4153.
[38] S. Arnott, *Prog. Biophys. Mol. Biol.* **1971**, *22*, 179–213.
[39] S. Millevoi, H. Moine, S. Vagner, *Wiley Interdiscip. Rev. RNA* **2012**, *3*, 495–507.
[40] J. Watson, F. Crick, *Nature* **1953**, *171*, 737–741, 964–967.
[41] M. C. Wahl, M. Sundaralingam, *Biopolymers* **1997**, *44*, 45–63.
[42] G. Dougherty, *Life Sci.* **1983**, *33*, 2049–2060.
[43] J. L. Leroy, K. Gehring, A. Kettani, M. Guéron, *Biochemistry* **1993**, *32*, 6019–6031.
[44] M. L. Bochman, K. Paeschke, V. A. Zakian, *Nat. Rev. Genet.* **2012**, *13*, 770–780.
[45] J. T. Grün, H. Schwalbe, *Biopolymers* **2022**, *113*, 1–15.
[46] C. E. Weinberg, Z. Weinberg, C. Hammann, *Nucleic Acids Res.* **2019**, *47*, 9480–9494.
[47] K. Kruger, P. J. Grabowski, A. J. Zaug, J. Sands, D. E. Gottschling, T. R. Cech, *Cell* **1982**, *31*, 147–157.
[48] K. Birikh, P. Heaton, F. Eckstein, *Eur. J. Biochem.* **1997**, *245*, 1–16.
[49] D. Treiber, J. Williamson, *Curr. Opin. Struct. Biol.* **2001**, *11*, 309–314.
[50] D. Herschlag, *J. Biol. Chem.* **1995**, *270*, 20871–20874.
[51] R. Russell, D. Herschlag, *J. Mol. Biol.* **2001**, *308*, 839–851.
[52] D. Treiber, M. Rook, P. Zarrinkar, J. Williamson, *Science (80-.).* **1998**, *279*, 1943–1946.
[53] B. Fürtig, P. Wenter, L. Reymond, C. Richter, S. Pitsch, H. Schwalbe, *J. Am. Chem. Soc.* **2007**, *129*, 16222–16229.
[54] D. Pörschke, in *Mol. Biol. Biochem. Biophys.*, **1977**, pp. 191–216.
[55] M. Rook, D. Treiber, J. Williamson, *J. Mol. Biol.* **1998**, *281*, 609–620.
[56] L. Rajkowitsch, D. Chen, S. Stampfl, K. Semrad, C. Waldsich, O. Mayer, M. Jantsch, R. Konrat, U. Bläsi, R. Schroeder, *RNA Biol.* **2007**, *4*, 118–130.
[57] R. Schroeder, A. Barta, K. Semrad, *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 908–919.
[58] R. Schroeder, R. Grossberger, A. Pichler, C. Waldsich, *Curr. Opin. Struct. Biol.* **2002**, *12*, 296–300.
[59] P. Redder, *Curr. Genet.* **2016**, *62*, 687–690.
[60] J. Bernstein, A. Khodursky, P. Lin, S. Lin-Chao, S. Cohen, *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 9697–9702.
[61] C. F. Higgins, R. S. McLaren, S. F. Newbury, *Gene* **1988**, *72*, 3–14.
[62] A. Lebreton, B. Séraphin, *Biochim. Biophys. Acta - Gene Regul. Mech.* **2008**, *1779*, 558–565.
[63] A. Carpousis, G. Van Houwe, C. Ehretsmann, H. Krisch, *Cell* **1994**, *76*, 889–900.
[64] H. Causton, B. Py, R. Mclaren, C. Higgins, *Mol. Microbiol.* **1994**, *14*, 731–741.
[65] T. Morita, H. Kawamoto, T. Inada, H. Aiba, *Mol. Microbiol.* **2004**, *54*, 1063–1075.
[66] J. Bernstein, P. Lin, S. Cohen, S. Lin-chao, *Proc. Natl. Acad. Sci.* **2004**.

[67] V. Khemici, A. Carpousis, *Mol. Microbiol.* **2003**, *51*, 777–790.
[68] M. Marcaida, M. Depristo, V. Chandran, A. Carpousis, B. Luisi, *Trends Biochem. Sci.* **2006**, *31*, DOI 10.1016/j.tibs.2006.05.005.
[69] S. Aït-Bara, A. Carpousis, *J. Bacteriol.* **2010**, *192*, 5413–23.
[70] S. Aït-Bara, A. Carpousis, Y. Quentin, *Mol. Genet. Genomics* **2015**, *290*, 847–862.
[71] Y. Kimhi, U. Littauer, *J. Biol. Chem.* **1968**, *248*, 231–240.
[72] J. Worrall, M. Górna, N. Crump, L. Phillips, A. Tuck, A. Price, V. Bavro, B. Luisi, *J. Mol. Biol.* **2008**, *382*, 870–883.
[73] V. Chandran, B. F. Luisi, *J. Mol. Biol.* **2006**, *358*, 8–15.
[74] T. Dendooven, G. Paris, A. Shkumatov, M. Islam, A. Burt, M. Kubańska, T. Yang, S. Hardwick, B. Luisi, *Mol. Microbiol.* **2022**, *117*, 102–120.
[75] C. Jain, *RNA* **2008**, *14*, 381–389.
[76] I. Iost, M. Dreyfus, *Nucleic Acids Res.* **2006**, *34*, 4189–97.
[77] Y. Ikeda, M. Yagi, T. Morita, H. Aiba, *Mol. Microbiol.* **2011**, *79*, 419–432.
[78] V. Carabetta, T. Silhavy, I. Cristea, *J. Bacteriol.* **2010**, *192*, 3713–3721.
[79] F. Lu, A. Taghbalout, *Biosci. Rep.* **2014**, *34*, 879–891.
[80] E. Blum, B. Py, A. J. Carpousis, C. F. Higgins, *Mol. Microbiol.* **1997**, *26*, 387–398.
[81] A. Miczak, V. Kaberdin, C. Wei, S. Lin-Chao, *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 3865–3869.
[82] A. Callaghan, J. Gu, Y. Redko, L. Ilag, M. Moncrieffe, M. Symmons, C. Robinson, K. Mcdowall, B. Luisi, *Biochemistry* **2003**, 13848–13855.
[83] M. Górna, A. Carpousis, B. Luisi, *Q. Rev. Biophys.* **2012**, *45*, 105–45.
[84] A. Carpousis, *Annu. Rev. Microbiol.* **2007**, *61*, 71–87.
[85] V. Khemici, L. Poljak, B. Luisi, A. Carpousis, *Mol. Microbiol.* **2008**, *70*, 799–813.
[86] L. Domínguez-Malfavón, L. Islas, B. Luisi, R. García-Villegas, J. García-Mena, *Biochimie* **2013**, *95*, 2034–41.
[87] A. Tejada-Arranz, V. de Crécy-Lagard, H. de Reuse, *Trends Biochem. Sci.* **2020**, *45*, 42–57.
[88] Y. Tsai, D. Du, L. Domínguez-Malfavón, D. Dimastrogiovanni, J. Cross, A. Callaghan, J. García-Mena, B. Luisi, *Nucleic Acids Res.* **2012**, *40*, 10417–10431.
[89] E. Jankowsky, M. Fairman, *Curr. Opin. Struct. Biol.* **2007**, *17*, 316–24.
[90] J. Lorsch, D. Herschlag, *Biochemistry* **1998**, *37*, 2180–2193.
[91] G. Rogers, N. Richter, W. Merrick, *J. Biol. Chem.* **1999**, *274*, 12236–12244.
[92] G. Rogers, W. Lima, W. Merrick, *J. Biol. Chem.* **2001**, *276*, 12598–12608.
[93] Q. Yang, E. Jankowsky, *Nat. Struct. Biol.* **2006**, *13*, 981–986.
[94] Q. Yang, M. Del Campo, A. Lambowitz, E. Jankowsky, *Mol. Cell* **2007**, *28*, 253–63.
[95] Q. Yang, E. Jankowsky, *Biochemistry* **2005**, 13591–13601.
[96] R. Russell, I. Jarmoskaite, A. Lambowitz, *RNA Biol.* **2013**, *10*, 44–55.
[97] P. Tijerina, H. Bhaskaran, R. Russell, *Proc. Natl. Acad. Sci. U. S. A.* **2006**.
[98] M. Del Campo, A. Lambowitz, *Mol. Cell* **2009**, *35*, 598–609.
[99] M. Fairman-Williams, U. Guenther, E. Jankowsky, *Curr. Opin. Struct. Biol.* **2010**, *20*, 313–324.
[100] A. Jankowsky, U. Guenther, E. Jankowsky, *Nucleic Acids Res.* **2011**, *39*, D338-41.
[101] J. Charollais, D. Pflieger, J. Vinh, M. Dreyfus, I. Iost, *Mol. Microbiol.* **2003**, *48*, 1253–1265.
[102] A. Turner, C. Love, R. Alexander, P. Jones, *J. Bacteriol.* **2007**, *189*, 2769–76.
[103] L. Sharpe Elles, M. Sykes, J. Williamson, O. Uhlenbeck, *Nucleic Acids Res.* **2009**, *37*, 6503–6514.

[104] P. Linder, P. Lasko, M. Ashburner, P. Leroy, P. Nielsen, K. Nishi, J. Schnier, P. Stonimski, *Nature* **1989**, *337*, 121–2.
[105] M. Hilbert, A. Karow, D. Klostermeier, *Biol. Chem.* **2009**, *390*, 1237–1250.
[106] I. Iost, M. Dreyfus, P. Linder, *J. Biol. Chem.* **1999**, *274*, 17677–17683.
[107] J. Banroques, M. Doère, M. Dreyfus, P. Linder, N. K. Tanner, *J. Mol. Biol.* **2010**, *396*, 949–966.
[108] C. Tsu, K. Kossen, O. Uhlenbeck, *RNA* **2001**, *7*, 702–709.
[109] M. Rudolph, D. Klostermeier, *Biol. Chem.* **2015**, *396*, 849–865.
[110] M. Varadi, S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon, A. Zídek, T. Green, K. Tunyasuvunakool, S. Petersen, J. Jumper, E. Clancy, R. Green, A. Vora, M. Lutfi, M. Figurnov, A. Cowie, N. Hobbs, P. Kohli, G. Kleywegt, E. Birney, D. Hassabis, S. Velankar, *Nucleic Acids Res.* **2022**, *50*, D439–D444.
[111] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. Kohl, A. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, *Nature* **2021**, DOI 10.1038/s41586-021-03819-2.
[112] J. Caruthers, E. Johnson, D. McKay, *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 13080–13085.
[113] R. Story, H. Li, J. Abelson, *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 1465–1470.
[114] B. Theissen, A. Karow, J. Köhler, A. Gubaev, D. Klostermeier, *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 548–553.
[115] A. Mallam, I. Jarmoskaite, P. Tijerina, M. Del Campo, S. Seifert, L. Guo, R. Russell, A. Lambowitz, *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 12254–12259.
[116] O. Cordin, N. K. Tanner, M. Doère, P. Linder, J. Banroques, *EMBO J.* **2004**, *23*, 2478–2487.
[117] K. J. Polach, O. C. Uhlenbeck, *Biochemistry* **2002**, *41*, 3693–3702.
[118] J. Banroques, O. Cordin, M. Doère, P. Linder, N. K. Tanner, *Mol. Cell. Biol.* **2008**, *28*, 3359–3371.
[119] F. Bono, J. Ebert, E. Lorentzen, E. Conti, *Cell* **2006**, 713–725.
[120] A. Henn, C. Wenxiang, D. Hackney, E. De La Cruz, *J. Mol. Biol.* **2008**, *377*, 193–205.
[121] F. Liu, A. Putnam, E. Jankowsky, *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 20209–14.
[122] F. Liu, A. Putnam, E. Jankowsky, *Biochemistry* **2014**, *53*, 423–33.
[123] C. M. Diges, O. C. Uhlenbeck, *EMBO J.* **2001**, *20*, 5503–5512.
[124] S. Stampfl, M. Doetsch, M. Beich-Frandsen, R. Schroeder, *RNA Biol.* **2013**, *10*, 149–56.
[125] Y. Chen, J. P. Potratz, P. Tijerina, M. Del Campo, A. M. Lambowitz, R. Russell, *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 20203–8.
[126] M. Doetsch, S. Stampfl, B. Fürtig, M. Beich-Frandsen, K. Saxena, M. Lybecker, R. Schroeder, *Nucleic Acids Res.* **2013**, *41*, 487–97.
[127] H. Bruce, D. Du, D. Matak-Vinkovic, K. Bandyra, R. Broadhurst, E. Martin, F. Sobott, A. Shkumatov, B. Luisi, *Nucleic Acids Res.* **2018**, *46*, 387–402.
[128] I. Jarmoskaite, R. Russell, *Wiley Interdiscip. Rev. RNA* **2011**, *2*, 135–152.
[129] R. Dickerson, H. Ng, *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 6986–6988.
[130] A. Gruber, R. Lorenz, S. Bernhart, R. Neuböck, I. Hofacker, *Nucleic Acids Res.* **2008**, *36*, 70–74.

[131] S. Bernhart, H. Tafer, U. Mückstein, C. Flamm, P. Stadler, I. Hofacker, *Algorithms Mol. Biol.* **2006**, *1*, 1–10.
[132] M. Andronescu, R. Aguirre-Hernández, A. Condon, H. Hoos, *Nucleic Acids Res.* **2003**, *31*, 3416–3422.
[133] M. Zuker, *Nucleic Acids Res.* **2003**, *31*, 3406–3415.
[134] R. Hartmann, *Handbook of RNA Biochemistry Volume 1*, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, **2005**.
[135] L. Marky, K. Breslauer, *Biopolymers* **1987**, *26*, 1601–1620.
[136] J. Reed, T. Reed, *Anal. Biochem.* **1997**, *254*, 36–40.
[137] N. Greenfield, G. Fasman, *Biochemistry* **1969**, *8*, 4108–4116.
[138] N. Greenfield, *Nat. Protoc.* **2007**, *1*, 2876–2890.
[139] M. Webb, *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 4884–4887.
[140] D. Wishart, C. Bigam, J. Yao, F. Abildgaard, H. Dyson, E. Oldfield, J. Markley, B. Sykes, *J. Biomol. NMR* **1995**, *6*, 135–140.
[141] W. Lee, M. Tonelli, J. Markley, *Bioinformatics* **2015**, *31*, 1325–1327.
[142] V. Sklenář, A. Bax, *J. Magn. Reson.* **1987**, *74*, 469–479.
[143] A. Shaka, P. Barker, R. Freeman, *J. Magn. Reson.* **1985**, *64*, 547–552.
[144] R. Wagner, S. Berger, *J. Magn. Reson. - Ser. A* **1996**, *123*, 119–121.
[145] S. Mori, C. Abeygunawardana, M. O'Neil-Johnson, P. Van Zljl, *J. Magn. Reson. Ser. B* **1995**, *108*, 94–98.
[146] W. Vuister, A. Bax, *J. Magn. Reson.* **1991**, *435*, 69.
[147] A. Palmer, J. Cavanagh, R. Byrd, M. Rance, *J. Magn. Reson.* **1992**, *96*, 416–424.
[148] V. Sklenář, R. Peterson, M. Rejante, J. Feigon, *J. Biomol. NMR* **1993**, *3*, 721–727.
[149] J. Schleucher, M. Schwendinger, M. Sattler, P. Schmidt, O. Schedletzky, S. Glaser, O. Sørensen, C. Griesinger, *J. Biomol. NMR* **1994**, *4*, 301–306.
[150] B. Simon, K. Zanier, M. Sattler, *J. Biomol. NMR* **2001**, *20*, 173–176.
[151] T. Hwang, A. Shaka, *J. Magn. Reson. - Ser. A* **1995**, *112*, 275–279.
[152] Z. Solyom, M. Schwarten, L. Geist, R. Konrat, D. Willbold, B. Brutscher, *J. Biomol. NMR* **2013**, *55*, 311–321.
[153] A. Favier, B. Brutscher, *J. Biomol. NMR* **2011**, *49*, 9–15.
[154] W. Bermel, I. Bertini, I. Felli, R. Peruzzini, R. Pierattelli, *ChemPhysChem* **2010**, *11*, 689–695.
[155] W. Bermel, I. Bertini, L. Duma, I. Felli, L. Emsley, R. Pierattelli, P. Vasos, *Angew. Chem. Int. Ed. Engl.* **2005**, 3089–3092.
[156] K. Mok, T. Nagashima, I. Day, J. Jones, C. Jones, C. Dobson, P. Hore, *J. Am. Chem. Soc.* **2003**, *125*, 12484–12492.
[157] M. Williamson, *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16.
[158] S. Freier, D. Alkema, A. Sinclair, T. Neilson, D. Turner, *Biochemistry* **1985**, *24*, 4533–4539.
[159] M. Serra, D. Turner, *Methods Enzymol.* **1995**, *259*, 242–261.
[160] M. Serra, J. Baird, T. Dale, B. Fey, K. Retatagos, E. Westhof, *RNA* **2002**, *8*, 307–323.
[161] S. Englander, T. Sosnick, J. Englander, L. Mayne, *Curr. Opin. Struct. Biol.* **1996**, *6*, 18–23.
[162] O. Zerbe, S. Jurt, *Applied NMR Spectroscopy for Chemists and Life Scientists*, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, **2014**.
[163] Y. Li, R. Breaker, *J. Am. Chem. Soc.* **1999**, *121*, 5364–5372.
[164] M. Oivanen, S. Kuusela, H. Lönnberg, *Chem. Rev.* **1998**, *98*, 961–990.
[165] K. Zhang, J. Hodge, A. Chatterjee, T. Moon, K. Parker, *Environ. Sci. Technol.* **2021**,

55, 8045–8053.
[166] H. Sternbach, R. Engelhardt, A. Lezius, *Eur. J. Biochem.* **1975**, *60*, 51–55.
[167] E. Genersch, C. Eckerskorn, F. Lottspeich, C. Herzog, K. Kühn, E. Pöschl, *EMBO J.* **1995**, *14*, 791–800.
[168] T. Middleton, B. Sugden, *J. Virol.* **1992**, *66*, 489–495.
[169] P. Romero, Z. Obradovic, K. Dunker, *Genome Inform. Ser. Workshop Genome Inform.* **1997**, *8*, 110–124.
[170] X. Li, P. Romero, M. Rani, A. K. Dunker, Z. Obradovic, *Genome Inform. Ser. Workshop Genome Inform.* **1999**, *10*, 30–40.
[171] A. Mallam, M. Del Campo, B. Gilman, D. Sidote, M. Alan, *Natu* **2013**, *490*, 121–125.
[172] V. Uversky, J. Gillespie, A. Fink, *Proteins Struct. Funct. Genet.* **2000**, *41*, 415–427.
[173] V. Uversky, *Int. J. Biochem. Cell Biol.* **2011**, *43*, 1090–1103.
[174] S. Mielke, V. Krishnan, *Prog. Nucl. Magn. Reson. Spectrosc.* **2009**, *54*, 141–165.
[175] C. Herbert, U. Schieborr, K. Saxena, J. Juraszek, F. De Smet, C. Alcouffe, M. Bianciotto, G. Saladino, D. Sibrac, D. Kudlinzki, S. Sreeramulu, A. Brown, P. Rigon, J. Herault, G. Lassalle, T. Blundell, F. Rousseau, A. Gils, J. Schymkowitz, P. Tompa, J. Herbert, P. Carmeliet, F. Gervasio, H. Schwalbe, F. Bono, *Cancer Cell* **2013**, *23*, 489–501.
[176] D. Sahu, M. Bastidas, S. Showalter, *Anal. Biochem.* **2015**, 17–25.
[177] M. Doetsch, S. Stampfl, B. Fu, K. Saxena, M. Lybecker, *Nucleic Acids Res.* **2013**, *41*, 487–497.
[178] O. Cordin, J. Banroques, N. Tanner, P. Linder, *Gene* **2006**, *367*, 17–37.
[179] L. Fielding, *Prog. Nucl. Magn. Reson. Spectrosc.* **2007**, *51*, 219–242.
[180] V. de Jesus, N. S. Qureshi, S. Warhaut, J. K. Bains, M. S. Dietz, M. Heilemann, H. Schwalbe, B. Fürtig, *Nat. Commun.* **2021**, *12*, 1–7.

Chapter 8 Appendix

8.1 Buffer, Media, and Solutions

Gel electrophoresis

Agarose gel solution
1% (w/v) agarose
TBE buffer

DNA loading buffer
0.1% (w/v) bromphenol blue
0.1% (w/v) xylene cyanol
25% (v/v) glycerol

TBE buffer
90 mM Tris
90 mM boric acid
2 mM EDTA
pH 8.3

TAE/TA buffer
40 mM Tris
0.1% (v/v) acetic acid
1 mM EDTA (omitted for TA buffer)
pH 8.0

Denaturing RNA-PAGE gel solution
10-15% (v/v) acrylamide/ bisacrylamide (29:1)
7 M urea
TBE buffer
0.1% (w/v) APS
0.1% (v/v) TEMED

Denaturing RNA loading buffer
0.1% (w/v) bromphenol blue
0.1% (w/v) xylene cyanol
100% formamide

Native RNA-PAGE gel solution
10-15% (v/v) acrylamide/ bisacrylamide (29:1)
TBE buffer
0.1% (w/v) APS
0.1% (v/v) TEMED

Native RNA loading buffer
0.1% (w/v) bromphenol blue
0.1% (w/v) xylene cyanol
25% (v/v) glycerol

GelRed Staining solution
0.001% (v/v) GelRed® Nucleic Acid Gel Stain (Biotium) in ddH_2O

Coomassie Staining Solution
10% (v/v) ethanol
5% (v/v) acetic acid
0.0025% (w/v) Coomassie brilliant blue G250
0.0025% (w/v) Coomassie brilliant blue R250

SDS-PAGE stacking gel solution
5% (v/v) acrylamide/ bisacrylamide (29:1)
0.1% (w/v) SDS
250 mM Tris/HCl
0.1% (w/v) APS
0.1% (v/v) TEMED
pH 6.8

SDS-PAGE resolving gel solution
10-15% (v/v) acrylamide/ bisacrylamide (29:1)
0.1% (w/v) SDS
375 mM Tris/HCl
0.1% (w/v) APS
0.1% (v/v) TEMED
pH 8.8

SDS-PAGE running buffer
25 mM Tris/HCl

SDS-PAGE sample loading buffer
250 mM Tris/HCl

250 mM glycine
1% (w/v) SDS
pH 6.8

4% (w/v) SDS
20% (v/v) glycerol
10% (v/v) 2-mercaptoethanol
0.25% (w/v) bromphenol blue
pH 6.8

Protein expression and purification

LB (lysogeny broth) medium
10 g/L tryptone
5 g/L yeast extract
10 g/L NaCl

SB (super broth) medium
35 g/L tryptone
20 g/L yeast extract
5 g/L NaCl
5 mL/L NaOH

TB (terrific broth) medium
20 g/L tryptone
24 g/L yeast extract
4% (v/v) glycerol
17 mM KH_2PO_4
72 mM K_2HPO_4
pH 7.4

SOC medium
2% (w/v) tryptone
0.5% (w/v) yeast extract
10 mM NaCl
2.5 mM KCl
10 mM $MgCl_2$
10 mM $MgSO_4$
20 mM glucose

LB agar plate
10 g/L tryptone
5 g/L yeast extract
10 g/L NaCl
15 g/L agar-agar

2000x vitamin mix
100 mM thiamine (vitamin B1)
100 mM biotin (vitamin B7)
100 mM niacin (vitamin B3)
10 mM cobalamin (vitamin B12)

M9 minimal medium
42 mM Na_2HPO_4
22 mM KH_2PO_4
8.5 mM NaCl
0.2 mM $CaCl_2$
2 mM $MgSO_4$
10 g/L glucose (2 g/L for ^{13}C glucose)
1 g/L NH_4Cl (1 g/L for ^{15}N NH_4Cl)
1x trace elements mix
1x vitamin mix
pH 7.4

5000x trace elements mix
50 mM $FeCl_3$
20 mM $CaCl_2$
10 mM $MnCl_2$
10 mM $ZnSO_4$
2 mM $CoCl_2$
2 mM $CuCl_2$
2 mM $NiSO_4$
2 mM Na_2MoO_4
2 mM Na_2SeO_3
2 mM H_3BO_3

Buffer A (RhlB, RNase E (628-843))
500 mM NaCl
50 mM Tris/HCl
10 mM 2-mercaptoethanol
10 mM imidazole

Buffer A (RNase E (694-790))
200 mM NaCl
100 mM KCl
50 mM Tris/HCl
5 mM $MgCl_2$

pH 8.3

Buffer B (RhlB, RNase E (628-843))
500 mM NaCl
50 mM Tris/HCl
10 mM 2-mercaptoethanol
500 mM imidazole
pH 8.3

Buffer C
450 mM KCl
75 mM Tris/HCl
5 mM DTT
pH 8.3

Buffer H
1 M KCl
25 mM Tris/HCl
5 mM DTT
pH 8.3

10 mM 2-mercaptoethanol
5 mM imidazole
pH 8.0

Buffer B (RNase E (694-790))
200 mM NaCl
100 mM KCl
50 mM Tris/HCl
5 mM $MgCl_2$
10 mM 2-mercaptoethanol
500 mM imidazole
pH 8.0

LiCl Buffer (RhlB, RNase E (628-843))
2 M LiCl
50 mM Tris/HCl
10 mM 2-mercaptoethanol
10 mM Imidazole
pH 8.3

RNA purification, EMSA and NMR experiments

HPLC buffer A
50 mM K_2HPO_4/KH_2PO_4
2 mM tetrabutylammonium bisulphate
pH 5.9

pH 6 NMR buffer
150 mM KCl
25 mM BisTris/HCl
pH 6.5

High-salt NMR buffer
450 mM KCl
25 mM Tris/HCl
5 mM DTT
pH 8.3

HPLC buffer B
50 mM K_2HPO_4/KH_2PO_4
2 mM tetrabutylammonium bisulphate
60% (v/v) acetonitrile
pH 5.9

pH 8 NMR buffer
150 mM KCl
25 mM Tris/HCl
5 mM DTT (for protein experiments)
pH 8.3

EMSA buffer
150 mM NaCl
25 mM Tris/HCl
5 mM DTT
1.5 mM $MgCl_2$
pH 8.3

8.2 Construct Sequences and Folding Predictions

8.2.1 Heterodimer Folding Predictions by Oligo Analyser

Depicted are the three most stable duplex folds predicted by OligoAnalyser™ and their corresponding Gibbs free energy.

J2h-M2

Delta G: -31.53 kcal/mol Base Pairs: 21

```
5'  UUAAUUUAAUGUUUUAGUUACUAGGGUUUAUGGCUGUUCGCCAUUU
    ||||||||||||||||||||
3'  UUAAAUUACAAAAUCAAUGAU
```

Delta G: -6.32 kcal/mol Base Pairs: 5

```
5'    UUAAUUUAAUGUUUUAGUUACUAGGGUUUAUGGCUGUUCGCCAUUU
      ||||| :  :  ::
3' UUAAAUUACAAAAUCAAUGAU
```

Delta G: -5.36 kcal/mol Base Pairs: 4

```
5' UUAAUUUAAUGUUUUAGUUACUAGGGUUUAUGGCUGUUCGCCAUUU
                                                ||||
3'                                             UUAAAUUACAAAAUCAAUGAU
```

J2-M2

Delta G: -31.53 kcal/mol Base Pairs: 21

```
5' AAUUUAAUGUUUUAGUUACUA
   ||||||||||||||||||||
3' UUAAAUUACAAAAUCAAUGAU
```

Delta G: -4.85 kcal/mol Base Pairs: 4

```
5'          AAUUUAAUGUUUUAGUUACUA
              |||| :    :
3' UUAAAUUACAAAAUCAAUGAU
```

Delta G: -4.85 kcal/mol Base Pairs: 4

```
5' AAUUUAAUGUUUUAGUUACUA
             |||| :    :
3'          UUAAAUUACAAAAUCAAUGAU
```

J2Δ8-M2

Delta G: -17.88 kcal/mol Base Pairs: 13

```
5'          GUUUUAGUUACUA
            |||||||||||||
3' UUAAAUUACAAAAUCAAUGAU
```

Delta G: -4.85 kcal/mol Base Pairs: 4
```
5' GUUUUAGUUACUA
     ||||  :    :
3' UUAAAUUACAAAAUCAAUGAU
```

Delta G: -3.89 kcal/mol Base Pairs: 3
```
5'            GUUUUAGUUACUA
               |||    :
3' UUAAAUUACAAAAUCAAUGAU
```

J2Δ8-M3

Delta G: -17.88 kcal/mol Base Pairs: 13
```
5' GUUUUAGUUACUA
   |||||||||||||
3' CAAAAUCAAUGAU
```

Delta G: -3.89 kcal/mol Base Pairs: 3
```
5'  GUUUUAGUUACUA
     |||    :
3' CAAAAUCAAUGAU
```

Delta G: -3.89 kcal/mol Base Pairs: 3
```
5' GUUUUAGUUACUA
     |||    :
3'  CAAAAUCAAUGAU
```

J2Δ14-M2

Delta G: -20.82 kcal/mol Base Pairs: 13
```
5' AAUUUAAUGUUUU
   |||||||||||||
3' UUAAAUUACAAAAUCAAUGAU
```

Delta G: -4.85 kcal/mol Base Pairs: 4
```
5'           AAUUUAAUGUUUU
              ||||  :    :
3' UUAAAUUACAAAAUCAAUGAU
```

Delta G: -3.89 kcal/mol Base Pairs: 3
```
5'  AAUUUAAUGUUUU
    : ::  :    |||
3' UUAAAUUACAAAAUCAAUGAU
```

J2Δ16-M2

Delta G: -23.38 kcal/mol Base Pairs: 15
```
5' AAUUUAAUGUUUUAG
```

```
    ||||||||||||||||
3'  UUAAAUUACAAAAUCAAUGAU
```

Delta G: -4.85 kcal/mol Base Pairs: 4

```
5'            AAUUUAAUGUUUUAG
                ||||  :    :
3'  UUAAAUUACAAAAUCAAUGAU
```

Delta G: -4.85 kcal/mol Base Pairs: 4

```
5'  AAUUUAAUGUUUUAG
                ||||
3'            UUAAAUUACAAAAUCAAUGAU
```

8.2.2 Single Strand Folding Prediction for Single Strand M2 by mFold

Depicted are the three most stable RNA folds as well as their corresponding thermodynamic parameters predicted at 288 K with a fixed ionic strength of 1M NaCl by mFold version 2.3.

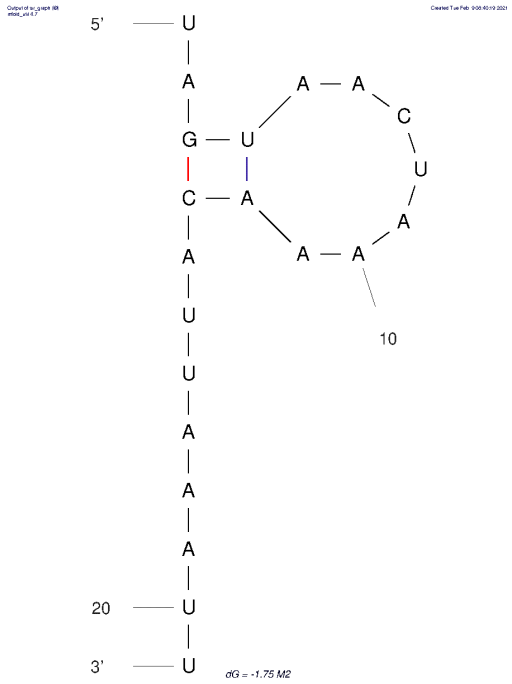

Figure 59. Structure 1 predicted by mFold 2.3 at 288 K and ionic strength of 1 M NaCl with ΔG= -1.75 kcal mol^{-1}, ΔH=-24.80 kcal mol^{-1}, ΔS= -79.99 kcal mol^{-1} K^{-1} and T_m=36.9°C.

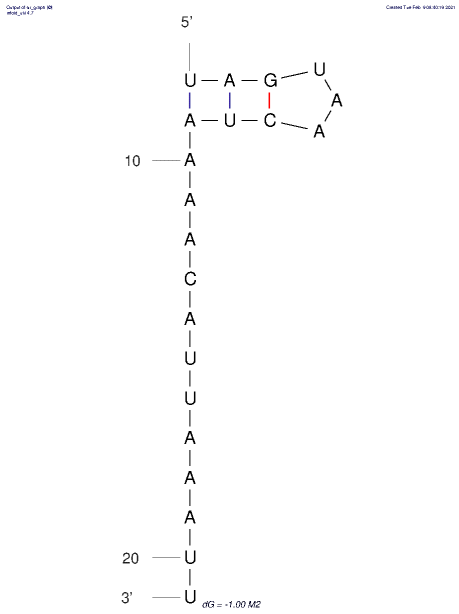

Figure 60. Structure 2 predicted by mFold 2.3 at 288 K and ionic strength of 1 M NaCl with ΔG= -1.00 kcal mol^{-1}, ΔH=-20.60 kcal mol^{-1}, ΔS= -68.02 kcal mol^{-1} K^{-1} and T_m=29.7°C.

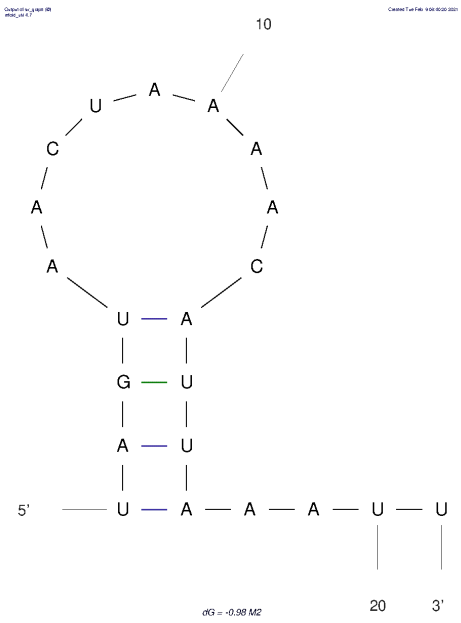

Figure 61. Structure 3 predicted by mFold 2.3 at 288 K and ionic strength of 1 M NaCl with ΔG= -0.98 kcal mol^{-1}, ΔH=-26.90 kcal mol^{-1}, ΔS= -89.95 kcal mol^{-1} K^{-1} and T_m=25.9°C.

8.2.3 Protein Constructs

Table 12. Amino acid sequences of RhlB, RNase E (694-790) and RNase E (628-843) based in wild type sequences from E. coli. Underlined amino acids were introduced with the TEV cleavage site and remained on the target proteins after cleavage.

Construct	Amino acid sequence N→C
RhlB	MSKTHLTEQKFSDFALHPKVVEALEKKGFHNCTPIQALALPLTLAGRDVAGQAQTGT GKTMAFLTSTFHYLLSHPAIADRKVNQPRALIMAPTRELAVQIHADAEPLAEATGLKLG LAYGGDGYDKQLKVLESGVDILIGTTGRLIDYAKQNHINLGAIQVVVLDEADRMYDLG FIKDIRWLFRRMPPANQRLNMLFSATLSYRVRELAFEQMNNAEYIEVEPEQKTGHRIK EELFYPSNEEKMRLLQTLIEEEWPDRAIIFANTKHRCEEIWGHLAADGHRVGLLTGDV AQKKRLRILDEFTRGDLDILVATDVAARGLHIPAVTHVFNYDLPDDCEDYVHRIGRTGR AGASGHSISLACEEYALNLPAIETYIGHSIPVSKYNPDALMTDLPKPLRLTRPRTGNGPR RTGAPRNRRRSGENLYFQ
RNase E (694-790)	GAKALNVEEQSVQETEQEERVRPVQPRRKQRQLNQKVRYEQSVAEEAVVAPVVEET VAAEPIVQEAPAPRTELVKVPLPVVAQTAPEQQEENNADNRD
RNase E (628-843)	GERTEGSDNREENRRNRRQAQQQTAETRESRQQAEVTEKARTADEQQAPRRERSRR RNDDKRQAQQEAKALNVEEQSVQETEQEERVRPVQPRRKQRQLNQKVRYEQSVAE EAVVAPVVEETVAAEPIVQEAPAPRTELVKVPLPVVAQTAPEQQEENNADNRDNGG MPRRSRRSPRHLRVSGQRRRRYRDERYPTQSPMPLTVACASPELASGKVW

Table 13. DNA sequences of RhlB, RNase E (694-790) and RNase E (628-843) constructs including His_6-Tag (blue), TEV-cleavage sites (red) and NdeI (CATATG) or BamHI (GGATCC) restriction sites (orange). When necessary, additional start codon ATG and stop codon TAA were added to the sequences.

Construct	DNA Sequence 5'→3'
RhlB	CATATGATGAGCAAAACCCATCTGACCGAACAGAAATTTAGCGATTTTGCGCTGCAT CCGAAAGTGGTGGAAGCGCTGGAAAAAAAAGGCTTTCATAACTGCACCCCGATTCA GGCGCTGGCGCTGCCGCTGACCCTGGCGGGCCGCGATGTGGCGGGCCAGGCGCAG ACCGGCACCGGCAAAACCATGGCGTTTCTGACCAGCACCTTTCATTATCTGCTGAGC CATCCGGCGATTGCGGATCGCAAAGTGAACCAGCCGCGCGCGCTGATTATGGCGCC GACCCGCGAACTGGCGGTGCAGATTCATGCGGATGCGGAACCGCTGGCGGAAGCG ACCGGCCTGAAACTGGGCCTGGCGTATGGCGGCGATGGCTATGATAAACAGCTGA AAGTGCTGGAAAGCGGCGTGGATATTCTGATTGGCACCACCGGCCGCCTGATTGAT TATGCGAAACAGAACCATATTAACCTGGGCGCGATTCAGGTGGTGGTGCTGGATGA AGCGGATCGCATGTATGATCTGGGCTTTATTAAAGATATTCGCTGGCTGTTTCGCCG CATGCCGCCGGCGAACCAGCGCCTGAACATGCTGTTTAGCGCGACCCTGAGCTATC GCGTGCGCGAACTGGCGTTTGAACAGATGAACAACGCGGAATATATTGAAGTGGA ACCGGAACAGAAAACCGGCCATCGCATTAAAGAAGAACTGTTTTATCCGAGCAACG AAGAAAAAATGCGCCTGCTGCAGACCCTGATTGAAGAAGAATGGCCGGATCGCGC GATTATTTTTGCGAACACCAAACATCGCTGCGAAGAAATTTGGGGCCATCTGGCGG CGGATGGCCATCGCGTGGGCCTGCTGACCGGCGATGTGGCGCAGAAAAAACGCCT GCGCATTCTGGATGAATTTACCCGCGGCGATCTGGATATTCTGGTGGCGACCGATGT GGCGGCGCGCGGCCTGCATATTCCGGCGGTGACCCATGTGTTTAACTATGATCTGCC GGATGATTGCGAAGATTATGTGCATCGCATTGGCCGCACCGGCCGCGCGGGCGCG

	AGCGGCCATAGCATTAGCCTGGCGTGCGAAGAATATGCGCTGAACCTGCCGGCGAT TGAAACCTATATTGGCCATAGCATTCCGGTGAGCAAATATAACCCGGATGCGCTGAT GACCGATCTGCCGAAACCGCTGCGCCTGACCCGCCCGCGCACCGGCAACGGCCCGC GCCGCACCGGCGCGCCGCGCAACCGCCGCCGCAGCGGCGAAAACCTGTATTTTCAA GGCCATCATCATCATCATCATTAAGGATCC
RNase E (694-790)	CATATGCATCATCATCATCATCATGAAAACCTGTATTTTCAAGGCGCGAAAGCGCTG AACGTGGAAGAACAGAGCGTGCAGGAAACCGAACAGGAAGAACGCGTGCGCCCG GTGCAGCCGCGCCGCAAACAGCGCCAGCTGAACCAGAAAGTGCGCTATGAACAGA GCGTGGCGGAAGAAGCGGTGGTGGCGCCGGTGGTGGAAGAAACCGTGGCGGCGG AACCGATTGTGCAGGAAGCGCCGGCGCCGCGCACCGAACTGGTGAAAGTGCCGCT GCCGGTGGTGGCGCAGACCGCGCCGGAACAGCAGGAAGAAAACAACGCGGATAA CCGCGATTAAGGATCC
RNase E (628-843)	CATATGCATCATCATCATCATCATGAAAACCTGTATTTTCAAGGCGAACGCACCGAA GGCAGCGATAACCGCGAAGAAAACCGCCGCAACCGCCGCCAGGCGCAGCAGCAGA CCGCGGAAACCCGCGAAAGCCGCCAGCAGGCGGAAGTGACCGAAAAAGCGCGCAC CGCGGATGAACAGCAGGCGCCGCGCCGCGAACGCAGCCGCCGCCGCAACGATGAT AAACGCCAGGCGCAGCAGGAAGCGAAAGCGCTGAACGTGGAAGAACAGAGCGTG CAGGAAACCGAACAGGAAGAACGCGTGCGCCCGGTGCAGCCGCGCCGCAAACAGC GCCAGCTGAACCAGAAAGTGCGCTATGAACAGAGCGTGGCGGAAGAAGCGGTGGT GGCGCCGGTGGTGGAAGAAACCGTGGCGGCGGAACCGATTGTGCAGGAAGCGCC GGCGCCGCGCACCGAACTGGTGAAAGTGCCGCTGCCGGTGGTGGCGCAGACCGCG CCGGAACAGCAGGAAGAAAACAACGCGGATAACCGCGATAACGGCGGCATGCCGC GCCGCAGCCGCCGCAGCCCGCGCCATCTGCGCGTGAGCGGCCAGCGCCGCCGCCG CTATCGCGATGAACGCTATCCGACCCAGAGCCCGATGCCGCTGACCGTGGCGTGCG CGAGCCCGGAACTGGCGAGCGGCAAAGTGTGGTAAGGATCC

8.2.4 Protein Structure Prediction

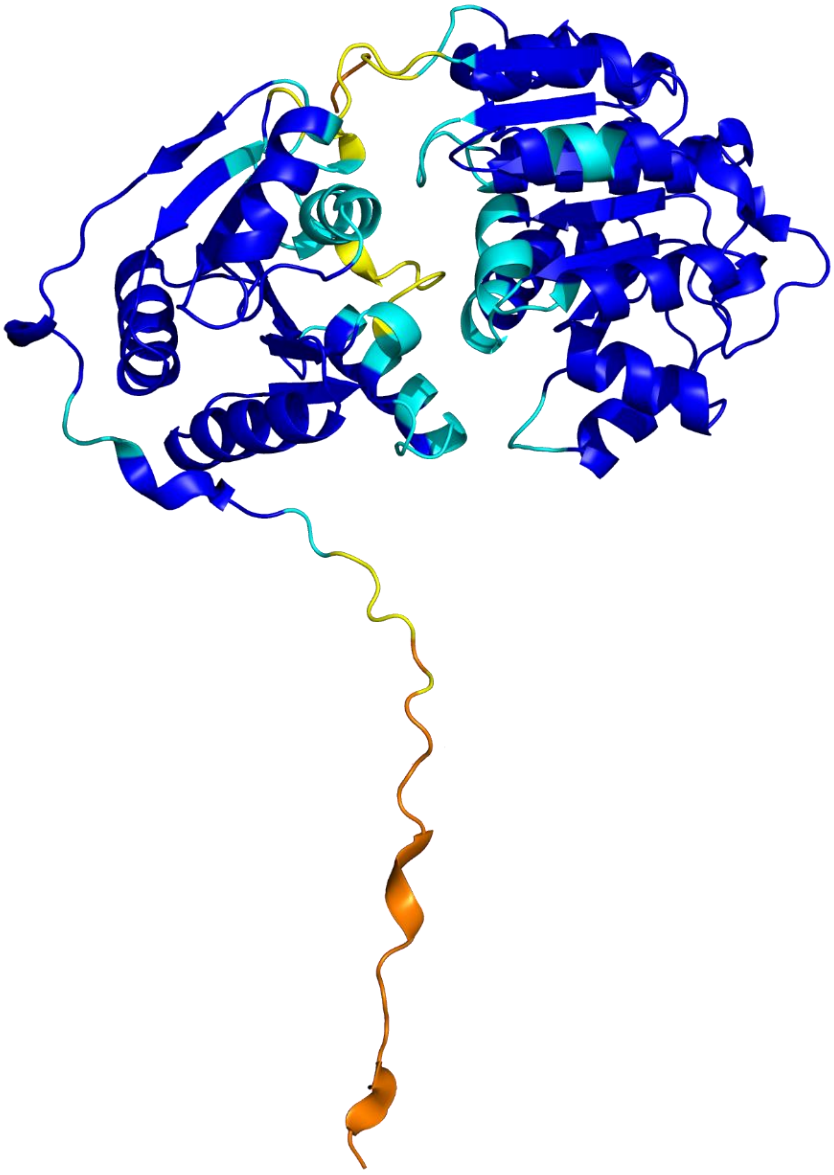

Figure 62. Alphafold structure prediction of Escherichia coli RhlB based on UniProt entry P0A8J8; AlphaFold DB version 2022-11-01, downloaded on 09. May 2023. AlphaFold produces a per-residue confidence score (pLDDT) between 0 and 100. Some regions below 50 pLDDT may be unstructured in isolation. Model confidence: blue: very high (pLDDT>90); cyan: confident (90>pLDDT>70); yellow: low (70>pLDDT>50); orange: very low (pLDDT<50)[110,111].

8.3 Supplementary Data Chapter 3

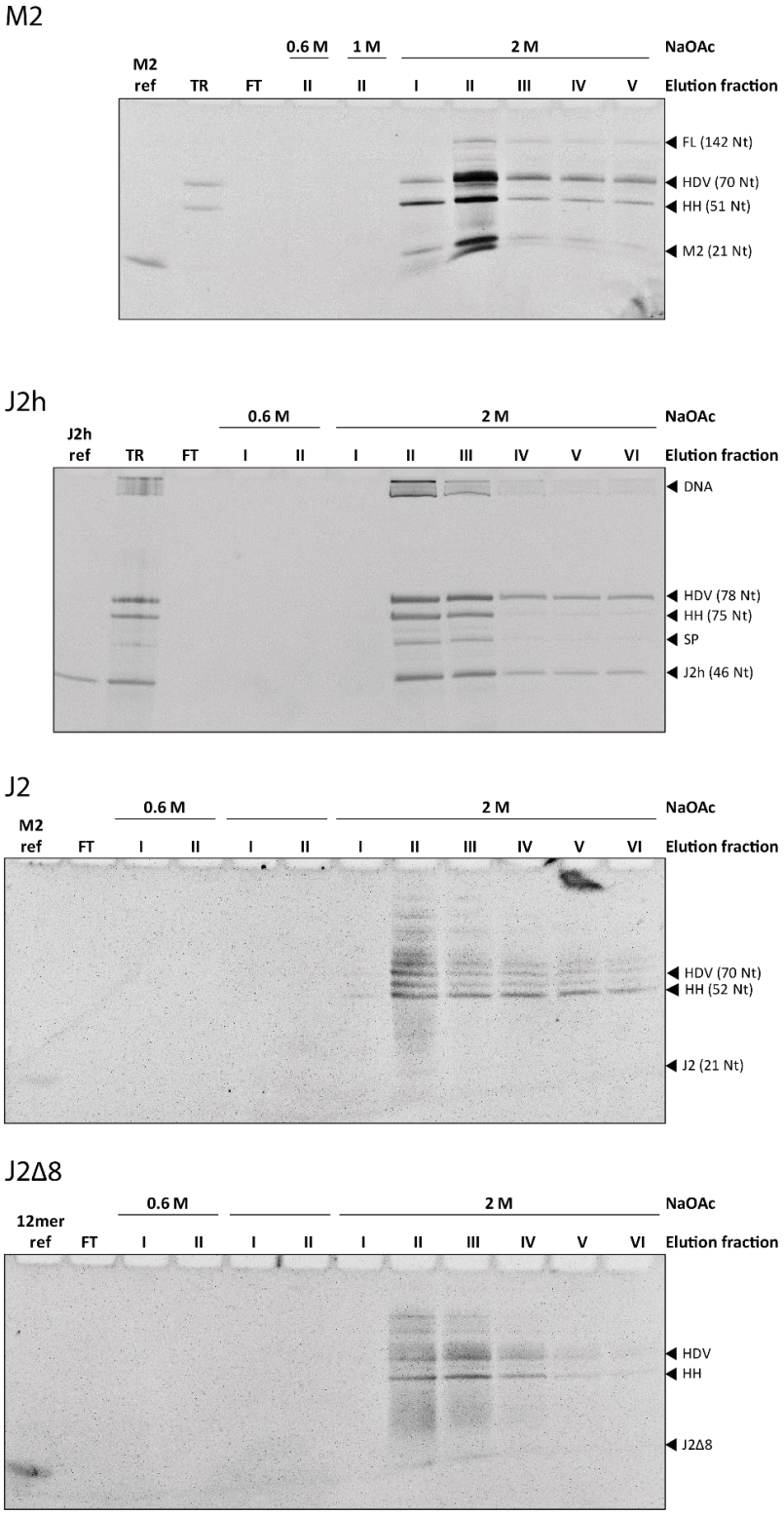

Figure 63. Denaturing RNA-PAGEs of DEAE-Chromatography fractions of synthesized RNA constructs M2, J2h, J2 and J2Δ8. Loaded onto the were RNA size references, the initial column flow through (FT) as well as elution fraction from the elution with 0.6 M, 1 M and 2 M NaOAc elution buffer respectively. Transcription products were annotated with black arrows.

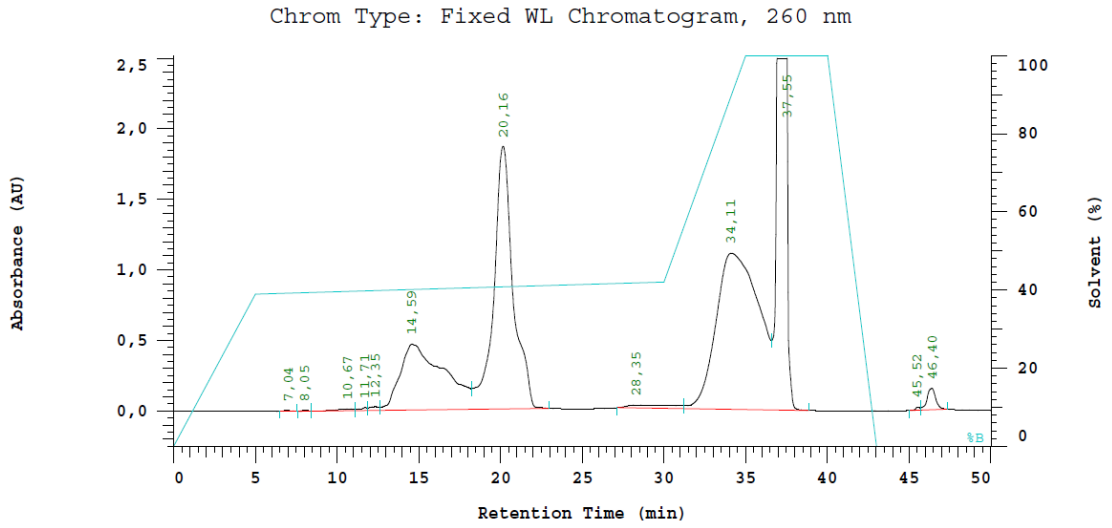

Figure 64. Exemplary reverse-phase HPLC chromatogram of M2 RNA as described in chapter 3.4.3. Absorption at 260 nm is plotted against retention volume. Gradient buffer B is shown in pale blue and significant elution peaks are annotated with retention time at maximum absorption.

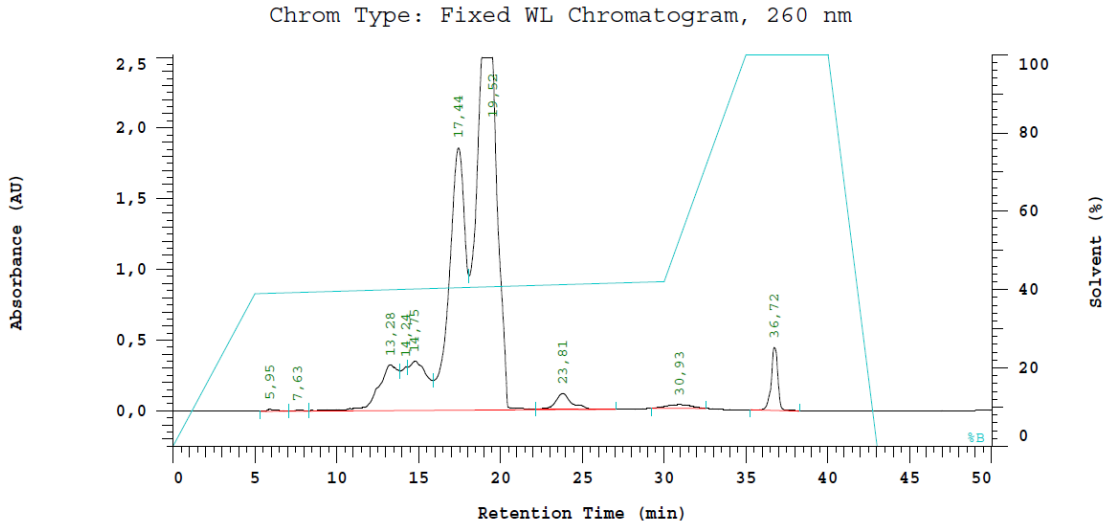

Figure 65. Exemplary HPLC chromatogram of J2h RNA as described in chapter 3.4.3. Absorption at 260 nm is plotted against retention volume. Gradient buffer B is shown in pale blue and significant elution peaks are annotated with retention time at maximum absorption.

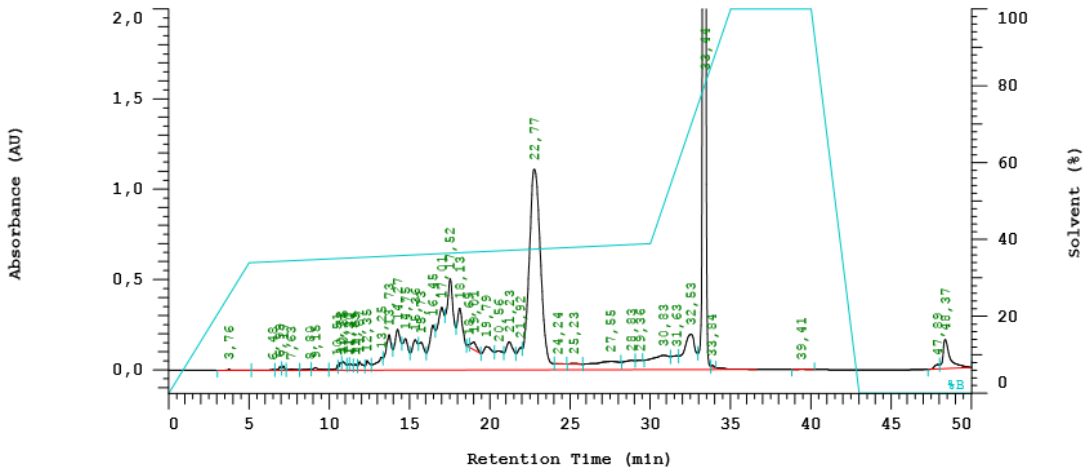

Figure 66. Exemplary HPLC chromatogram of J2 RNA as described in chapter 3.4.3. Absorption at 260 nm is plotted against retention volume. Gradient buffer B is shown in pale blue and significant elution peaks are annotated with retention time at maximum absorption.

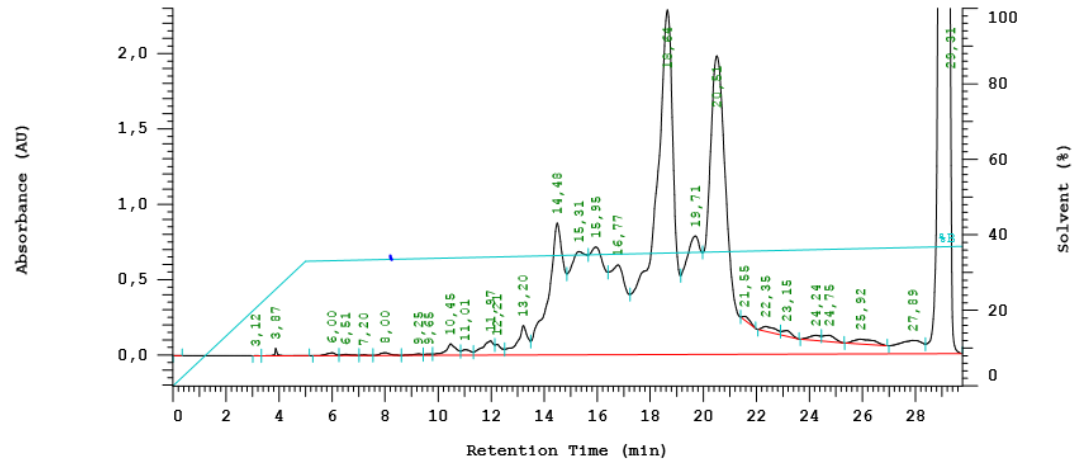

Figure 67. Exemplary HPLC chromatogram of J2Δ8 as described in chapter 3.4.3. Absorption at 260 nm is plotted against retention volume. Gradient buffer B is shown in pale blue and significant elution peaks are annotated with retention time at maximum absorption.

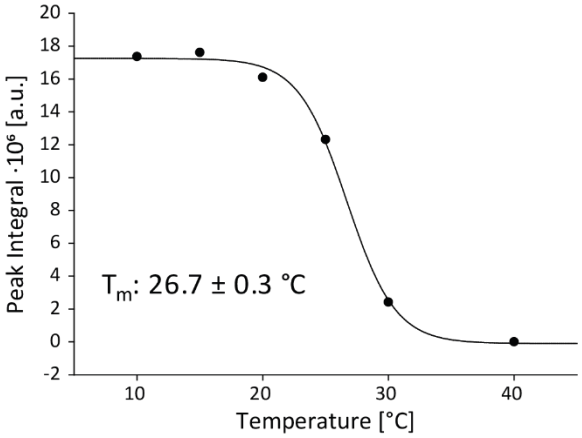

Figure 68. Melting curve of internal hairpin of M2 RNA strand. Peak integral of M2 imino proton G3 were plotted against measurement temperature and T_m of 26.7°C is extracted from the sigmoidal fit.

Table 14. Melting temperatures T_m of RNA duplexes measured via CD spectroscopy under different buffer conditions as described in 4.2.2. The double sigmoidal fit of J2h-M2 provided two melting points.

RNA duplex	Buffer	$MgCl_2$	Measured T_m [°C]	
J2h-M2	pH 6 buffer[a]	-	55.74 ± 0.23	58.67 ± 0.80
		+ 1.5 mM	63.09 ± 0.53	65.20 ± 0.43
J2-M2	pH 8 buffer[b]	-	55.27 ± 0.07	
		+ 4.5 mM	62.80 ± 0.07	
J2Δ8-M3	pH 8 buffer[b]	-	46.71 ± 0.12	
		+ 4.5 mM	53.36 ± 0.004	
J2Δ8-M2	pH 8 buffer[b]	-	53.26 ± 0.19	
		+ 4.5 mM	58.25 ± 0.11	
J2Δ14-M2	pH 8 buffer[b]	-	37.45 ± 0.48	
		+ 4.5 mM	42.55 ± 0.26	
J2Δ16-M2	pH 8 buffer[b]	-	43.47 ± 0.36	
		+ 4.5 mM	50.38 ± 0.07	

[a] 150 mM KCl, 25 mM KH_2PO_4/K_2HPO_4, pH 6.2

[b] 150 mM KCl, 25 mM Tris, pH 8.3

Errors given are the mean standard deviation of the fit.

8.4 Supplementary Data Chapter 4

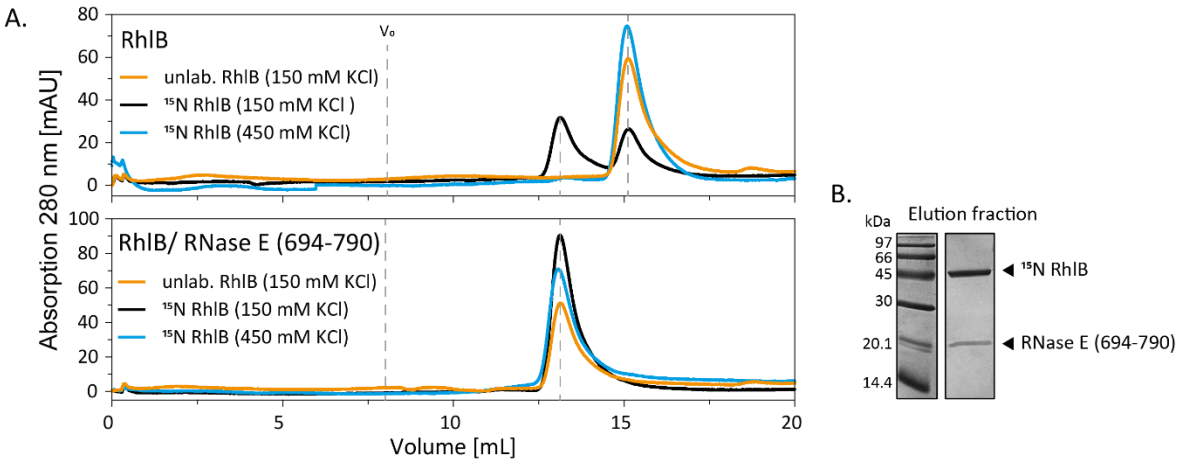

Figure 69. Analytical size exclusion chromatograms of different RhlB batches alone and in complex with RNase E (694-790). A. Upper panel: S200 analytical size exclusion chromatograms of 20 µM unlabelled and ^{15}N labelled RhlB in NMR buffer with 150 mM KCl vs 450 mM KCl. While unlabelled protein shows one elution peak, indicating that it adopts one single stable fold at 150 mM salt concentration, the helicase batch prepared from M9 medium shows a conformational heterogeneity with two peaks in elution profile. The protein conformation can be stabilized with higher salt concentration, resulting in one peak. Lower panel: S200 analytical size exclusion chromatograms of 20 µM unlabelled and ^{15}N labelled RhlB in NMR buffer with 150 mM KCl vs 450 mM KCl in complex with RNase E (694-790). Both proteins are mixed in a 1.1 ratio. Both proteins elute together in one peak, indicating a stable homogenous complex formation for all different RhlB batches. B. SDS-PAGE of elution fraction of ^{15}N RhlB in complex RNase E (694-790) at 150 mM KCl. Despite the same relative elution volume as misfolded RhlB population in top panel, gel confirms coelution of both proteins and therefore complex formation.

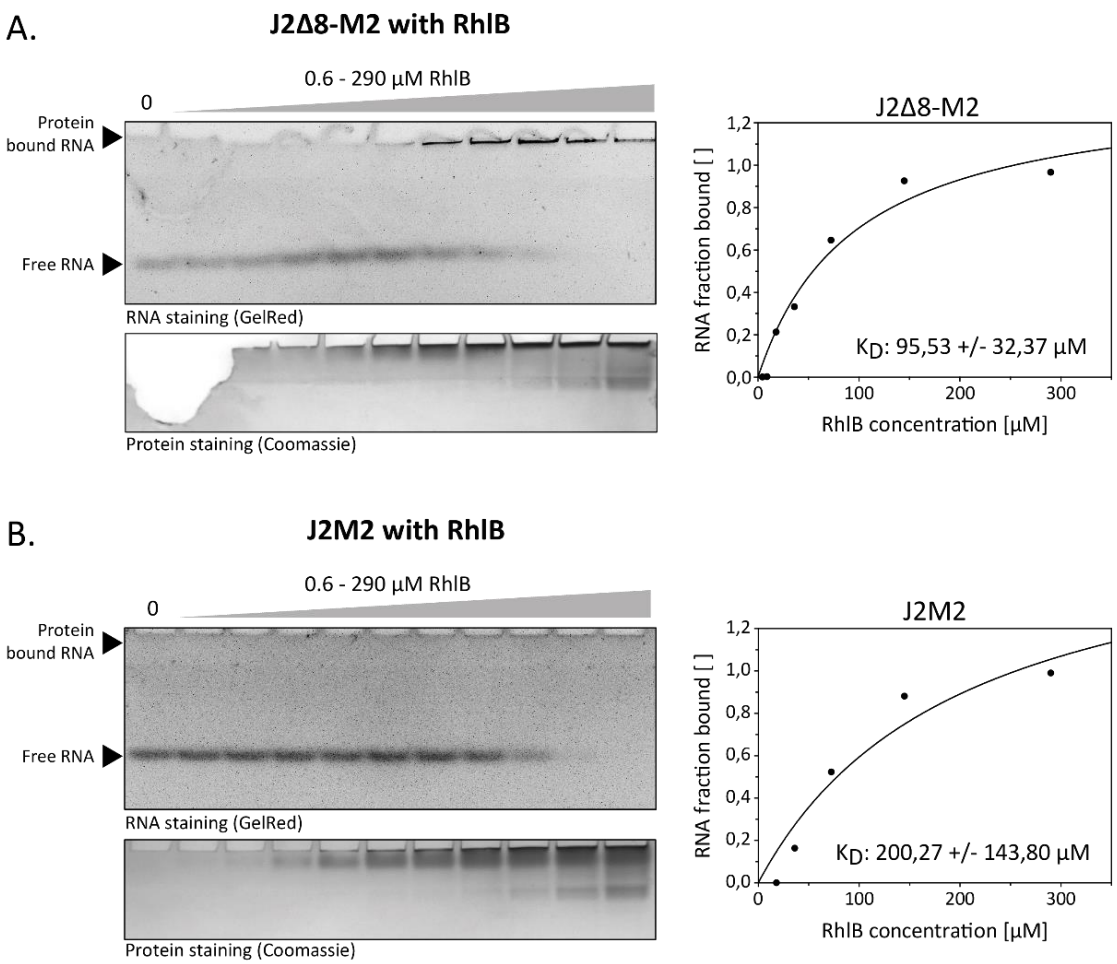

Figure 70. Electrophoretic mobility shift assays (EMSA) of RhlB with J2Δ8-M2 RNA duplex (A) and J2-M2 blunt end duplex (B). Loaded onto the gel are the RNAs alone (0) and with increasing concentrations of RhlB (0.6 – 290 µM). The free RNA is visible as a single band on the gel, while the protein-RNA complex is remaining in the gel pocket as the helicase does not migrate into the gel under the given native PAGE conditions. The gels were stained with GelRed for RNA visualisation (upper gel panel) and Coomassie for protein visualisation (lower gel panel). Quantification of the free RNA with ImageJ was performed as described in 3.6.1. The fraction of RNA in complex determined in gel shift assays are plotted as a function of protein concentration. Curves are best fit to a simple single binding site model. Due to slightly irregular gel staining some datapoints from (A) and (B) were omitted from the analysis to obtain a more accurate curve fit.

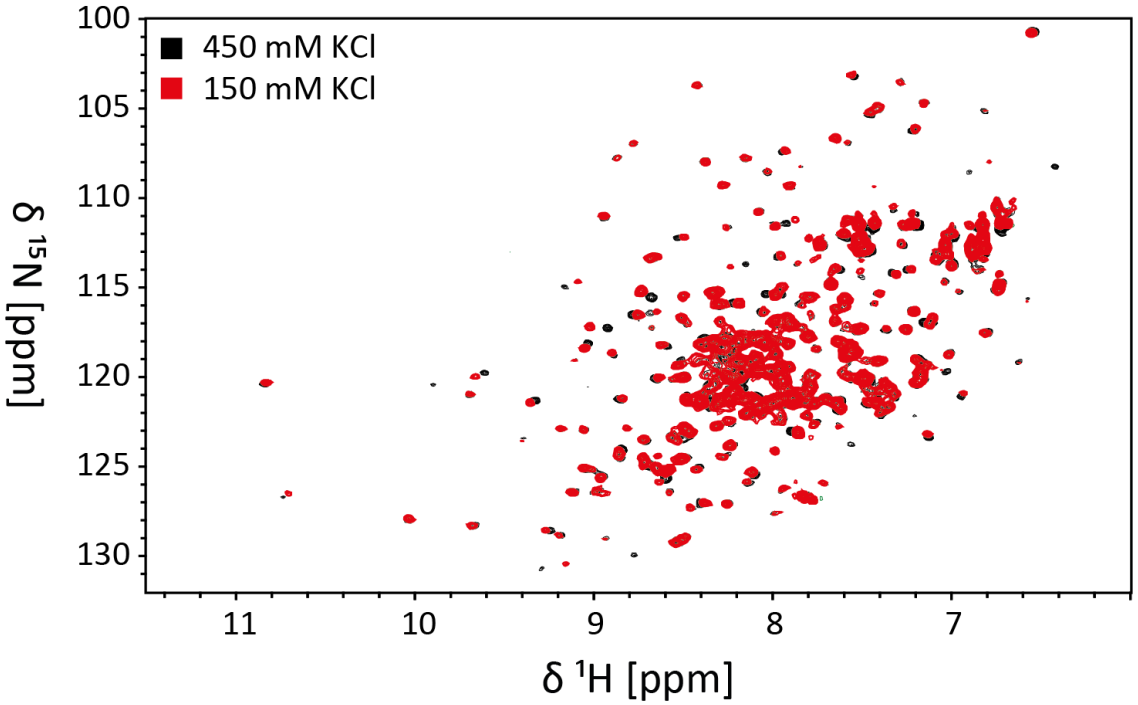

Figure 71. Comparison RhlB/RNase E (694-790) complex (1:1) in NMR buffer with 150 mM KCl and 450 mM KCl. Overlay of ^{1}H, ^{15}N HSQC spectra of 100 µM ^{15}N labelled RhlB in presence and absence of 1 equivalent unlabelled RNase E (694-790). Both spectra were measured with 104 scans in 150 mM or 450 mM KCl, 25 mM Tris, 5 mM DTT, 100 µM DSS and 10% D_2O (pH 8.3) at 305 K and 800 MHz. Resonances were referenced to DSS.

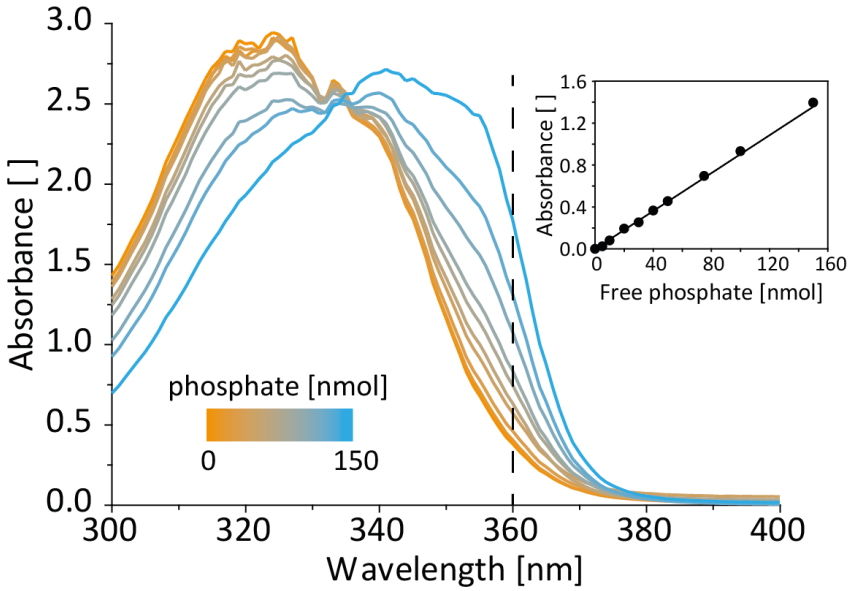

Figure 72. Free phosphate standard curve for Molecular Probes™ EnzChek™ Phosphate Assay. Absorbance of 0 to 30 pmol orthophosphate (0-150 nmol) in reaction buffer (0.2 mM MESG, 1 U/mL PNP, 111 mM KCl, 68.5 mM Tris, 1 mM $MgCl_2$, 0.1 mM sodium azide at pH 8.1) were measured from 300 to 400 nm wavelength. Increased orthophosphate concentrations in the reaction mix result in an increased concentration of 2-amino-6-mercapto-7-methylpurine, which is accompanied by an increase in absorbance at 360 nm. Linear correlation of total orthophosphate against the absorbance at 360 nm provides a phosphate standard curve, that allows quantification of inorganic phosphate (P_i) released during ATP hydrolysis.

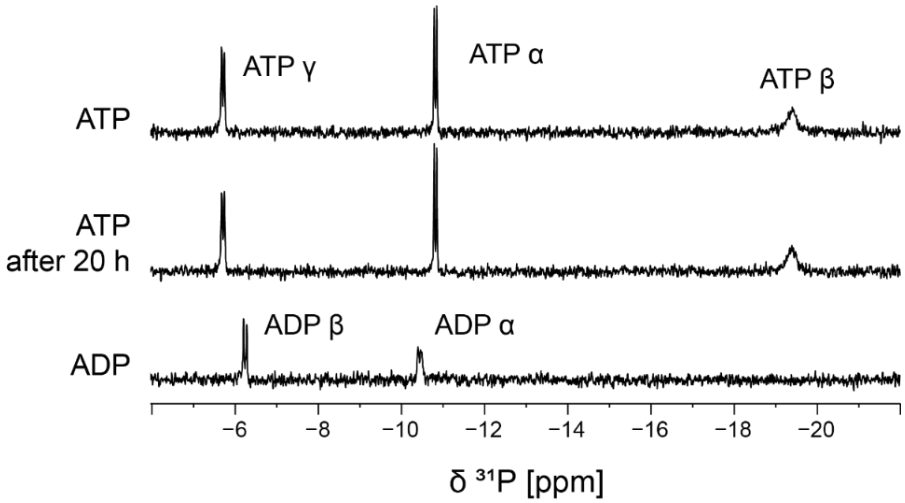

Figure 73. ATP stability test. ^{31}P 1D spectra of 3 mM ATP in NMR buffer directly after preparation and 20 h after incubation at room temperature (295 K) as well as 100 µM ADP in NMR buffer for reference. The spectra clearly show that no spontaneous hydrolysis of ATP to ADP took place under those buffer conditions, since no ADP signals could be detected in the ATP spectrum after 20 h. Spectra were recorded at 288 K and 600 MHz with 265 scans and a spectra width of 50 ppm.

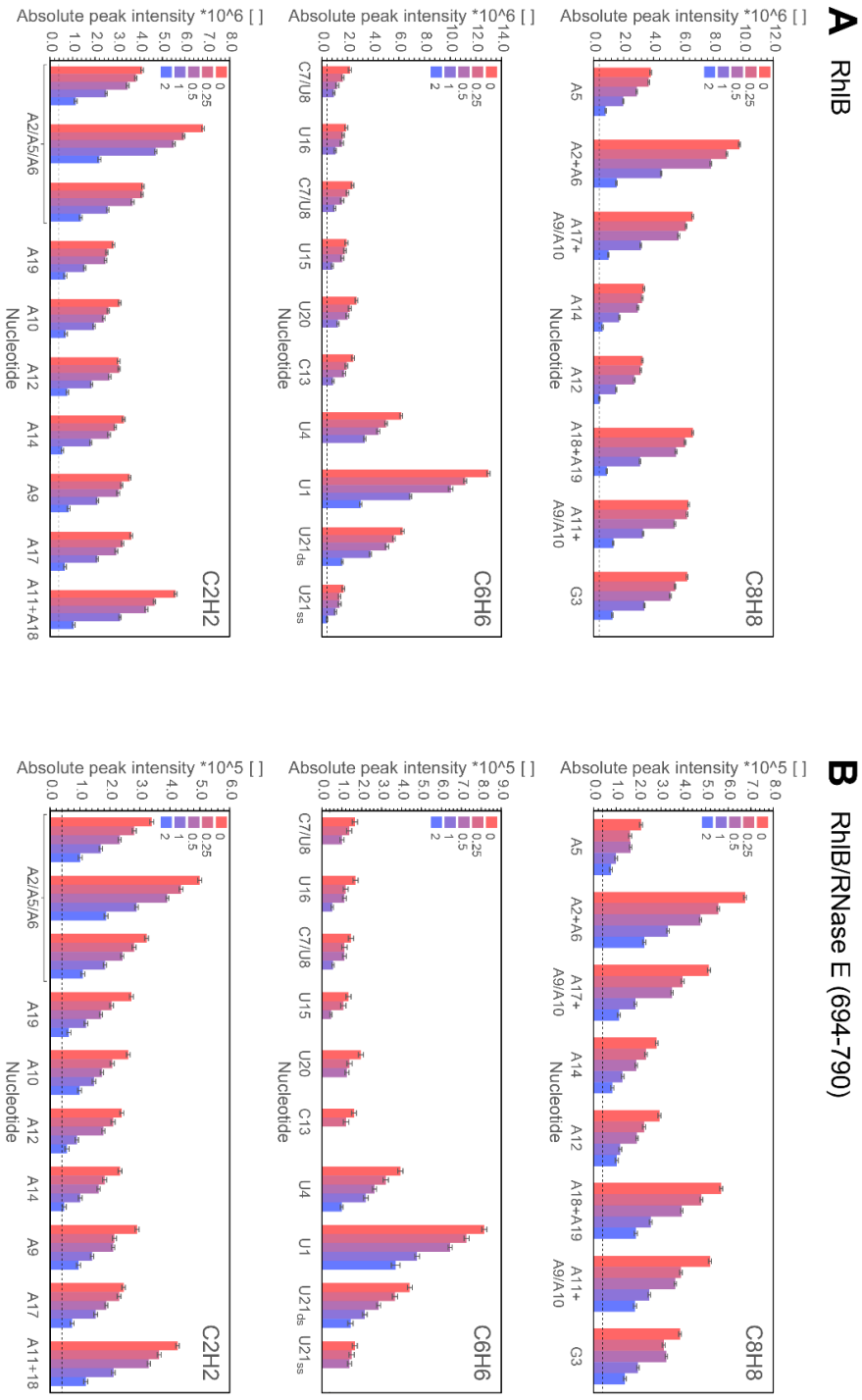

Figure 74. Nucleobase peak intensities of 100 µM 5'-OV RNA during titration with up to 2 equivalents of RhlB or RhlB/RNase E (694-790) complex. ^{13}C HSQC spectra of each titration step recorded with 52 scans at 700 and 800 MHz for RhlB and RhlB/RNase E (694-790), respectively, at 288 K. Spectra were referenced against the chemical shift of DSS. Peak intensities were extracted from Sparky and errors calculated from the S/N ratio of the corresponding spectra. The dashed lines represent the sensitivity threshold, at which the peaks were detected. Resonance assignment is matched with the assignment depicted in Figure 57.

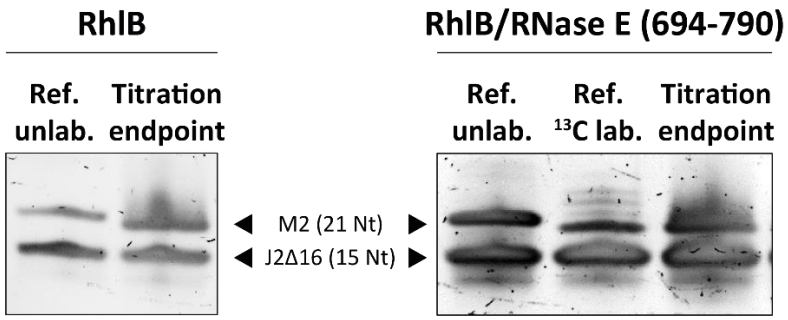

Figure 75. Denaturing RNA PAGE of NMR samples after titration of 4 equivalents RhlB or RhlB/RNase E (694-790) to 5' tailed duplex J2Δ16-M2 (M2 strand ^{13}C labelled) in section 5.6. For reference, unlabelled and ^{13}C labelled RNA sample were loaded in the same gels. The additional molecular weight causes the ^{13}C labelled M2 strand to migrate slightly further in the gel compared to the unlabelled counterpart. Both gels show that the RNA strands were still intact after titration and now degradation could be observed.

Chapter 9 Pulse Programs

Pulse program of *rt_zgig_2D.hz*

```
;$CLASS=HighRes
;$DIM=2D
;$TYPE=
;$SUBTYPE=
;$COMMENT=

prosol relations=<triple>

#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

"d11=30m"
"d12=20u"
"d28=d1-d27-20u"

"l0=0"
"l31=cnst10*ns+1"          ; td2=15n increments, ns=number scans

1 ze
2 d11 do:f2
3 d12 pl12:f2
  ;50u UNBLKGRAD
 3m iu0

 if "l0 == l31"
     {
  10u setnmr3|10
  d27
  10u setnmr3^10
  d28
     }
  else
     {
  d1
     }

  10u pl1:f1
  p1 ph1
  ;4u BLKGRAD

  go=2 ph31 cpd2:f2
  d11 do:f2 mc #0 to 2
     F1QF(zd)
exit

ph1=0 2 2 0 1 3 3 1
ph31=0 2 2 0 1 3 3 1

;cnst10: number of exp before trigger
```

Acknowledgements - Danksagung

An aller erster Stelle möchte ich meinem Doktorvater **Boris Fürtig** danken. Dafür, dass er mir im Rahmen des CliC Graduiertenkollegs als allererste Doktorandin des AK Fürtig die Möglichkeit gegeben hat, diese Arbeit unter seiner Betreuung anzufertigen. Danke Boris für deine Unterstützung bei jeglichen wissenschaftlichen Fragen und Problemen und dein kontinuierliches Feedback. Und vor allem danke ich dir für das unerschöpfliche Vertrauen, die Geduld und deinen unerschütterlichen Optimismus, den du mir und meinem Projekt entgegengebracht hast und ohne den diese Arbeit das jetzige Ziel nicht gefunden hätte.

Ein herzlicher Dank geht an **Harald Schwalbe** für die Möglichkeit, diese Arbeit mit der Ausstattung und der fantastischen Arbeitsatmosphäre des AK Schwalbe durchführen zu dürfen. Danke für die großzügige Förderung, die es mir unter Anderem ermöglicht hat, an so vielen Konferenzen teilnehmen zu können. Vielen Dank für dein konstruktives Mitwirkung an meinem Projekt und für die freundliche Übernahme des Zweitgutachters.

Martin Hengesbach möchte ich für all seinen wertvollen wissenschaftlichen Input und die spannenden Diskussionen danken. Und nicht zuletzt für eine konstruktive, aber auch unschlagbar lustige Atmosphäre im Büro. In diesem Zusammenhang muss und möchte ich auch unbedingt der alten, mittelalten sowie neuen Besetzung des **Taylor-Rooms** danken: **Martin Hengesbach, Isam Elamri, Kathi Hohmann, Tom Landgraf, Vanessa de Jesus, Julius Blechar, Florian Sochor und Florian Buhr**. Wir waren schon mit Abstand das nerdigste Büro. Danke für den großartigen Austausch von Ideen und all den Spaß, den wir über die Jahre hatten. Danke an die beiden Flos, die dafür gesorgt haben, dass das Humor-Niveau stets noch unter der Tür durchpasst.

Ein riesiger Dank geht auch an **Sridhar Sreeramulu, Christinan Richter, Jan Ferner, Julia Wirmer-Bartoschek und Anna Wacker**, dafür dass Sie alle mich in den verschiedensten Bereichen meines Projektes mit ihrem Input unterstützt haben, von NMR-Experimenten bis hin zu Fragen rund um RNA.

Ich danke **Dinh Du Tran** und **Prof. Wachtveitl** für die tolle Kollaboration und Zusammenarbeit rund um das Helikase-Projekt.

Ein unendlich großes Dankeschön geht an **Nusrat Qureshi**. Von der Zeit, als ich für mein erstes Praktikum im AK anfing, über meine Diplomarbeit bis hin zur Promotion hast du mir so unzählig viele Male geholfen und mir rund ums Labor und das Spektrometer so viel beigebracht. Ich danke dir für deine bedingungslose gute Laune, deinen wundervollen zynischen Humor und dass du so viel von deinem Wissen mit mir geteilt hast. Und dafür, dass du im Notfall immer irgendwo einen Sekt oder Baileys zur Hand hattest.

Unbedingt danken möchte ich außerdem **Elke Stirnal** und **Kerstin Witt**, nicht nur für die HPLC-Aufreinigungen meiner vielen RNA-Konstrukte, sondern auch dafür, dass Sie wie zwei Felsen in der Brandung immer die Versorgung der Labore mit jeglichem Equipment sichergestellt habt.

Anna Paulus, Daniela Dworak und besonders **Kerstin Dathe** möchte ich für all die organisatorische Unterstützung im AK danken. Kerstin, ich danke dir dafür, dass du mir einfach zu jeder Tages- und Nachtzeit eine unerschöpfliche Positivität und Hilfsbereitschaft entgegengebracht hast und beim Projekt jedes einzelnen Doktoranden so mitfieberst, als wäre es dein eigenes.

Kerstin Witt, Denis Kudlinzki und **Verena Linhard** danke ich für all die Hilfe mit den Kristallisations-Experimenten. Auch wenn es am Ende leider nicht funktioniert hat, habe ich dank euch sehr viel Neues gelernt und hatte einen großartigen Trip ans BESSY nach Berlin.

Ich möchte mich bei all meinen AK Schwalbe CliC-Teamkollegen der ersten Kohorte **György Pinter, Linda Schulte, Sara Keyhani** und **Isam Elamri** aber auch bei allen anderen „CliCies" bedanken für den wunderbaren wissenschaftlichen und interdisziplinären Austausch und die gegenseitige Unterstützung. **Patrick Seyfried** danke ich für die spannende gemeinsame Zeit im Steering Board.

Ein riesengroßes Dankeschön geht an **Martin Hähnke** für den großartigen Support bei allen großen und kleinen IT-Problemen, für den anderswo ein ganzes Team zuständig gewesen wäre. Danke für den ganzen Spaß am 3D Drucker und dass du immer für jeglichen Bastel-Schabernack zu haben bist. Und natürlich für die vielen fantastischen Pizzaabende.

Ein großer Dank geht auch an die Besetzung des **RNA-Labors** über die Jahre: es sind zu viele, um sie alle aufzuzählen, immerhin waren wir im Labor dicht gepackt, aber ein besonderer Dank geht an **Nusrat, Feli, Kathi, Tassilo, Flo Buhr und Flo Sochor**. Danke für die fabelhafte Arbeitsatmosphäre zu jeder Tag- und Nachtzeit und die noch besseren Musik-Playlists.

Ich möchte mich auch bei **Kathi Hohmann** bedanken. Es war eine riesige Freude, mit dir zusammen den AK Fürtig zu begründen und auf dem Helikase Projekt zusammenzuarbeiten. Deine positive Art ist immer ein Sonnenschein im Büro gewesen.

Besonderer Dank geht auch an alle meine Bachelor- und Master-Studenten, die ich das Vergnügen hatte, betreuen zu dürfen: **Clara Roh, Florian Walter, Guanzhu Götze, Kathi Hohmann, Laura Raschke, Lynn Buyachuian, Tobi Matzel**. Ihr alle habt tolle Arbeit geleistet und besonderen Dank an Tobi und Laura, die noch weiter als HiWis so viele wertvolle Beiträge zu meinem Projekt geleistet haben. Ihr wart echt die besten HiWis, die man sich hätte wünschen können.
Danke an meine beiden Schüler-Praktikanten **Konstantin Völkert** und **Laura Almena-Rodriguez.** Konstantin danke ich dafür, dass er mich mit seiner unbedarften Neugier immer mal wieder daran erinnert hat, wie cool Wissenschaft doch eigentlich ist.

Und da ich bestimmt sonst jemanden vergessen würde, möchte ich von Herzen dem **kompletten Arbeitskreis** für den einmaligen Zusammenhalt danken. Ihr alle sorgt dafür, dass die kleinen und großen Erfolge jedes einzelnen gleich doppelt so schön sind und ihr

sorgt auch dafür, dass schwere Tage gleich viel leichter zu ertragen sind. Danke für unzählige tolle Erinnerungen an 11Uhr-Mittagsessen-Gruppen, Keeler Clubs, Pizza Abende, Hut-Bastel-Aktionen, Group Trips, Cocktail-Abende, Weihnachtsfeiern, Geburtstagskuchen, Konferenzen und alles dazwischen.

Darüber hinaus möchte allen meinen lieben Freunden in und um den **Gaming Monday** von Herzen danken. Danke an **Marcus Erdt**, **Miriam Emde**, **Julius Vinnemann**, **Flo Buhr**, **Max Riepl**, **Feli Kutz**, **Daniel Hymon**, **Martin Hähnke**, **Jan Heering**, **Marcus Liebermann**, **Mathis Müller-Ohldach**, **Beata Berlin**, **Korbi Obermann** und **Albrecht Völklein**. Für die schöne und unterhaltsame Zeit jeden Montagabend und den willkommenen Ausgleich zum Laboralltag.

Ein ganz besonderer Dank geht an **Tine Molenda** und **Maria Grötzinger**. Danke, dass Ihr bei unseren gemeinsamen „Zoom-Focus-Sessions“ mein Durchhaltevermögen angespornt habt. Ohne euch und eure Unterstützung wäre diese Arbeit nicht zustande gekommen.

Zuletzt möchte ich meiner **Familie** danken. Ihr habt immer mit mir mitgefiebert oder mich wieder aufgebaut, wenn es mal nicht so gut lief. Danke an meinen Opa, den stolz zu machen mich nochmal extra angespornt hat.

Und von allen Menschen gilt mein größter Dank an **Marcus Erdt**. Danke für dein Vertrauen in mich und dass du mir die Kraft gegeben hast, das Ganze durchzuziehen. Ich danke dir für deine unendliche Geduld und widme dir wie versprochen den Punkt meines „Dr.“.

Eidesstattliche Erklärung und Versicherung

Ich erkläre hiermit, dass ich mich bisher keiner Doktorprüfung im Mathematisch-Naturwissenschaftlichen Bereich unterzogen habe.

Frankfurt am Main, den

Heidi Zetzsche

Ich erkläre hiermit, dass ich die vorgelegte Dissertation über

Biochemical and NMR-Spectroscopic Investigation of the RNA-Unwinding Mechanism of E. coli DEAD-Box Helicase RhlB

selbständig angefertigt und mich anderer Hilfsmittel als der in ihr angegebenen nicht bedient habe, insbesondere, dass alle Entlehnungen aus anderen Schriften mit Angabe der betreffenden Schrift gekennzeichnet sind.
Ich versichere, die Grundsätze der guten wissenschaftlichen Praxis beachtet, und nicht die Hilfe einer kommerziellen Promotionsvermittlung in Anspruch genommen zu haben.

Frankfurt am Main, den

Heidi Zetzsche

Publications

Fürtig, B.; Schnieders, R.; Richter, C.; **Zetzsche, H**.; Keyhani, S.; Helmling, C.; Kovacs, H.; Schwalbe, H. Direct 13 C-Detected NMR Experiments for Mapping and Characterization of Hydrogen Bonds in RNA. *J. Biomol. NMR* **2016**, 207–221. https://doi.org/10.1007/s10858-016-0021-5.

Bains, J.*; Blechar, J.*; de Jesus, V.*; Meiser, N.*; **Zetzsche, H**.; Fürtig, B.; Schwalbe, H.; Hengesbach, M. Combined SmFRET and NMR Analysis of Riboswitch Structural Dynamics. *Methods* **2019**, *153*, 22–34. https://doi.org/10.1016/j.ymeth.2018.10.004.

Fürtig, B.; Oberhauser, E.; **Zetzsche, H**.; Klötzner, D.-P.; Heckel, A.; Schwalbe, H. Refolding through a Linear Transition State Enables Fast Temperature Adaptation of a Translational Riboswitch. *Biochemistry* **2020**, *59* (10), 1081–1086. https://doi.org/10.1021/acs.biochem.9b01044.

Schnieders, R.*; Knezic, B.*; **Zetzsche, H.***; Sudakov, A.; Matzel, T.; Richter, C.; Hengesbach, M.; Schwalbe, H.; Fürtig, B. NMR Spectroscopy of Large Functional RNAs: From Sample Preparation to Low-Gamma Detection. *Curr. Protoc. Nucleic Acid Chem.* **2020**, *82* (1). https://doi.org/10.1002/cpnc.116.

Wacker, A.; Weigand, J.; Akabayov, S.; Altincekic, N.; Kaur Bains, J.; Banijamali, B.; Binas, O.; Castillo-Martinez, J.; Cetiner, E.; Ceylan, B.; Chiu, L.-Y.; Davila-Calderon, J Dhamotharan, K.; Duchardt-Ferner, E.; Ferner, J.; Frydman, L.; Fürtig, B.; Gallego, J.; Grün, J.; Hacker, C.; Haddad, C.; Hähnke, M.; Hengesbach, M.; Hiller, F.; Hohmann, K.; Hymon, D.; de Jesus, V.; Jonker, H.; Keller, H.; Knezic, B.; Landgraf, T.; Löhr, F.; Luo, L.; Mertinkus, K.; Muhs, C.; Novakovic, M.; Oxenfarth, A.; Palomino-Schatzlein, M.; Petzold, K.; Peter, S.; Pyper, D.; Qureshi, N.; Riad, M.; Richter, C.; Saxena, K.; Schamber, T.; Scherf, T.; Schlagnitweit, J.; Schlundt, A.; Schnieders, R.; Schwalbe, H.; Simba-Lahuasi, A.; Sreeramulu, S.; Stirnal, E.; Sudakov, A.; Tants, J.-N.; Tolbert, B.; Vogele, J.; Weiß, L.; Wirmer-Bartoschek, J.; Wirtz Martin, M.; Wöhnert, J.; **Zetzsche, H**. Secondary Structure Determination of Conserved SARS-CoV-2 RNA Elements by NMR Spectroscopy. *Nucleic Acids Res.* **2020**. https://doi.org/10.1093/nar/gkaa1013.

Zetzsche, H.; Raschke, L.; Fürtig, B. Allosteric activation of RhlB by RNase E induces partial duplex opening in substrate RNA. *Front. Mol. Biosci*. **2023**, *Submitted.*

* Shared first authorship

Conference Contributions

Zetzsche, H.; Du Tran, D.; Wachtveitl, J.; Fürtig, B. A new RNA model system for investigation of helicases unwinding mechanism via NMR spectroscopy,
Poster Presentation, VII. Nucleinsäurechemie-Treffen, Berlin (Germany), Sept. 2015

Zetzsche, H.; Fürtig, B. A new RNA model system for investigation of helicases unwinding mechanism via NMR spectroscopy,
Poster Presentation, GDCH Fachgruppentagung Magnetische Resonanz, Frankfurt (Germany), Sept. 2015

Zetzsche, H.; Du Tran, D.; Wachtveitl, J.; Fürtig, B. Investigation of RNA helicase unwinding mechanism via NMR and fluorescence spectroscopy,
Poster Presentation, EMBO Workshop – RNA structure meets function, Stockholm (Sweden), Jun. 2016

Zetzsche, H.; Du Tran, D.; Wachtveitl, J.; Fürtig, B. Investigation of RNA helicase unwinding mechanism via NMR and fluorescence spectroscopy,
Poster Presentation, Design and Light Control conference, Niedernberg (Germany), Aug. 2017

Zetzsche, H.; Fürtig, B. Investigating the RNA unwinding mechanism of DEAD-Box helicase RhlB with NMR Spectroscopy,
Poster Presentation, Experimental Nuclear Magnetic Resonance Conference, Pacific Grove, California (USA), Apr.2019

Zetzsche, H.; Fürtig, B. Investigating the RNA unwinding mechanism of DEAD-Box helicase RhlB with NMR Spectroscopy,
Poster Presentation, EUROISMAR, Berlin (Germany), Aug. 2019